Dissecting host-pathogen interactions of *Clostridioides difficile*

By

Lucy R. Frost

A thesis presented for the degree of Doctor of Philosophy in Biological Sciences

School of Life Sciences, University of Warwick

March 2021

Table of Contents

List of Figures

List of Tables

Acknowledgements

It is with immense gratitude that I acknowledge the help and support of my supervisor, Dr Meera Unnikrishnan. I am particularly grateful for her patience and kindness towards me throughout the PhD project, as well as her enthusiasm about science which left me feeling motivated, no matter how overwhelmed or frustrated I felt prior to our conversations. Meera has also taught me the value of perseverance and resilience, without which I would have never finished this PhD.

It is also a pleasure to give a special thanks my second supervisor, Chrystala Constantinidou, and Richard Stark for their help with bioinformatics and interpretation of my dual RNA-seq data. Their technical expertise and patience were vital for achieving reliable results from large-scale, complex datasets. I would also like to thank Dr Ludmila Rodrigues Pinto Ferreira from The Federal University of Minas Gerais, Brazil for her excellent canonical pathway analysis with my human RNA-seq data. Ludmila very generously took the time to show me exactly how the pathway analysis was conducted to facilitate my learning and understanding.

I am also very grateful for the kindness and patience of Dr Blessing Anonye, a former postdoctoral researcher who trained me in laboratory techniques during the early stages of my PhD. Finally, I would like to thank everyone in the M116 laboratory and the Microbiology and Infection Unit of Warwick Medical School for always keeping my spirits high and making the past 4 and a half years a much more enjoyable and memorable experience.

Author's declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for another Degree.

The work presented (including data generated and data analysis) was carried out by myself except where duly acknowledged in the text and legends.

Signed: 

Date: 17-03-2021

Abstract

C. difficile is the most common cause of hospital-acquired diarrhoea and is a major global health concern. *C. difficile* is a particularly difficult infection to diagnose and treat, demonstrating the urgency for new strategies against the disease. Bacterial interactions with the gut mucosa are crucial for colonisation and establishment of infection, yet key events including bacterial attachment and gut penetration are poorly defined. The aim of this project was to investigate the transcriptomic responses of *C. difficile* infection to identify genes involved in host-*C. difficile* interactions. An *in vitro* human gut model was used to facilitate *C. difficile* infection of human intestinal epithelial cells. A dual RNA-seq approach resolved the transcriptomic profiles of both the host cells and bacteria simultaneously during infection. Several timepoints captured the temporal changes in gene expression over the course of infection.

Many interesting differentially expressed genes and pathways were identified which may have important roles in the pathogenesis of *C. difficile* infection. I demonstrated a potential role for the secreted protease, PPEP-1, in *C. difficile* adhesion to epithelial cells. I also demonstrated that the human mucosal immunity protein, MUC13, was highly upregulated during infection which may have a role in the regulation of inflammatory cytokines in response to *C. difficile*. Interestingly, modulation of purine and pyrimidine ribonucleotide pathways was observed in both the human and bacterial transcriptomic analyses, demonstrating a potential role for inter-kingdom nucleotide signalling during *C. difficile* infection. The identification of genes and pathways modulated during infection could greatly improve our understanding of *C. difficile* colonisation and

contribute towards the development of novel diagnostic markers, as well as new strategies for the treatment of the disease.

Abbreviations

AMPK - 5′-AMP-activated protein kinase

BLAST - Basic Local Alignment Search Tool

c-di-GMP - cyclic diguanylate (3=,5=-cyclic diguanylic acid)

CCL - C-C motif chemokine ligand

CDI – *C. difficile* infection

CDK - cyclin-dependent kinase

CDT – *C. difficile* transferase toxin

CROPs - carbohydrate-binding combined repetitive oligopeptides

CSPG4 - chondroitin sulphate proteoglycan 4

Ct - cycle threshold

CV – crystal violet

CWP – cell wall protein

CXCL - C-X-C motif ligand

DDS - dextran sulfate sodium

DEGs – differentially expressed genes

eATP – extracellular ATP

ECM – extracellular matrix

EGF - epidermal growth factor

EIAs - enzyme immunoassays

ENHO - energy homeostasis associated gene

ERK - extracellular signal-regulated kinase

F-actin - filamentous actin

FACS – fluorescence activated cell-sorting

FISH - fluorescent in situ hybridization

FU – fluorescent units

FXR - farnesoid X receptor

FZDs - Wnt Frizzled receptor

G-actin – globular actin

GAPDH - glyceraldehyde 3-phosphate dehydrogenase

GO – gene ontology

HGF - hepatocyte growth factor

IBD – inflammatory bowel disease

IFN-γ - interferon gamma

Ig – immunoglobulin

IL – interleukin

ILCs - innate lymphoid cells

IκBα - inhibitor of nuclear factor kappa B, alpha

JAK-STAT - Janus Kinase/Signal Transducer and Activator of Transcription

KEGG - Kyoto Encyclopaedia of Genes and Genomes

LPS – lipopolysaccharide

LRP - low-density lipoprotein receptor

LSR - lipolysis-stimulated lipoprotein receptor

LXR - liver X receptor

MAPK - mitogen-activated protein kinase

MIF - macrophage migration inhibitory factor

MMPs - matrix metalloproteinases

mRNA – messenger RNA

NAATs - Nucleic acid amplification tests

NF-κB - nuclear factor-κB

Nod - nucleotide-binding oligomerization domain

PAMPs - pathogen-associated molecular patterns

PCA – principal component analysis

PET - polyethylene terephthalate

PKC - protein kinase C

PRRs - pattern recognition receptors

PVRL3 - Poliovirus receptor-like 3 / NECTIN3

RNA-seq – RNA sequencing

ROS - reactive oxygen species

rRNA – ribosomal RNA

RT – room temperature

RXR - retinoid X receptor

SD – standard deviation

sGAGs - sulfated glycosaminoglycans

SLP – surface layer protein

ssDNA – single-stranded DNA

sST-2 - tumorigenicity 2 receptor

T3SS - type III secretion system

TEER - trans-epithelial electrical resistance

TLR – Toll-like receptor

TNF – tumor necrosis factor

TRAIL - TNF-related apoptosis-inducing ligand

TWEAK - TNF-like weak inducer of apoptosis

UPEC - uropathogenic *Escherichia coli*

VDC - vertical diffusion chamber

Chapter 1: Introduction and Literature Review

1.1 *Clostridioides difficile* infection

Clostridioides difficile is a Gram-positive, anaerobic, spore forming bacterium and one of the most frequently reported nosocomial intestinal pathogens (Rupnik, Wilcox and Gerding, 2009). *C. difficile* is an opportunistic pathogen capable of causing a toxin-mediated disease in humans and animals. Clinical manifestations of *C. difficile* infection (CDI) can range from mild diarrhoea to pseudomembranous colitis and toxic megacolon, which are often life-threatening. In the UK 12,275 cases were reported in England between April 2018 and March 2019 (Public Health England report, July 2019). However, in the USA the number of cases is currently very high, with approximately half a million cases being reported each year, and an estimated 14,000 which result in death. The economic burden of CDI is estimated to set back the US health systems up to $4.8 billion each year (McGlone *et al.*, 2012). *C. difficile* is a particularly difficult infection to treat, and approximately 1 in 5 of patients with hospital-acquired CDI suffer from a recurrence of the infection despite receiving appropriate therapy (Lessa *et al.*, 2015). This is attributed to the ability of *C. difficile* to form metabolically inactive spores, which are associated with ineffective treatment (Rupnik, Wilcox and Gerding, 2009). The Centers for Disease Control and Prevention has recently highlighted *C. difficile* as an urgent antibiotic-resistant threat and a major public health concern.

C. difficile was first described by Hall and O'Toole in 1935 in a study investigating microbiome changes of healthy infants. The bacterium was newly identified as a strict anaerobe, with subterminal, non-bulging, elongated spores. It was particularly difficult to isolate and study, leading it to originally be named *Bacillus difficilis* (Hall, Ivan and O'Toole, 1935). A later study in 1937 found that *B. difficilis* toxins caused edema in subcutaneous tissues of guinea pigs, rabbits, cats, dogs, rats and pigeons and also lead to respiratory-related deaths (Snyder, 1937). In 1938, *B. difficilis* was reclassified into the Clostridium genus and renamed *Clostridium difficile* (Prevot, 1938). In the 1960s it was noted that the majority of pseudomembranous colitis cases occurred after the use of antibiotics. It was not until the 1970s that *C. difficile* was isolated from a number of hospital patients, and in 1978 *C. difficile* was identified as a cause for diarrhoea and pseudomembranous colitis (Gorbach and Thadepalli, 1975; Bartlett *et al.*, 1978; Larson *et al.*, 1978). In late 1978, vancomycin was recognised as a treatment to eliminate *C. difficile* from the colon and rapidly improve the clinical manifestations of pseudomembranous colitis (Keighley *et al.*, 1978). Since then, the number of cases of hospital-acquired infections has increased and *C. difficile* was described as the pathogen of the 90s after a number of outbreaks worldwide (Riley, 1998). The early 2000s saw a rise in the incidence, severity and mortality rate associated with CDI in many developed countries across Europe and North America. This exacerbated disease was associated with the emergence of a novel 'hypervirulent' strain, PCR-ribotype 027 (Kuijper *et al.*, 2008). *C. difficile* is now a global public health concern and the most common cause for antibiotic-associated diarrhoea in healthcare settings (Rodriguez *et al.*, 2016).

There are many factors associated with CDI, the main pharmacological risk factor is being on broad-spectrum antibiotics such as ampicillin, amoxicillin, cephalosporins, clindamycin and fluoroquinolones, which are frequently associated with the infection (Freeman and Wilcox, 1999). The highest rates of infection and mortality are in elderly patients undergoing broad-spectrum antibiotic treatment, with over 80% of deaths occurring in patients over 65 years old (Vardakas *et al.*, 2012; Leffler and Lamont, 2015). Spending a period of time in a healthcare environment such as a hospital or care home is also a major CDI risk factor (Eze *et al.*, 2017, Pepin *et al.*, 2005). Whilst these risk factors are frequently seen in CDI patients, a rise in community-acquired CDI has been observed, where up to 50% of patients do not have the typical risk factors (Dial *et al.*, 2008).

The gut microbiota provides colonisation resistance against *C. difficile* through several mechanisms, including depletion of the levels of available nutrients, prevention of access to mucosal adherence sites and the production of inhibitory compounds, such as secondary bile acids (Rosa, Donskey and Munoz-Price, 2018). Antibiotic treatments lead to an alteration in the species composition of the gut microbiota which can result in susceptibility to CDI. In addition to antibiotics, there are many other causes of dysbiosis of the microbiota which have also been linked to CDI, such as gastrointestinal disturbances (Vaninsberghe *et al.*, 2020). Specifically, a microbiota composition which increases in the availability of amino acids, particularly proline which is utilised by *C. difficile* in Stickland metabolism, may increase susceptibility to *C. difficile* infection (Jackson *et al.*, 2006; Bouillaut, Self and Sonenshein, 2013; Battaglioli *et*

al., 2018). Recently, other dietary components have been reported to impact *C. difficile* infection. Dietary zinc has been associated with an altered microbiota and an exacerbation in *C. difficile* infection severity, however, the mechanism by which zinc impacts *C. difficile* virulence remains unclear (Zackular *et al.*, 2017; Lopez *et al.*, 2019). Dietary trehalose has been linked to the dramatic increase in the number of cases of severe hospital-acquired CDI and emergence of *C. difficile* ribotype 027 in the early 2000's, as it was widely used as a food supplement during this period (Collins *et al.*, 2018). A proposed mechanism for this increased virulence is the presence of an additional trehalose metabolism operon in ribotype 027 strains (Rao *et al.*, 2015). Additionally, several studies suggest that the use of proton pump inhibitors is linked to an increase in the number of CDI cases in recent years. However, a study to prove causality and a plausible mechanism of action are yet to be established (Patil and Blankenship, 2013). A recent study investigated the effect of diarrhoeal events on susceptibility to *C. difficile* colonisation. It was reported that *C. difficile* colonisation occurred during recovery from diarrhoeal events, including gastrointestinal infections from pathogens such as *Vibrio cholerae*. Laxative treatment was used to simulate gastrointestinal disturbances in mice. These mice were found to have an increased susceptibility to *C. difficile* colonisation. This study suggests that microbiota disturbances caused by diarrhoeal events can trigger a window of susceptibility to *C. difficile* colonisation (Vaninsberghe *et al.*, 2020).

Interestingly, *C. difficile* colonises the gastrointestinal tract of approximately 50% of healthy infants. These infants are asymptomatic despite toxin levels in faeces being comparable to those seen during infection in adults (Rolfe and Iaconis,

1983). Newborn rabbits lacked brush boarder receptors for toxin A, suggesting that human infants may also lack the toxin A receptor in their colon, facilitating asymptomatic carriage (Eglow *et al.*, 1992). A later study found that human milk inhibited toxin A binding to purified hamster brush border membrane receptors (Rolfe and Song, 1995). Therefore, the consumption of human milk may also contribute to asymptomatic carriage in infants. The microbiota in the gut of infants matures with age, reducing the numbers of *C. difficile*, and by the age of 2 the gut microbiota resembles that of an adult with the absence of *C. difficile* (Holdeman, Good and Moore, 1976). It has been suggested that carriage in infants could act as a potential reservoir for pathogenic strains (Rousseau *et al.*, 2012).

Clinical diagnosis of *C. difficile* infection is critical for the successful treatment of the disease and prevention of transmission. However, *C. difficile* infection can be a challenge to diagnose and there is a lack of effective tools. Cell culture cytotoxicity neutralisation assay and toxigenic stool culture can be used to detect toxins and toxigenic strains of *C. difficile*. However, they require long turnaround times and excessive labour requirements, meaning they have limited utilisation in clinical settings. Toxin enzyme immunoassays (EIAs) can also be used to detect and diagnose *C. difficile* toxins, but potentially lack the appropriate sensitivity, resulting in concerns for false negative results. Nucleic acid amplification tests (NAATs) which target chromosomal toxin genes (usually the toxin B gene) are now the most commonly used tests for *C. difficile*, which are highly sensitive and can provide rapid results. However, NAATs cannot distinguish between toxigenic *C. difficile* infection or asymptomatic colonisation (Kociolek, Polage and Riley, 2016). *C. difficile* is also a particularly challenging infection to treat due to its

resistance to most classes of antibiotics and ability to form resistant spores in the gut. The recommended treatment for CDI is typically oral administration of vancomycin, metronidazole or fidaxomicin (Vardakas *et al.*, 2012). In recent years, fecal microbiota transplantation (FMT) has been a highly effective treatment for recurrent *C. difficile* infections, which functions by replenishing the gut microbiota. The mechanisms by which the microbiota inhibit *C. difficile* have not been fully elucidated, but it is likely that the presence of the microbiota inhibits spore germination and outcompete *C. difficile* for nutrients (Wilson, 1983; Sorg and Sonenshein, 2008; Kelly *et al.*, 2015). However, there are several risks associated with FMT treatment and it is reserved for severe cases with at least 3 episodes of disease recurrence. Firstly, there is a risk from the possible transmission of infectious agents from asymptomatic donors. Gut microbiota alterations have also been associated with the development of long-term diseases or conditions such as obesity, diabetes, atherosclerosis, irritable bowel disease (IBD) and colon cancer (Kelly *et al.*, 2015). A better understanding of the mechanisms by which FMT is effective against *C. difficile* will aid in the improvement of safety for FMT procedures. The identification of novel diagnostic markers for the disease is essential for the development of effective tools for the diagnosis and treatment of CDI.

1.2 How does *C. difficile* cause disease?

1.2.1 Spore germination

C. difficile is a strict anaerobe and in order to survive outside of an anaerobic environment the bacterium forms a metabolically inactive spore. Spores are

resistant to heat, chemicals, radiation and antibiotics, making a contaminated environment difficult to sterilise (Eckstein *et al.*, 2007; Anderson *et al.*, 2017). *C. difficile* infection is initiated by the ingestion of a spore from the environment into the gastrointestinal tract. Once ingested, *C. difficile* spores withstand the digestive enzymes and gastric acid in the stomach and pass through to the anaerobic environment of the duodenum. A germinant molecule is recognised by a receptor on the spore inner membrane and an irreversible spore germination reaction is triggered, leading to Ca^{2+}-dipicolinic acid release, water uptake and spore cortex degradation (Paredes-Sabja, Shen and Sorg, 2014). In the duodenum, bile is secreted by the gall bladder which consists mainly of cholate and chenodeoxycholate salts conjugated with either taurine or glycine. Many species of the microbiota, such as *Clostridium perfringens*, express bile salt hydrolases on their cell surfaces, which remove the conjugated amino acid from the primary bile salts in the small intestines (Gopal-Srivastava and Hylemon, 1988). These conjugated and deconjugated bile salts induce the germination of *C. difficile* spores into metabolically active vegetative cells (Wilson, 1983). It has also been demonstrated that cholate derivatives, such as taurocholate, activate spore germination more efficiently when combined with glycine or other amino acid co-germinants (Sorg and Sonenshein, 2008). However, the primary bile salt chenodeoxycholate competitively inhibits taurocholate-induced spore germination and vegetative outgrowth, which suggests a potential mechanism by which *C. difficile* uses the relative concentrations of various bile salts as cues for germination in the host (Sorg and Sonenshein, 2009, 2010). Unconjugated primary bile salts are taken up by species such as *Clostridium scindens* in the cecum, and are converted to secondary bile salts through a series of enzymatic

reactions in the cytoplasm (Wells and Hylemon, 2000). Secondary bile salts derived from gut microbiota species, such as deoxycholate, inhibit spore germination and bacterial outgrowth in the large intestines (Sorg and Sonenshein, 2008; Thanissery, Winston and Theriot, 2017). Therefore, reduction of secondary bile acid production during antibiotic-induced microbiota disruption has been linked to increased *C. difficile* spore germination and outgrowth in the large intestines (Theriot, Bowman and Young, 2016; Thanissery, Winston and Theriot, 2017).

Antibiotic-associated alterations of the gut microbiota increases the availability of nutrients required for *C. difficile* outgrowth. Precolonisation of a susceptible host with non-toxigenic *C. difficile* inhibited toxigenic *C. difficile* colonisation in a mouse model (Wilson and Sheagren, 1983; Gerding *et al.*, 2015). Cecal metabolomics following *C. difficile* colonisation in a mouse model were investigated to define the nutrients required for *C. difficile* colonisation and pathogenesis. It was suggested that *C. difficile* colonisation is associated with decreased levels of N-acetylated amino acids, carbohydrates, and sugar alcohols, as well as increased levels of metabolic by-products and lipids (Wilson and Perini, 1988; Fletcher *et al.*, 2018). These studies suggest that antibiotic-induced decreases in the abundance of commensal microbiota community members with similar nutritional requirements to *C. difficile*, such as commensal species of Clostridia, may be linked to increased susceptibility to *C. difficile* colonisation. The mechanism of *C. difficile* colonisation is summarised in Figure 1.

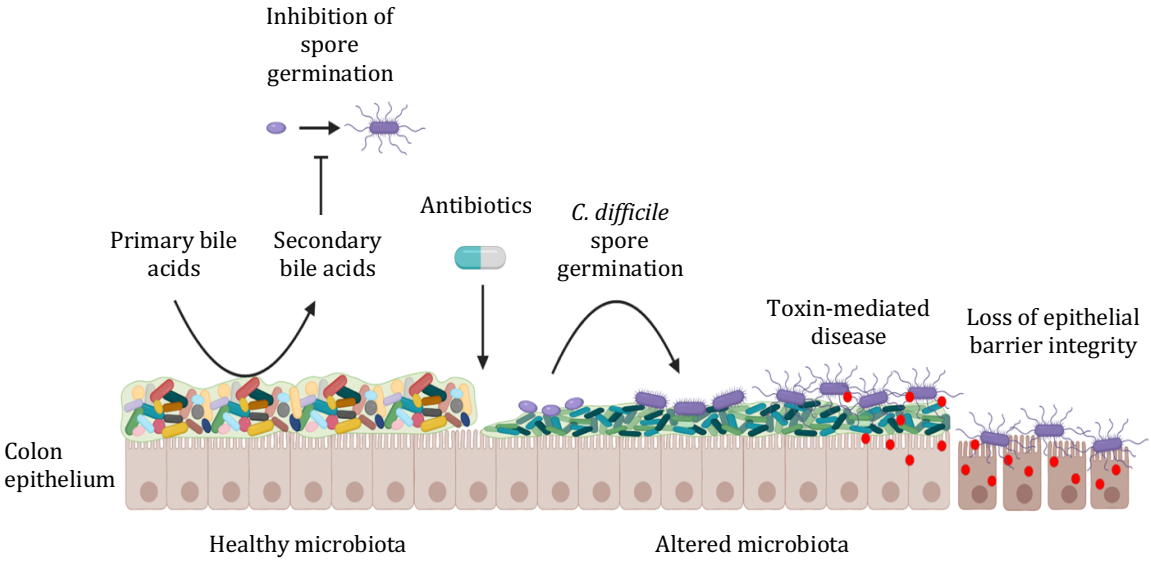


Figure 1. *C. difficile* colonisation of the gut epithelium.

A healthy gut microbiota provides colonisation resistance against *C. difficile*, inhibiting spore germination and preventing colonisation. Antibiotic treatments cause alterations in the microbiota which confers susceptibility to *C. difficile* colonisation. Spores germinate into vegetative *C. difficile* cells and toxin-mediated disease leads to disruption of the epithelial barrier.

1.2.2 Colonisation factors

Following spore germination, *C. difficile* colonises the gut mucosa for successful establishment of disease. Several colonisation factors have been identified which have exhibited adhesive properties to host extracellular matrix (ECM) components, epithelial cell layers or gastrointestinal tissues.

1.2.2.1 *C. difficile* surface proteins

Surface layer proteins (SLPs) play a major role in *C. difficile* adhesion and host-cell colonisation. The *C. difficile* surface layer is comprised of 2 major SLPs, which

are encoded by the same gene *slpA* which encodes a precursor protein, SlpA. SlpA is cleaved by the Cwp84 protease to generate the two mature proteins, one with a high molecular weight (~40 kDa), HMW-SLP, and one with a low molecular weight (~35 kDa), LMW-SLP (Karjalainen *et al.*, 2001). *In vitro* studies demonstrated that *C. difficile* SLPs bind to Hep-2 cells, Vero cells, and human gastrointestinal tissues. The HMW-SLP also bound to immobilised components of the ECM, such as collagen and thrombospondin. However, there was no observed binding to collagen IV, fibronectin or laminin (Calabi *et al.*, 2002). The chemical removal of SLPs or treatment of *C. difficile* bacterial cells with anti-SLP Fab fragments abolished the adherence of *C. difficile* to mouse 929 and human HeLa cells (Takeoka *et al.*, 1991). These studies demonstrated the importance of SLPs as colonisation factors.

1.2.2.2 Cell wall proteins

Cell wall proteins (CWPs) are also important colonisation factors of *C. difficile*. The Cwp84 (for "clostridial wall protein 84 kDa") protease not only cleaves the SlpA precursor protein, but also exhibited ECM-degrading activity on fibronectin, laminin, and vitronectin. It has been suggested that degradation of the basement membrane may compromise the integrity of host tissues and contribute towards the necrosis of the intestinal epithelium and bacterial dissemination (Janoir *et al.*, 2007). Cwp84 has also been reported to have a role in the modulation of *C. difficile* biofilms *in vitro* (Dawson *et al.*, 2012; Dapa *et al.*, 2013; Pantaléon *et al.*, 2015). Cell wall proteins, such as Cwp66, have also been reported to play a role in *C. difficile* colonisation of gastrointestinal tissues. Interestingly, *C. difficile* attachment to Vero cells was increased under certain stress conditions, including

heat shock, high osmolarity, acid shock and iron starvation (Waligora *et al.*, 1999). Anti-Cwp66 antibodies partially inhibited heat-shocked *C. difficile* adherence to Vero cells through competitive inhibition. However, the adherence of non-heat shocked *C. difficile* was not inhibited by anti-Cwp66 antibodies. The partial inhibition of adherence demonstrated in this study suggested that multiple adhesins contribute to the cellular adherence of *C. difficile*. This study also suggested that heat shock may induce a conformational change in Cwp66 which gains it adhesive properties (Waligora *et al.*, 2001). The highly conserved heat shock protein GroEL, is released extracellularly after heat shock and has also been reported to exhibit an adhesive function (Hennequin, Collignon and Karjalainen, 2001).

1.2.2.3 Flagella

As well as bacterial motility, the flagella has also been implemented in *C. difficile* adhesion to host tissues. Specifically, the flagellar cap protein, FliD, and the flagellin, FliC have been suggested to contribute towards *C. difficile* R20291 colonisation of host tissues. Recombinant FliC and FliD, as well as crude flagella were reported to bind to axenic mouse cecal mucus. This suggests there is a receptor for flagella binding in mouse cecal mucus and the flagella may play a role in the attachment to mucus in the colon, facilitating the establishment of infection in the gut (Tasteyre *et al.*, 2001). However, later studies reported that *fliC* and *fliD* mutants in *C. difficile* 630, which lack flagella and are non-motile, had increased adherence to intestinal Caco-2 colonocytes (Dingle, Mulvey and Armstrong, 2011; Baban *et al.*, 2013). There was also no observed differences in virulence or intestinal colonisation of *ΔfliC* or *ΔfliD* compared to the WT in a

hamster model (Dingle, Mulvey and Armstrong, 2011). Therefore, the role of the flagella in *C. difficile* colonisation and pathogenesis remains a controversial subject and may differ between strains (Baban *et al.*, 2013). Interestingly, *fliC* and *fliD* mutants also exhibited increased toxin production in both strains, suggesting these flagella proteins may also play a role in *C. difficile* virulence (Dingle, Mulvey and Armstrong, 2011; Baban *et al.*, 2013).

1.2.2.4 Pili

Type IV pili have also been reported to be important for intestinal colonisation in mice, as a type IV pili mutant was cleared more quickly by the immune response and exhibited reduced attachment to the colon mucosa compared to the WT (McKee *et al.*, 2018). Pili have also been reported to modulate bacterial autoaggregation, a process which may be important for host-cell colonisation and the early stages of biofilm formation. This further demonstrates the adhesive properties of pili and the important role they may play in the pathogenesis of *C. difficile* infection (Bordeleau *et al.*, 2015).

1.2.2.5 Extracellular matrix-binding proteins

Vegetative *C. difficile* cells also bind to several components of the ECM, including fibronectin, collagen types I, III, IV and V, fibrinogen and vitronectin, all of which are located in the basement membrane that underlies the gut intestinal epithelium (Cerquetti *et al.*, 2002). *C. difficile* expresses many cell-surface exposed ECM binding proteins, including fibronectin binding proteins such as Fbp68 (Hennequin *et al.*, 2003) and FbpA (Barketi-Klai *et al.*, 2011). In a moxogenic mouse model, a *fbpA* mutant had impaired colonisation capacity in

the cecum compared to the WT, demonstrating a potential role for this protein in adhesion to the gastrointestinal tissues (Barketi-Klai *et al.*, 2011). *C. difficile* also expresses cell-surface collagen binding proteins, including CD2831 and CbpA (Tulli *et al.*, 2013; Arato *et al.*, 2019). Recombinant CbpA was reported to adhere to immobilised collagen I and V, as well as murine intestines. However, a *cbpA* knockout strain exhibited no significant difference in its adhesion to immobilised collagen V or IMR-90 lung fibroblast cells, as well as its colonisation capacity in a dixenic mouse model of infection (Janoir *et al.*, 2013; Tulli *et al.*, 2013). This is likely to be due to redundancies in the adhesive bacterial surface proteins. To overcome the adhesion redundancies in *C. difficile* and assess the functionality of CbpA, this protein was expressed in *Lactococcus lactis*. CbpA expression in *L. lactis* increased bacterial adherence to immobilised collagen V and IMR-90 cells *in vitro*, demonstrating that CbpA may contribute towards *C. difficile* colonisation (Tulli *et al.*, 2013). The collagen-binding protein CD2831 has been recently shown to be important for *C. difficile* biofilm formation and host colonisation (Arato *et al.*, 2019; Dawson *et al.*, 2021). Overexpression of this adhesin significantly increased biofilm formation *in vitro* and resulted in greater adhesion to both immobilised collagen and collagen produced by human IMR-90 cells, highlighting the importance of CD2831 in adhesion to host ECM components. CD2831 has also been reported to bind to the human complement component C1q, acting as a potential inhibitor of the complement activation pathway (Arato *et al.*, 2019). This suggests that the CD2831 adhesin may have multiple roles in *C. difficile* pathogenesis, including colonisation of host tissues and evasion of the immune response. This property places CD2831 in the Microbial Surface Components Recognising Adhesive Matrix Molecule (MSCRAMMs) family, a

collection of virulence factors which have a role in ECM binding and modulation of the host immune response (Wann, Gurusiddappa and Höök, 2000; Arato *et al.*, 2019).

1.2.2.6 Secreted colonisation factors

The CD2831 collagen-binding protein is proteolytically cleaved by a zinc metalloprotease secreted by *C. difficile*. Bacterial zinc-dependent metalloproteases often have a role in pathogenesis, such as the neurotoxins from the related species *Clostridium botulinum* and *Clostridium tetani*, and the lethal factor from *Bacillus anthracis* (Simpson, Maksymowych and Hao, 2001; Sherer *et al.*, 2007). *Treponema pallidum* pallilysin is a zinc metalloprotease which cleaves laminin and fibrinogen, implicating a potential role for this protease in bacterial dissemination during infections (Houston *et al.*, 2012). A novel extracellular zinc metalloprotease CD2830, was identified by Cafardi *et al.* through a proteomic analysis of *C. difficile* culture supernatants (Cafardi *et al.*, 2013). This protein had a high degree of structural similarity to the lethal factor of *Bacillus anthracis*, an important toxin for anthrax pathogenesis, despite the sequence identity between the two proteins being relatively low (22%) (Sherer *et al.*, 2007; Cafardi *et al.*, 2013; Rubino *et al.*, 2016). The protease activity of CD2830 was dependent on its binding to Zn^{2+}, leading it to be named zinc metalloprotease 1 (Zmp1) (Cafardi *et al.*, 2013). The Zmp1 protease has a unique cleavage site which preferentially cleaves proteins with a (V, L, I)NPPVP motif between the P3-P4' proline residues. Although Proline-Proline bond hydrolysis has been previously observed, Zmp1 was the first identified protease to exhibit a strong preference for these peptide bonds, leading Zmp1 to be renamed proline-proline endopeptidase 1 (PPEP-1)

due to its unique cleavage site specificity (Hensbergen *et al.*, 2014). A PPEP-1 homologue was discovered in *Paenibacillus alvei*, which was denoted PPEP-2. Like PPEP-1, PPEP-2 is a secreted protease which cleaves a cell-surface adhesin located adjacent to the PPEP-2 gene. However, the cleavage motif of the two PPEP proteases are distinct, and substrates optimally cleaved by one PPEP protease cannot be cleaved by the other (Klychnikov *et al.*, 2018). The PPEP-1 signal sequence, which targets PPEP-1 for secretion, has also been fused to synthetic reporter constructs to produce new secreted proteins in *C. difficile*. These secreted reporter constructs can potentially be used for gene expression analyses, as their expression is detectable in solid and liquid cultures (Oliveira Paiva *et al.*, 2016).

Multiple human substrates cleaved by PPEP-1 have been identified, including fibronectin, a component of the ECM, and the heat-shock protein, HSP90β (Cafardi *et al.*, 2013). The exact function of HSP90β is not clear, but it may play a role in stabilising extracellular proteins against external stresses and host inflammatory responses to bacterial infections (Prodromou *et al.*, 1997; Byrd *et al.*, 1999; Triantafilou *et al.*, 2001). IgA1/2, antibodies also contain a motif in the bend region which may be cleaved by the PPEP-1 protease. These antibodies have a crucial role in mucosal immune responses (Smith *et al.*, 2006; Mantis, Rol and Corthésy, 2011). Therefore, PPEP-1-mediated cleavage of IgA antibodies may aid in *C. difficile* evasion of the host immune response in the gut mucosa (Hensbergen *et al.*, 2014). Another human substrate for PPEP-1 is fibrinogen (*β* chain), a soluble glycoprotein which circulates in the bloodstream, is incorporated in the ECM, and plays a central role in blood clot formation and

wound healing (Mosesson, 2012; Cafardi *et al.*, 2013; Schacherl *et al.*, 2015). These findings suggest that PPEP-1 may contribute towards *C. difficile* dissemination through the basement membrane of the gastrointestinal tissue during infection. The PPEP-1 cleavage motif was investigated further using a set of biochemical and *in vitro* proteolytic assays with a synthetic peptide library. Screening for peptides with this cleavage site identified two *C. difficile* targets for PPEP-1, CD2831 and CD3246. Both substrates are putative cell-surface collagen-binding proteins which covalently linked to the peptidoglycan cell wall by a sortase (SrtB)-mediated reaction. Interestingly, CD2831 is located adjacent to the *PPEP-1 (CD2830)* gene in the strain 630 *C. difficile* genome. Both substrates contain multiple consecutive cleavage sites, located adjacent to the peptidoglycan cell wall anchor motif (Hensbergen *et al.*, 2014). The activity of recombinant PPEP-1 was sufficient to completely remove the cell-surface CD2831 when cells were in stationary phase during *in vitro* growth, even in a system where CD2831 is overexpressed (Hensbergen *et al.*, 2015; Peltier *et al.*, 2015). Moreover, uncleaved endogenous CD2831 in a PPEP-1 mutant could be removed from the bacterial cell-surface through cleavage with recombinant PPEP-1, further demonstrating the crucial role of PPEP-1 in cleavage of CD2831 (Hensbergen *et al.*, 2015). Interestingly, a PPEP-1 mutant also exhibited an attenuation in virulence in a hamster model, demonstrating a potential role for this protease in *C. difficile* virulence (Hensbergen *et al.*, 2015).

Cyclic diguanylate (3=,5=-cyclic diguanylic acid) (c-di-GMP) is a small-molecule second messenger used by prokaryotes to regulate gene expression by binding to riboswitches located upstream of the target genes. In *C. difficile*, two types of

c-di-GMP riboswitches have been identified which are involved in *C. difficile* colonisation and pathogenesis. Riboswitch type I occurs when there are low intracellular concentrations of c-di-GMP and is motility associated, resulting in the upregulation of genes involved in processes such as flagellar assembly and chemotaxis. Riboswitch I also indirectly modulates the expression of *tcdA* and *tcdB* through modulation of the *C. difficile* flagellar alternative sigma factor, SigD, demonstrating a link between the c-di-GMP riboswitch and virulence (McKee *et al.*, 2013). Riboswitch type II occurs when there are high intracellular concentrations of c-di-GMP and is adhesion-associated. Cell-surface adhesins, such as CD2831 and CD3246, as well as pilin biosynthesis are upregulated, all of which may contribute towards bacterial adherence and host-cell colonisation. It has been established that *C. difficile* motility is repressed, and cells aggregate together in high levels of c-di-GMP (Purcell *et al.*, 2012). Peltier *et al.* demonstrated that PPEP-1 transcription is inversely regulated by c-di-GMP, where transcription levels of PPEP-1 are low in the presence of high c-di-GMP concentrations. However, the PPEP-1 substrates, CD2831 and CD3246, are elevated in the presence of high intracellular c-di-GMP concentrations. This evidence links PPEP-1 expression to the motility-associated riboswitch type I phenotype and suggests it may be involved in *C. difficile* virulence (Peltier *et al.*, 2015). However, the molecular function of PPEP-1 has not been investigated in the context of infection.

1.2.3 *C. difficile* biofilms

It has been suggested that biofilms could play a role in the different phases over the course of infection, and may be a cause for disease recurrence, as biofilms

may aid *C. difficile* in resisting host immune responses and antibiotics (Frost, Cheng and Unnikrishnan, 2021). Formation of *C. difficile* aggregates around damaged tissue in mice suggests that microcolonies may form on the surface of the epithelial cell layer (Lawley *et al.*, 2009). *C. difficile* was first shown to form mono-species biofilms *in vitro* by Donelli *et al.*, in 2012, and since then many factors which modulate *C. difficile* biofilm formation have been identified, including flagella proteins, the S-layer, and sporulation (Dawson *et al.*, 2012; Donelli *et al.*, 2012; Dapa *et al.*, 2013; Semenyuk *et al.*, 2014). In a mixed species biofilm community in a mouse model, the inner mucus layer and epithelial cell surface were very rarely inhabited by *C. difficile*. The *C. difficile*-bearing multispecies communities were found most commonly with members of *Bacteroidaceae* and *Enterobacteriaceae* in the proximal colon in the outer MUC2 mucus layer (Semenyuk *et al.*, 2015). It has also been demonstrated that *C. difficile* forms biofilm-like structures in a germ-free murine model. These structures were glycan-rich and are mostly localised outside of the mucus layer (Soavelomandroso *et al.*, 2017). CDI patients have been reported to exhibit decreased secretion of MUC2, and an increased secretion of MUC1, which is typically adhered to the cell surface. Decreased expression of MUC2 may allow for increased *C. difficile* binding to MUC1 on the cell surface and may provide a site for *C. difficile* attachment to the epithelial cell layer (Engevik *et al.*, 2015). A recent study demonstrated that *C. difficile* adhered to the O-linked glycan component of MUC2. A bioreactor inoculated with healthy human faeces, was used to model the microbiota before being treated with clindamycin and infected with *C. difficile* to generate a model for *C. difficile* infection. Human MUC2-coated coverslips were added to the bioreactor and 16S rRNA sequencing revealed the

profile of the multispecies biofilm formed on the coverslip. A substantial co-colonisation with *Fusobacterium* was observed. *C. difficile* was found to aggregate with *Fusobacterium nucleatum*, and blockage or mutation of the *F. nucleatum* RadD adhesin or *C. difficile* flagella inhibited this effect. Interestingly, *C. difficile* biofilm formation and extracellular polysaccharide production were enhanced after the addition of *F. nucleatum* (Engevik *et al.*, 2020). This study highlighted the complex interactions of *C. difficile* and the gut microbiota and its importance in the pathogenesis of *C. difficile* infection.

A recent study demonstrated that *C. difficile* biofilms are composed of vegetative cells and spores encased in a protective matrix of extracellular DNA (eDNA), cell surface proteins and intracellular proteins. Two c-di-GMP regulated proteins, *C. difficile* gene 2831 (CD2831) and CD3246 were identified as having roles in *C. difficile* biofilm formation, both of which are substrates for PPEP-1. There was no observed difference in biofilm formation between WT and *ΔPPEP-1*. However, an insertion mutant in the PPEP-1 substrate, CD2831, exhibited significantly reduced biofilm formation compared to the WT and complemented strain. It was also reported that overexpression of CD2831 and CD3246 in *ΔPPEP-1* resulted in a significant increase in biofilm formation. Two c-di-GMP-independent cell surface proteins, CD3392 and CD0183, were also identified as having roles in *C. difficile* biofilm formation, as mutations in these genes significantly reduced early biofilm biomass (Dawson *et al.*, 2021). This study demonstrated the numerous intracellular and extracellular components that are required for the formation and structural integrity of *C. difficile* biofilms. It has been suggested that biofilms may provide a mechanism which contributes towards *C. difficile* persistence in

the gut, however the specific role of biofilms in *C. difficile* pathogenesis has not been investigated (Frost, Cheng and Unnikrishnan, 2021).

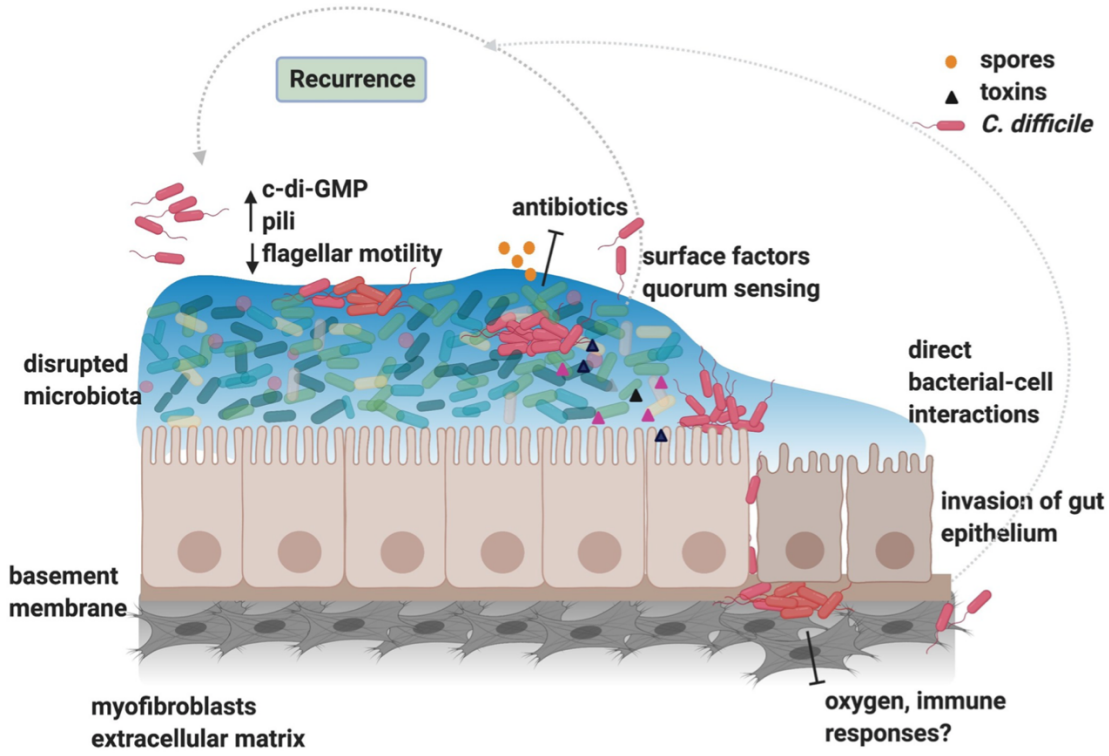


Figure 2. A model for C. difficile biofilms during infection.

After the native gut microbiota is altered from broad spectrum antibiotics, *C. difficile* attaches to the gut mucosa. An increase in intracellular c-di-GMP levels results in a decrease in bacterial motility enables attachment and establishment of microaggregates or communities. These communities could exist as single species or in close association with the gut microbiota, serve as a niche for production of spores and toxins and provide protection from oral antibiotics in the lumen used during treatment (e.g., vancomycin, metronidazole). *C. difficile* cell surface factors (e.g., pili, flagella, S-layer), quorum sensing (e.g., LuxS), and regulators (e.g., Spo0A, CD630_2214) control biofilm/aggregate formation. The action of *C. difficile* toxins trigger cell death and disruption of the epithelial barrier, allowing bacteria to penetrate the epithelial cell layer to the underlying basement membrane. *C. difficile* may form communities in underlying tissue

which may protect bacteria from oxygen and immune responses. Bacterial communities in the gut mucosa may facilitate bacterial persistence and may lead to recurrence of the infection. Figure reproduced with permission from Frost, Cheng and Unnikrishnan (2021).

1.2.4 *C. difficile* Toxins

1.2.4.1 Regulation of toxin production

Following successful colonisation of the gut epithelium, *C. difficile* replicates and secretes toxins: the enterotoxin TcdA, the cytotoxin TcdB, as well as a binary ADP-ribosylating toxin, CDT in ribotype 027 strains. These toxins are primarily responsible for epithelial barrier disruption, tissue damage and fluid accumulation during *C. difficile* infection (Voth and Ballard, 2005). Despite being similar in structure (N-terminal domains have 74% homology), TcdA and TcdB perform independent, nonredundant functions (Voth and Ballard, 2005). Additionally, hypervirulent ribotype 027 strains produce a third toxin termed the binary ADP-ribosyltransferase toxin, *Clostridium difficile* transferase (CDT). The toxin genes are encoded on the 19.6 Kb pathogenicity locus (PaLoc) which is highly conserved between toxigenic strains of *C. difficile*. The PaLoc consists of five genes: *tcdA*, *tcdB*, *tcdC*, *tcdE* and *tcdR*. The two toxin genes, *tcdA* and *tcdB* are separated by the *tcdE* gene, which is important for toxin secretion from cells. This protein has pore-forming activity and shares homology to phage holin, a protein responsible for the release of phage from a host cell. TcdC is an anti-sigma factor which negatively regulates the expression of TcdR, an alternative sigma factor which directs transcription of TcdA and TcdB, as well as its own promoter. The

expression of the toxin genes is regulated by CodY, a global gene regulator which represses toxin gene expression when nutrients are in abundance by binding to the *tcdR* promotor region. However, when nutrients are diminished during stationary phase, CodY dissociates from *tcdR* which stimulates expression of TcdA and TcdB (Voth and Ballard, 2005; Dineen *et al.*, 2007; Carter, Rood and Lyras, 2012). The *C. difficile* PaLoc is illustrated in Figure 3.

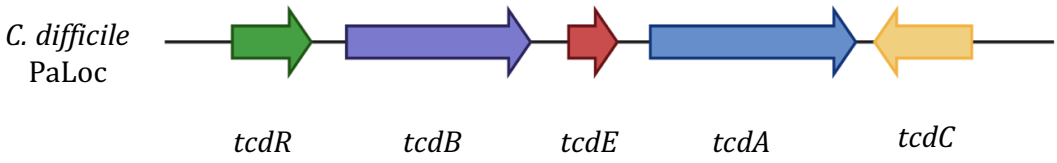


Figure 3. *C. difficile* pathogenicity locus.

1.2.4.2 Toxin receptors

Intoxication occurs when toxins bind to host cell receptors and are internalised via receptor-mediated endocytosis. Multiple receptors for TcdA and TcdB have been proposed and vary for different cell types. First, TcdA was reported to bind to the trisaccharide Galal-3GaI1-4GlcNAc which occurs in large amounts on the intestinal epithelium of hamsters, but not humans (Krivan *et al.*, 1986). Since then, various human cell-surface oligosaccharides, including carbohydrate antigens I, X, and Y, have been identified as TcdA binding sites (Tucker and Wilkins, 1991; Greco *et al.*, 2006). TcdA contains a carbohydrate-binding combined repetitive oligopeptides (CROPs) domain which mediates attachment and cytotoxicity in epithelial cell-surfaces (von Eichel-Streiber and Sauerborn, 1990). However, the CROPs domain is not essential for TcdA cellular uptake,

demonstrating the existence of CROPs-independent receptors (Olling *et al.*, 2011). Gp96, a member of the heat shock protein family, is another receptor for TcdA cellular uptake which is expressed on human colonocyte apical membranes as well as in the cytoplasm (Na *et al.*, 2008). Recent studies have reported that TcdA binds to sulfated glycosaminoglycans (sGAGs) and low-density lipoprotein receptors (LRP) in a CROPs-independent manner (Tao *et al.*, 2019). Multiple receptors for TcdB attachment and cellular uptake have also been identified. Chondroitin sulphate proteoglycan 4 (CSPG4) has been identified as a cellular TcdB receptor and mediator of TcdB-induced cytopathic and apoptotic events in HeLa and HT29 cells (Yuan *et al.*, 2015). Poliovirus receptor-like 3 (PVRL3/NECTIN3), a cellular adhesin molecule highly expressed on the surface of the colonic epithelium, has also been identified as an epithelial cell receptor for TcdB in Caco-2 cells (LaFrance *et al.*, 2015). Wnt receptor Frizzled (FZDs) family members have also been reported to act as receptors for TcdB. Specifically, FZD1/2/7 triple-knockout (KO) HeLa cells were highly resistant to TcdB and colonic epithelia in FZD7$^{-/-}$ mice were less susceptible to TcdB-induced tissue damage *in vivo* (Tao *et al.*, 2016). These studies demonstrate the numerous receptors which facilitate toxin translocation into host cells.

1.2.4.3 Cellular effects of toxins

The Rho family of GTPases proteins are molecular switches which play key roles in numerous processes, including the regulation of tight junction assembly and cellular integrity of intestinal epithelial layers (Walsh *et al.*, 2001). Once TcdA and TcdB are translocated into the cytosol of target cells, dithiothreitol and *β*-mercaptoethanol induced auto-cleavage of the catalytic domain occurs (Egerer

et al., 2007). The *C. difficile* toxins then permanently inactivate Rho GTPases by mono-glucosylation at the Thr-35 and Thr-37 residues (Nusrat *et al.*, 2001; Egerer *et al.*, 2007; Gerhard *et al.*, 2008; Chen *et al.*, 2015). Glucosylation of Rho proteins inhibits critical signalling pathways, including reorganisation of actin microfilaments and major shrinking of the cell (Ottlinger and Lin, 1988). Cell-cell contacts and cell adhesion are also disrupted through dissociation of occludin, ZO-1, and ZO-2 in tight junctions, which are regulated by Rho proteins and actin. Disruption of these proteins ultimately increases paracellular permeability and results in loss of epithelial barrier integrity (Nusrat *et al.*, 2001; Walsh *et al.*, 2001). The glucosyltransferase activity of TcdA and TcdB also induces apoptotic cell death of target cells. TcdA triggers apoptosis through the activation of caspase-3, 8, and 9, while TcdB triggers apoptosis through inactivation of RhoA (Fiorentini *et al.*, 1998; Hippenstiel *et al.*, 2002; Nottrott *et al.*, 2007; Gerhard *et al.*, 2008). It has been shown that TcdB alone is sufficient to induce intestinal inflammation and elicit epithelial layer damage in a hamster model (Lyras *et al.*, 2009). TcdA and TcdB also trigger inflammasome activation in an ASC (apoptosis-associated speck-like protein with a caspase recruitment domain (CARD))-dependent manner. Inflammasome activation stimulates the secretion of interleukin-1β (IL-1β) in macrophages, a pro-inflammatory cytokine which is commonly detected in CDI patients (Ng *et al.*, 2010; Czepiel *et al.*, 2014; Liu *et al.*, 2018). The severity of *C. difficile*-induced colitis has been linked to the magnitude of the immune response which can have an impact on the clinical manifestations and outcomes of the disease (Ng *et al.*, 2010).

The *C. difficile* binary toxin, CDT, binds to the lipolysis-stimulated lipoprotein receptor (LSR) and is translocated into the cytosol of target cells (Papatheodoroua *et al.*, 2011). This toxin ADP-ribosylates monomeric globular actin (G-actin) which inhibits actin polymerisation (Popoff *et al.*, 1988; Perelle *et al.*, 1997; Barth *et al.*, 2004). ADP-ribosylated actin also acts as a capping protein for F-actin which prevents elongation and destabilises actin filaments (Wegner and Aktories, 1988; Aktories and Wegner, 1992). CDT has also been reported to change the organisation and dynamics of microtubules, inducing formation of elongated microtubule-based protrusions and increasing microtubule bundling in the cytoplasm. These protrusions increased the adherence of *C. difficile* to colon epithelial cells, presenting a potential role for CDT in bacterial colonisation (Schwan *et al.*, 2009). The CDT binary toxin also disrupts the host eosinophilic response, a Toll-like receptor 2 (TLR2)-dependent pro-inflammatory pathway. This study further demonstrated the advanced virulence mechanisms of CDT-expressing strains as well as clarified the protective role of the eosinophil response in *C. difficile* infection (Cowardin *et al.*, 2017).

Together, toxin-induced modulation of these essential cellular functions are strongly associated with loss of epithelial barrier integrity (Fiorentini *et al.*, 1998; Pothoulakis, 2000; Nusrat *et al.*, 2001). It has been suggested that toxin-mediated actin myofilament disaggregation and tight junction opening allows *C. difficile* to penetrate the epithelial cell layer and bind to components the underlying ECM (Pothoulakis, 2000). The combined effect of these alterations is the loss of fluid into the intestinal lumen which manifests as diarrhoea (Nusrat *et al.*, 2001; Voth and Ballard, 2005).

1.2.4.4 Antiproliferative effects

C. difficile TcdA and TcdB also exert antiproliferative effects as they block the G_1-S and the G_2-M transitions, in part due to the reorganisation of actin and subsequent inhibition of contractile ring formation during cytokinesis (Fiorentini *et al.*, 1998; Ando *et al.*, 2007; Nottrott *et al.*, 2007; D'Auria *et al.*, 2012). G_1 phase is arrested even at low toxin concentrations by inhibiting the expression of cyclin D1 and blocking its binding to Cyclin-dependent kinase 4 (CDK4) and CDK6 (D'Auria *et al.*, 2012). The cyclin D1 gene is transcriptionally regulated by several mechanisms, including the Wnt/β-catenin pathway (Klein and Assoian, 2008). The Wnt/β-catenin signalling pathway plays a major role in the development and self-renewal of human intestine cells, and is a primary driving force of proliferation (van der Flier and Clevers, 2009). *C. difficile* TcdA binding to FZD Wnt receptors has been reported to attenuate Wnt signalling *in vitro*, and therefore is likely to contribute towards the antiproliferative effects observed in intestinal epithelial cells during *C. difficile* infection. However, the role of these antiproliferative effects in the pathogenesis of *C. difficile* are currently not understood (Lima *et al.*, 2014). It has been suggested that inhibition of proliferation and self-renewal may contribute towards disruption of epithelial barrier integrity which leads to diarrhoea, as this is also a common side effect of anti-cancer drugs which supress cellular proliferation (Major *et al.*, 2007). The main cellular activities of TcdA, TcdB and CDT are illustrated in Figure 4.

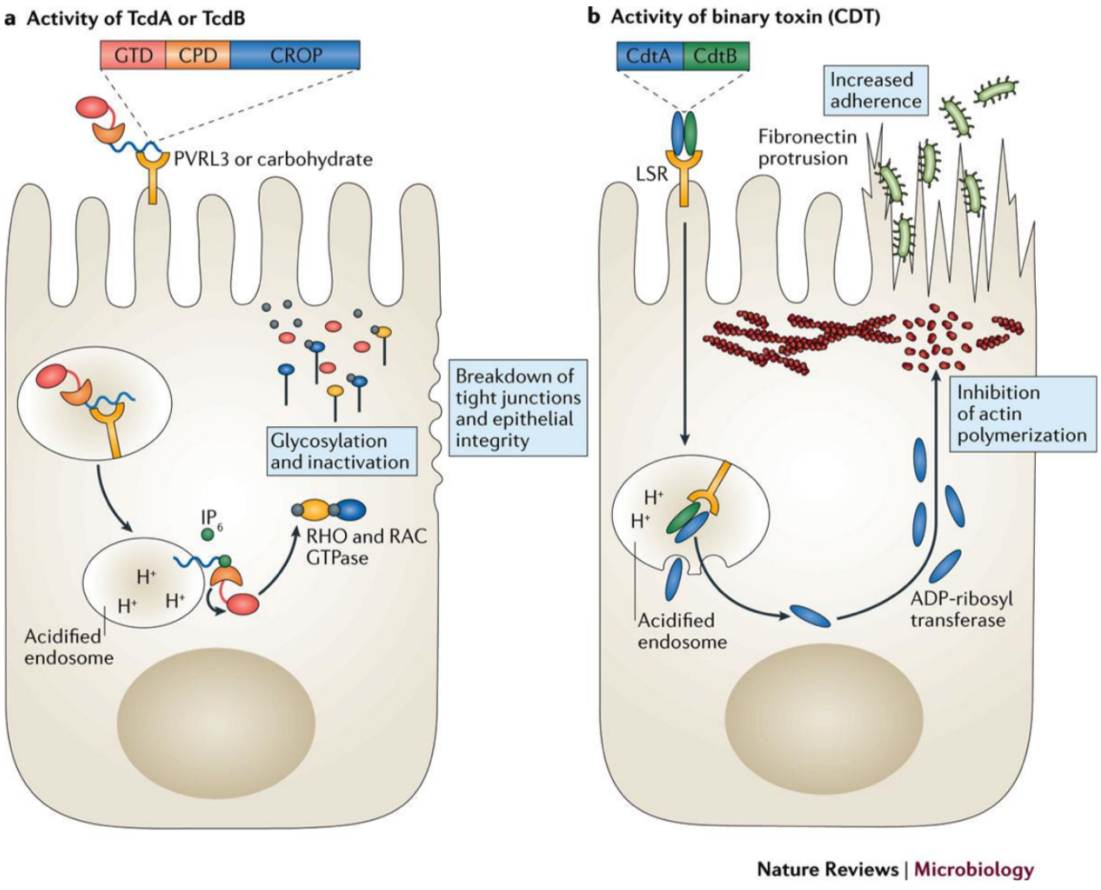


Figure 4. The mode of action of *C. difficile* toxins.

A. the CROPS domain of TcdA binds to carbohydrates on the apical surface of epithelial cells, whereas TcdB binds to a poliovirus receptor-like 3 (PVRL3) expressed on colonic epithelial cells. Toxins are internalised, where the acidification of the endosome enables the CROP domain to embed into the endosomal membrane. The cysteine protease domain (CPD) and the glucosyl transferase domain (GTD) are consequently transported into the cytosol. The cysteine protease is activated by inositol hexakisphosphate (IP6) to release the toxin glucotransferase. RHO or RAC GTPases are inactivated by glucosylation, which results in the breakdown of tight junctions and epithelial barrier integrity. B. *C. difficile* transferase (CDT) is internalised by binding to the lipolysis-stimulated lipoprotein receptor (LSR). The CdtB subunit creates pores in the acidified endosome releasing the CdtA subunit into the cytosol. The ADP-ribosyl transferase activity of the CdtA subunit inhibits actin polymerisation near the cell membrane. This enables the fibronectin microtubules to elongate and protrude through microvilli, enhancing *C.*

difficile adherence to the epithelial layer (Abt et al., 2016). Figure reproduced with permission from Springer Nature.

1.2.5 Sporulation

C. difficile relies on sporulation for disease transmission. Due to the strict anaerobic growth requirements of *C. difficile*, it is unable to survive in aerobic conditions in its vegetative form and relies on the dormant spore morphotype to survive in aerobic environments. *C. difficile* spores are composed of several structural layers: the spore core, inner membrane, germ cell wall, cortex, outer membrane, coat and exosporium. The structure of the spores contribute to their resistance properties (Pereira *et al.*, 2013; Paredes-Sabja, Shen and Sorg, 2014). Spores can exist in the environment for long periods of time and can resist extreme environmental factors such as high heat, pH stress, mechanical forces and aerobic conditions (Paredes-Sabja, Shen and Sorg, 2014). Importantly, *C. difficile* spores are also resistant to antibiotics, the host immune system and bleach-free disinfection protocols commonly used in hospital and care home settings (Lawley *et al.*, 2009). Disease transmission occurs when patients with CDI excrete a large number of spores in their faeces which disseminate into the environment before being ingested by a susceptible recipient (Deakin *et al.*, 2012). A number of spores may also remain in the gastrointestinal tract which can germinate once treatments have ceased. The germination of these persistent spores and subsequent recolonisation of the gastrointestinal tract may result in a relapse of the disease (Paredes-Sabja and Sarker, 2012).

The signals which trigger *C. difficile* sporulation have not been fully identified, but in related *Bacillus* species the initiation of sporulation relies on environmental stimuli such as nutrient starvation, quorum sensing, and other unidentified stress factors (Higgins and Dworkin, 2012). The sporulation pathway is triggered by the phosphorylation of the master transcriptional regulator, Spo0A, by 5 orphan histidine kinases (Underwood *et al.*, 2009). The stimulation of the sporulation pathway results in a cascade of activation of 4 sporulation-specific RNA polymerase sigma factors, σF, σG, σE and σK. These sporulation sigma factors operate by compartment-specific activation with σF and σG activity being restricted to the forespore and σE and σK activity being restricted to the mother cell (Pereira *et al.*, 2013). As well as the sporulation cascade, Spo0A also regulates several other cellular processes, including key virulence factors such as *tcdA* expression, biofilm formation and metabolic pathways (Underwood *et al.*, 2009; Deakin *et al.*, 2012; Pettit *et al.*, 2014). An R20291 *Δspo0A* mutant strain exhibited an asporogenous phenotype but exhibited increased virulence in mice compared to the WT, due to an increase in toxin production. However, *Δspo0A* was severely impaired in its ability to persist in the gut and no relapses of infection occurred after antibiotic treatment. Mice infected with *Δspo0A* also failed to transmit the infection, demonstrating that sporulation is required for host-to-host transmission (Deakin *et al.*, 2012). This study further highlighted the key role of sporulation in the infection cycle of *C. difficile*. Sporulation has also been reported to be involved in biofilm formation, as a *spo0A* mutant formed biofilms with reduced depth and breadth compared to the WT (Dawson *et al.*, 2012; Dapa *et al.*, 2013; Semenyuk *et al.*, 2014). *Δspo0A* biofilms also had significantly less resistance to oxygen stress than the WT (Dawson *et al.*, 2012). This suggests that

biofilm formation may also be regulated by the Spo0A master transcriptional regulator.

1.2.6 Other virulence factors

Interestingly, there have been several reports of infections which have been associated with non-toxigenic strains of *C. difficile*, suggesting that there may be other factors which contribute towards *C. difficile* virulence (Pron *et al.*, 1995; Castagliuolo *et al.*, 1997; Gravisse *et al.*, 2003; G. Martirosian *et al.*, 2004). During *C. difficile* infection, the abundance of haemoglobin in the gastrointestinal lumen increases, and the reactivity of heme can have toxic effects to both eukaryotic and prokaryotic cells. A study reported that heme activated transporter system (HatRT, R = regulator, T = transporter) mutants were not defective in colonisation or persistence in a mouse model of infection, suggesting that *C. difficile* has additional mechanisms to survive excess heme conditions. However, HatRT mutants could not exhibit full virulence, highlighting a potentially important role for heme detoxification in *C. difficile* pathogenicity (Knippel *et al.*, 2018). A recent study identified a *C. difficile* heme-inducible operon, HsmRA, which detects heme and contributes towards its detoxification through sequestration. The utilisation of exogenous heme also aided in *C. difficile* resistance to oxidative stress and antibiotics during *C. difficile* infection (Knippel *et al.*, 2020). These studies demonstrated the role of heme toxicity in the pathogenesis of *C. difficile* infection.

1.3 Host responses to *C. difficile* infection

When a host cell interacts with an infecting pathogen, it elicits mechanisms to detect the pathogen and trigger an immune response. The following section highlights our current knowledge of host responses to *C. difficile* infection.

1.3.1 Innate immune response to *C. difficile* infection

1.3.1.1 Pathogen Associated Molecular Patterns (PAMPs)

During microbial colonisation, host cells recognise conserved signatures called Pathogen Associated Molecular Patterns (PAMPs) using Pattern Recognition Receptors (PRRs) which are present on the host cell-surface. The binding of a PAMP to a PRR initiates an innate response against the target pathogen. Common bacterial PAMPs include, flagellin, lipopolysaccharide, peptidoglycan and lipoteichoic acids. A subset of PRRs, the Toll-like Receptors (TLRs), have been reported to recognise PAMPs from *C. difficile* and contribute to the initiation of an innate immune response. The TLR protein, MyD88, is a common adaptor molecule for almost all TLR signalling (except TLR3), which has been strongly associated with the host innate immune response to *C. difficile* (Arancibia *et al.*, 2007; Lawley *et al.*, 2009; Jarchum *et al.*, 2012). MyD88 deletion in mice resulted in an immunocompromised phenotype which had decreased survival rates for *C. difficile* infection, demonstrating the important role of TLR signalling in the immune response against *C. difficile* (Ryan *et al.*, 2011; Jarchum *et al.*, 2012). MyD88 signalling is essential for C-X-C motif ligand 1 (CXCL1)-mediated recruitment of neutrophils to the colonic lamina propria during *C. difficile* infection, a pathway which is impaired in MyD88$^{-/-}$ mice (Jarchum *et al.*, 2012).

MyD88 is a common adaptor molecule for TLR signalling, but specifically, TLR4 has been reported to recognise SLPs from *C. difficile* which induce maturation of bone marrow derived dendritic cells and cytokine production. TLR4$^{-/-}$ mice also exhibited an increased susceptibility to *C. difficile* infection (Ryan *et al.*, 2011). TLR5 has also been reported to recognise flagellin from *C. difficile*, and initiate nuclear factor-κB (NF-κB) and p38 mitogen-activated protein kinase (MAP kinase) activation, as well as interleukin 8 (IL-8) secretion (Jarchum *et al.*, 2011; Yoshino *et al.*, 2013). This response was also elevated following pre-treatment with *C. difficile* toxin B (Yoshino *et al.*, 2013). Interestingly, TLR5 stimulation with flagellin from *Salmonella typhimurium* prior to *C. difficile* infection in mice resulted in a delay in *C. difficile* growth and toxin production. This study highlighted the role of TLR5 in the innate immune response to *C. difficile* colonisation and potentially provides a novel approach for reducing *C. difficile*-induced pathology (Jarchum *et al.*, 2011).

Nucleotide-binding oligomerisation domain 1 (Nod1) is a PRR which recognises peptidoglycan-related small molecules from *C. difficil*e. Nod1 stimulation activates the secretion of the chemokine CXCL1 which induces the recruitment of neutrophils. Nod1$^{-/-}$ mice showed reduced infiltration of neutrophils to the intestines and a high rate of mortality during *C. difficile*-induced colitis. Interestingly, Nod1$^{-/-}$ mice also had an increase in translocation of commensal species from the damaged intestines to other organs due to reduced neutrophil recruitment, which is important for the elimination of translocated commensals induced by *C. difficile* infection (Hasegawa *et al.*, 2011). Translocated commensal bacteria produce lipopolysaccharide (LPS), a TLR4-stimulatory molecule, which

is important for the induction of IL-1β secretion. Since IL-1β is required for CXCL1 production, it has been suggested that a positive-feedback loop of IL-1β and CXCL1 secretion is required for neutrophil recruitment in *C. difficile* infected intestines (Hasegawa *et al.*, 2012).

1.3.1.2 Inflammatory responses

The glucosyltransferase activity of *C. difficile* toxins A and B as well as PAMPs elicit a multifaceted inflammatory response in the infected host (Jafari *et al.*, 2013). An intense acute inflammatory response is evoked and signalling pathways such as NF-κB and MAP kinase are activated (Jefferson, Smith and David, 1999; Warny *et al.*, 2000; Kim *et al.*, 2005; Chae *et al.*, 2006; Chen *et al.*, 2006). NF-κB and MAP kinase activation stimulates the release of IL-8, a potent chemoattractant for neutrophils which has fecal and serum levels correlated with CDI severity (Linevsky *et al.*, 1997a; Steiner *et al.*, 1997; He *et al.*, 2002; Hasegawa *et al.*, 2011; El Feghaly *et al.*, 2013; Rao *et al.*, 2014). An IL-8 gene polymorphism has also been linked to an increased susceptibility to *C. difficile* infection, demonstrating the protective role of IL-8 during *C. difficile*-induced colitis (Jiang *et al.*, 2007). TcdA and TcdB also activate the production of reactive oxygen species (ROS) in the intestinal epithelium. Overproduction of ROS is part of the inflammatory response to defend against bacterial pathogens and their derivatives can cause oxidative damage to DNA and proteins (Frädrich, Beer and Gerhard, 2016; Hinchy *et al.*, 2018). TcdA induces alterations in mitochondrial functions and ROS production which triggers NF-κB-mediated IL-8 production (He *et al.*, 2002; Kim *et al.*, 2005). TcdB also stimulates NADPH oxidase-mediated ROS production at high concentrations (Farrow *et al.*, 2013; Frädrich, Beer and

Gerhard, 2016). The induction of ROS production by both TcdA and TcdB was independent of toxin-mediated glucosyltransferase activity (He *et al.*, 2002; Kim *et al.*, 2005; Farrow *et al.*, 2013).

During *C. difficile* infection an increase in an abundance of a variety of proinflammatory cytokines is observed, including IL-8, IL-1β, IL-23, IL-6, IL-33 and tumor necrosis factor *a* (TNF-*a*) (Linevsky *et al.*, 1997a; Jefferson, Smith and David, 1999; Warny *et al.*, 2000; Sun *et al.*, 2009; Hasegawa *et al.*, 2012; Buonomo *et al.*, 2013; Sun and Hirota, 2015; McDermott *et al.*, 2016; Frisbee *et al.*, 2019), as well as chemokines, such as C-C motif chemokine ligand 5 (CCL-5) and CCL-2 (Kelly and Kyne, 2011; Rao *et al.*, 2014). Early recruitment of immune cells, such as neutrophils, eosinophils and interferon gamma (IFN-γ)-producing type 1 innate lymphoid cells (Abt *et al.*, 2015), as well as leptin (Madan *et al.*, 2014) and IL-33 (Frisbee *et al.*, 2019) have also been associated with protection against *C. difficile* infection.

The severity of *C. difficile* infection has been strongly associated with the immune response which can have an impact on the clinical manifestations and outcome of the disease (Ng *et al.*, 2010). Leukocytosis is a key feature of *C. difficile* infection and contributes towards intestinal injury via the formation of pseudomembranes on the walls of the large intestines which are mostly composed of neutrophils and cellular debris (Savidge *et al.*, 2003). It has also been suggested that intestinal inflammation may present a hostile environment for the gut microbiota, further favouring *C. difficile* outgrowth (Pechine and Collignon, 2016). The anti-inflammatory agent ketotifen has been reported to reduce intestinal injury

during *C. difficile* infection in a rat model, demonstrating the potentially detrimental role of the inflammatory response in the prognosis of the infection (Pothoulakis *et al.*, 1993). However, the opposite effect was observed for the anti-inflammatory drug indomethacin, which increased disease severity in mice (Muñoz-Miralles *et al.*, 2018). Therefore, the use of anti-inflammatory agents in conjunction with antibiotics to treat *C. difficile* infection requires greater investigation before it can be considered as a potential therapeutic strategy.

Immune profiling of CDI patients has identified biomarkers associated with *C. difficile* infections, including many pro-inflammatory cytokines, such as hepatocyte growth factor (HGF), macrophage migration inhibitory factor (MIF), IL-1β, IL-2, IL-5, IL-6, IL-8, IL-10, IL-13, IL-15, IL-16, IL-17A, and TNF-α and CCL-5 (Czepiel *et al.*, 2014; Rao *et al.*, 2014; Yu *et al.*, 2017; Limsrivilai *et al.*, 2018; Abhyankar *et al.*, 2020). An increase in MIF, HGF, IL-1β, IL-2, IL-4, IL-8, IL-15 and IL-16 have been associated with severe CDI (Steiner *et al.*, 1997; Yu *et al.*, 2017; Abhyankar *et al.*, 2020), whereas an increase in IFN-γ, IL-5, CCL-5, EGF, and CCL-4 have been negatively correlated with CDI disease severity (Yu *et al.*, 2017; Abhyankar *et al.*, 2020). Ninety-day mortality rates have also been associated with increased serum concentrations of IL-6, IL-8, IL-15, suppression of tumorigenicity 2 receptor (sST-2) and TNF-α and decreased levels of CCL-5, CCL-4 and epidermal growth factor (EGF) (Abhyankar *et al.*, 2020). A better understanding the molecular drivers of inflammation and biochemical markers associated with disease severity is essential for predicting disease prognosis and developing novel strategies to prevent and treat *C. difficile* infection.

1.3.2 Adaptive immunity induced by *C. difficile* infection

1.3.2.1 Immune response to toxins

C. difficile infection also elicits an adaptive immune response in the host which can provide long-term immunity. TcdA and TcdB are immunologically distinct, despite their homologous structures (Libby and Wilkins, 1982; Leuzzi *et al.*, 2013). Many healthy children and adults possess systemic immunoglobulin G (IgG) and immunoglobulin A (IgA) antibodies to *C. difficile* TcdA and TcdB in their serum, possibly due to the high incidence of asymptomatic colonisation in infants (Rousseau *et al.*, 2012). It has been reported that high TcdB IgG antibody titres correlate with improved clinical recovery without disease relapse (Aronsson *et al.*, 1985; Wullt *et al.*, 2012). CDI patients also produce systemic antibodies to TcdA, but these antibodies did not appear to have an effect on the outcome of the infection (Johnson, Gerding and Janoff, 1992). However, it has also been reported that asymptomatic carriers had significantly higher anti-TcdA IgG antibody concentrations compared to those who developed diarrhoea within the first 3 days of colonisation. This suggests that the production of IgG antibodies against TcdA after exposure to *C. difficile* may play a role in protecting against *C. difficile*-associated colitis (Kyne *et al.*, 2001). Furthermore, another study demonstrated that patients who suffered a recurrence of the infection were lacking the IgG2 and IgG3 subclasses of anti-TcdA IgG antibodies compared to patients who experienced a single episode (Katchar *et al.*, 2007). These studies demonstrated the protective role of systemic IgG antibody production to *C. difficile* toxins.

Secretory IgA antibodies play an important role in mucosal immunity and are produced in digestive tract in response to *C. difficile* infection (Johnson, Gerding and Janoff, 1992). IgA antibodies isolated from stool have been reported to have a neutralising effect on *C. difficile* toxins and inhibited toxin binding to intestinal receptors (Kelly *et al.*, 1992; Warny *et al.*, 1994). It has also been reported that CDI patients who only experienced a single episode with no relapse of disease had higher anti-TcdA IgA antibody concentrations in their faeces compared to patients who suffered a recurrence of the infection (Warny *et al.*, 1994). Low levels of anti-TcdA IgA antibody levels in CDI patient faeces and low numbers of IgA-producing cells in the colon have been associated with prolonged disease and increased incidence of disease recurrence (Johal *et al.*, 2004). Low anti-TcdB IgA titres have also been linked to increased susceptibility to CDI (Islam *et al.*, 2014). These studies highlight the importance of the host mucosal immune response against *C. difficile* toxins and the effects on the outcome of the disease.

1.3.2.2 Immune response to *C. difficile* surface components

CDI patients have also been reported to produce antibodies against several *C. difficile* surface components (Pantosti *et al.*, 1989; Mulligan *et al.*, 1993; Péchiné *et al.*, 2005). Mulligan *et al.* (1993) reported that the levels of immunoglobulins against *C. difficile* cell surface proteins were higher in asymptomatic carriers than symptomatic patients, suggesting antibodies against surface components may impact disease outcome (Mulligan *et al.*, 1993). *C. difficile* SLPs are the most abundant surface proteins and bind to TLR4, inducing a pro-inflammatory innate immune response in the host (Ryan *et al.*, 2011). In addition, the S-layer precursor, SlpA, and the two mature SLPs are immunogenic and induce an

adaptive immune response during *C. difficile* infection (Drudy *et al.*, 2004; Bruxelle *et al.*, 2016). Antibody responses to SLPs were not significantly different between CDI patients, asymptomatic carriers and healthy controls. However, unlike patients who had a single episode, patients who suffered from a recurrence of the disease did not produce an anti-SLP immunoglobulin M (IgM) antibody response (Drudy *et al.*, 2004). These findings suggest that anti-SLP antibodies may contribute towards protection against disease relapses. Other commonly reported immunogenic antigens include FliC, FliD, Cwp84, Cwp66, Fbp68 (Péchiné *et al.*, 2005). The mean levels of serum antibodies to these proteins were significantly lower in CDI patients compared to healthy controls (Péchiné *et al.*, 2005). This suggests that a low antibody levels against cell-surface proteins may facilitate disease and demonstrates the protective role of these antibodies during *C. difficile* infection.

1.3.2.3 Immunotherapy development

The important role of antibody responses to *C. difficile* infection and their impact on disease prognosis suggests that they could potentially be developed as therapies for the prevention and treatment for *C. difficile* infection. The first developed immunotherapies were focused on targeting toxins, such as the CDIFFENSE™ vaccine developed by Sanofi Pasteur as well as the TcdB-targeting monoclonal antibody therapy, Bezlotoxumab developed by Merck & Co. These therapies initially had promising results with decreases in *C. difficile* infection incidence and recurrences in treated individuals. However, neither the vaccine or passive immunotherapy could prevent *C. difficile* colonisation of the gastrointestinal tract and still led to the dissemination of spores into the

environment (Navalkele and Chopra, 2018). Therefore, another approach to block *C. difficile* colonisation may be required to provide more effective immunisation and prevention strategies. Immunotherapies targeting cell-surface components may also have the capacity to influence the outcome of *C. difficile* infection. Passive immunisation with polyclonal antibodies against *C. difficile* SLPs prolonged survival in a hamster lethal model. However, there was no delay the clinical onset of the disease or prevention of death (O'Brien *et al.*, 2005). Mice immunised with FliC-specific hyper-immune serum had a higher rate of survival than non-immunised mice. In the same study, mice and hamsters vaccinated with recombinant FliC as an antigen exhibited a reduction in disease incidence and death rate in a dose-dependent manner. Importantly, immunised mice also shed significantly lower number of spores compared to the non-immunised control group (Ghose *et al.*, 2016).

1.4 Models for *C. difficile* infection

The majority of the *C. difficile* pathogenesis and immune response studies have been conducted using animal models. The development of both *in vivo* and *in vitro* models has greatly expanded our knowledge of the molecular processes which occur during *C. difficile* pathogenesis. This section reviews the current models for the study of *C. difficile* infection.

1.4.1 *in vivo* models

Experimental animal models for the study of infectious diseases have many advantages over clinical studies. These include ready availability, easy standardisation of bacterial inoculums and the ability to perform invasive tests,

such as tissue sampling as well as the ability to examine the effectiveness of novel treatment strategies. The widespread use of animal models has further facilitated the ease of their use, such as a greater availability of reagents for testing (e.g., murine antibodies), relatively cheap prices and ease of handling. Many different animal models have been used to study *C. difficile* infection, including small animals such as hamsters, mice, rats, rabbits, guinea pigs, prairie dogs and quails (Muller, Pitt and George, 1987; Blankenship-Paris *et al.*, 1995; Castagliuolo *et al.*, 1997; Butel *et al.*, 1998; Xia *et al.*, 2000; Alcantara *et al.*, 2001; Chen *et al.*, 2008). There have also been some studies conducted with large animals, including piglets, fouls and monkeys (Arnon *et al.*, 1984; Arroyo, Weese and Staempfli, 2004; Steele *et al.*, 2010). Zebrafish embryos have also been used to investigate the mechanism of *C. difficile* toxins (Hamm, Voth and Ballard, 2006).

1.4.1.1 Hamster model

C. difficile infection in Syrian hamsters (*Mesocricetus auratus*) is induced by the administration of antibiotics to disrupt the intestinal microbiota, and following ingestion of *C. difficile* spores, hamsters display many of the pathophysiological manifestations seen in humans, such as inflammation of the colon, fluid accumulation and pseudomembranous colitis (Fekety *et al.*, 1979). The disease in hamsters is rapidly fatal after a number of days if left untreated. Therefore, in experiments involving potential treatment strategies for *C. difficile* infection, the model is typically focused on the prevention of death. Hamsters have been used to model *C. difficile* infection since the late 1960's, and since then the hamster model has been applied to investigate numerous aspects of *C. difficile* infection, such as strain and toxin characterisation, disease induction, population

dynamics, colonisation capacity, virulence factors and immune responses as well as new prevention and treatment strategies (Chang *et al.*, 1978; Borriello *et al.*, 1987; Sambol *et al.*, 1999; Best, Freeman and Wilcox, 2012; Semenyuk *et al.*, 2015).

1.4.1.2 Murine models

Similar to the hamster model, murine models have also been used extensively to study the pathogenesis of *C. difficile* infection. Several mouse models for *C. difficile* infection have been used, including the conventional mouse model, monoaxenic mice, gnotobiotic mice and the human microbiota-associated mouse model (Butel *et al.*, 1998; Janoir *et al.*, 2013; Collins *et al.*, 2015). Molecular processes such as *C. difficile* colonisation, toxin production, spore germination and biofilm production have been investigated using murine models (Chen *et al.*, 2008; Lawley *et al.*, 2009; Giel *et al.*, 2010; Barketi-Klai *et al.*, 2011; Soavelomandroso *et al.*, 2017). Mouse models overcome several of the drawbacks associated with hamster models, and mouse-specific reagents are more readily available and genetically modified animals are also more easily obtainable (Best, Freeman and Wilcox, 2012).

1.4.2 *In vitro* models

In vitro models have been used to study *C. difficile* since the late 1970s and offer several practical advantages over the use of animals, such as elimination of ethical issues surrounding animal use, as well as easier control and manipulation of the systems (Onderdonk, Lowe and Bartlett, 1979).

1.4.2.1 Fecal emulsion model

Fecal emulsion models involve batch culturing *C. difficile* in faeces from human donors. Borriello and Barclay investigated the behaviour of *C. difficile* in fecal emulsions from healthy donors, patients receiving antibiotics who do not have diarrhoea, patients not receiving antibiotics with diarrhoea, patients with antibiotic associated diarrhoea and patients with diagnosed *C. difficile* infection (Borriello and Barclay, 1986). It was reported that *C. difficile* growth and cytotoxin production was highest in fecal emulsions from *C. difficile*-infected patients and there was no growth in fecal emulsions from healthy donors. The credibility of this model was investigated by comparing the results to a hamster model and a close correlation between the *in vitro* and *in vivo* results was observed (Borriello and Barclay, 1986).

1.4.2.2 Continuous culture systems

Continuous culture systems offer a more analogous model of the human gut than batch culturing models such as fecal emulsions. Continuous culture systems not only provide a more accurate representation of the human gut, but also facilitate controlled manipulation and monitoring of the environment and allow the experiment to continue over extended time periods. In 1979 a continuous culture model was used to investigate the effects of environmental stress on *C. difficile* toxin levels. It was found that changes in oxidation and reduction levels (Eh), temperature and sub-inhibitory concentrations of vancomycin and penicillin affected *C. difficile* toxin production (Onderdonk, Lowe and Bartlett, 1979). This study suggested that a number of factors control *C. difficile* toxin secretion and

demonstrated the potential of continuous culture systems towards the study of anaerobic gut bacteria.

1.4.2.3 Triple-stage chemostat models

The human colon contains hundreds of different species of bacteria in a complex ecosystem. *In vitro* triple-stage chemostat models simulate the conditions of the human gut and have been used to culture microbiota species and can therefore be adapted for the study of *C. difficile* infection (Macfarlane, Macfarlane and Gibson, 1998; Freeman *et al.*, 2005). This model involves the use of three vessels organised in a weir cascade system in an oxygen-free nitrogen atmosphere. Growth medium is fed into the top of the system at a controlled rate, and each vessel operates at a controlled pH to reflect the increasing pH along the digestive tract. Pooled emulsified fecal samples are incubated in the system for 2 weeks for bacterial populations to equilibrate (Macfarlane, Macfarlane and Gibson, 1998). Using this system as a model for the human gut, *C. difficile* infection can be simulated by treating the system with clindamycin, then introducing *C. difficile* spores into it. This gut model has been used to test antimicrobials for their ability to induce *C. difficile* infection (Freeman, O'Neill and Wilcox, 2003; Baines, Freeman and Wilcox, 2005; Baines *et al.*, 2006, 2009; Moura *et al.*, 2019), the efficacies of treatments (Freeman *et al.*, 2005, 2007; Baines *et al.*, 2008, 2009), and relative fitness of various *C. difficile* isolates (Freeman *et al.*, 2007). The triple-stage chemostat gut model was also adapted to study multispecies biofilms of gut microbes by incorporating glass rods into a fermentation chamber. Biofilm structures of bacteria from fecal emulsion samples were allowed to form before *C. difficile* spores were added to the system. Consistent with previous studies, it

was demonstrated that *C. difficile* spores only germinate at low levels in the absence of antimicrobial intervention. However, the *C. difficile* spores persisted in the biofilm structures preferentially compared to planktonic communities. This study demonstrated the use of the adapted triple-stage chemostat model as a tool for studying intestinal multispecies biofilm communities and their role in *C. difficile* infection (Crowther *et al.*, 2014).

1.4.3 Cellular gut models

Fecal emulsion-based models have improved our knowledge of the role of antibiotics in *C. difficile* induction and treatment, as well as the exploration of microbiota population dynamics. However, host-specific factors such as immunological or secretary events cannot be represented. A major obstacle to studying host-anaerobe interactions arises from the difficulty of culturing strictly anaerobic microbes with an oxygen-requiring gut epithelium. This has resulted in host-*C. difficile* interactions at a molecular level being understudied. However, multiple *in vitro* cellular gut models have been developed to study host-pathogen interactions of *C. difficile*.

1.4.3.1 Epithelial layer models

A commercially available systems used to mimic the aerobic/anaerobic interface of the human gut is a vertical diffusion chamber (VDC). The VDC consists of two chambers which act as compartments for the aerobic and anaerobic gaseous environments. Cell layers are incubated at the interface of the two chambers which facilitates the co-culture of anaerobic bacterial species with human cells. The trans-epithelial electrical resistance (TEER) is commonly used in parallel to

monitor the integrity and permeability of the epithelial cell layers along the course of infection. Quantification of the TEER is a useful tool for measuring the severity of the host cell damage as a result of bacterial infection. Schüller *et al.* (2010) used the VDC system to investigate host-pathogen interactions of enterohaemorrhagic *Escherichia coli* (EHEC). It was reported that under anaerobic conditions EHEC colonisation was enhanced, type III secretion system (T3SS) effector protein expression and translocation were increased in the presence of host cells in the VDC system (Schüller and Phillips, 2010). Host-pathogen interactions of *C. difficile* have also previously been studied using the VDC system. Jafari *et al.* (2016) investigated the role of toxins in host-pathogen interactions using the VDC system. It was reported that the toxins did not have an effect on *C. difficile* adherence to host cells, but caused tight junction disruption and a reduction in TEER values (Jafari *et al.*, 2016).

2D epithelial layer models are a simple method for improving our understanding of host-pathogen interactions. However, the insights which can be established with these models are limited, as the human gut is home to a variety of cell-types which are organised in a multi-layer structure. Interactions between host cells can occur through direct cellular connections such as gap junctions, tight junctions, and desmosomes, as well as through secreted signalling molecules (Harris and Tepass, 2010). The large intestine holds a dense mucus layer which prevents inflammation by shielding the underlying epithelial layer from luminal bacteria and food antigens. Mucus in the large intestine is mostly made up from the mucin MUC2, a large net-like polymer which is secreted by goblet cells (Shan *et al.*, 2013). The MUC2 mucin contains numerous *O*-glycans which act as a

nutrient source and attachment site for the microbiota (Sonnenburg *et al.*, 2005). To reach the epithelial cells in the colon, gastrointestinal pathogens must develop strategies to penetrate the inner mucus layer. Fibroblast cells reside in the basement layer of the intestines and play a role in the regulation of epithelial cell proliferation and differentiation through the production of signalling molecules such as growth factors and pro-inflammatory cytokines (Smith *et al.*, 1997; Göke, Kanai and Podolsky, 1998; Skibinski, Elborn and Ennis, 2007; Higuchi *et al.*, 2015). Intestinal fibroblast cells differentiate into myofibroblasts which play a central role in inflammatory responses in the gut (Higuchi *et al.*, 2015). 3-dimentional (3D) multicellular models can be used as a valuable tool to replicate human epithelial cell layers for the study of host-pathogen interactions. 3D models have been shown to encourage better growth and differentiation of epithelial layers, where proteins such as growth factors are secreted by an underlying fibroblast cell layer (Morris *et al.*, 2014). Electrospinning can be employed to generate a fibrous network from polymers with a similar structure to the ECM (Lu, Li and Chen, 2013). A 3D model was recently applied to investigate host-pathogen interactions of *C. difficile* by Anonye *et al.* (2019), where electrospun fibres made from polyethylene terephthalate (PET) were used in conjunction with the VDC system (3D-VDC) to model the human gut (Figure 5). It was reported that *C. difficile* exhibited a greater adhesion to host epithelial cells in the 3D-VDC system compared to the conventional 2-dimentional VDC system (2D-VDC) (Anonye *et al.*, 2019).

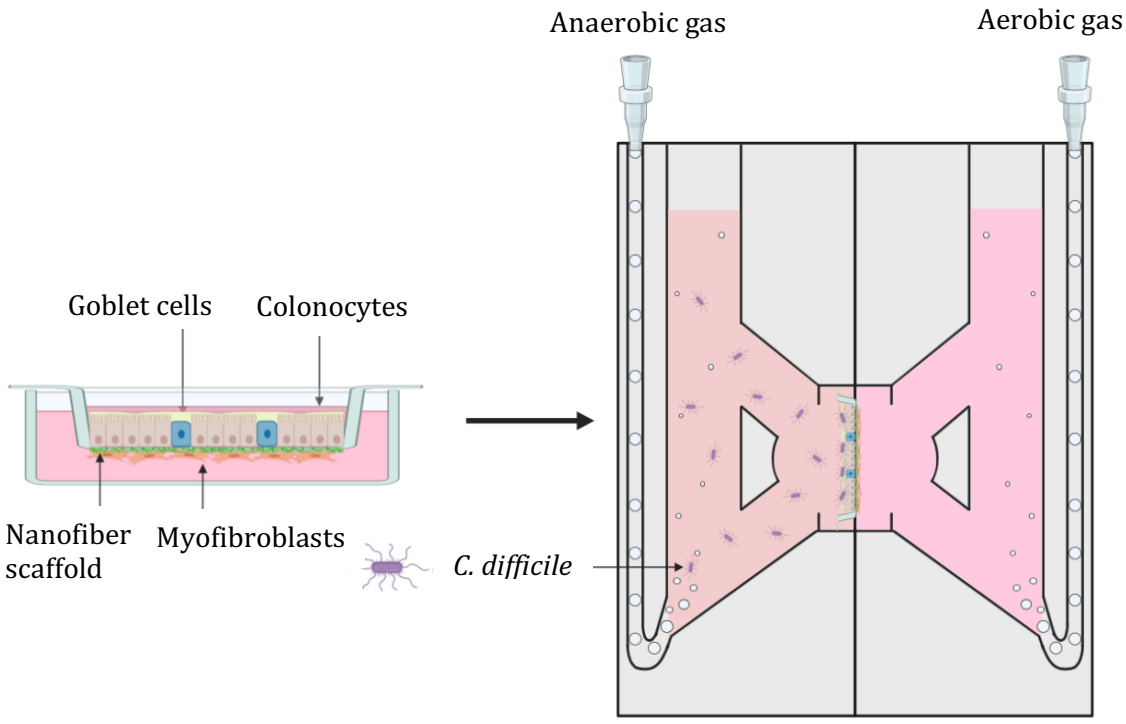


Figure 5. A 3-dimensional vertical diffusion chamber (3D-VDC) model used to model *C. difficile* infection.

1.4.3.2 Organoid models

Shaban *et al.* developed a bioengineered 3D human intestinal tissue model using tubular silk biomaterial scaffolds. The hollow chamber of the silk scaffolds was seeded with Caco-2 and HT29 intestinal epithelial cells and the outer surface was seeded with myofibroblast cells. This model generated an anaerobic environment in the lumen which could support *C. difficile* spore germination and outgrowth over a 48 hour timecourse. Toxin activity and epithelium damage were increased in the 3D model compared to a 2D transwell model (Shabana *et al.*, 2018).

1.5 Transcriptomic responses during *C. difficile* infection

A holistic view of cell responses is important to learn about infection in *in vitro* and *in vivo* models. A transcriptome of a cell is considered to be all of the genes which are expressed and their transcriptional levels. The study of transcriptomic profiles can resolve the overall modification of gene expression in response to environmental changes and can therefore contribute towards an improved understanding of the molecular processes which occur in cells. Infection initiates a cascade of events which result in altered host and bacterial gene expression. These changes lead to adaptations from the host to protect against the pathogen, and from the pathogen to resist clearance from the host immune response and cause disease. A greater understanding of transcriptomic responses to infection can identify host cell pathways which respond to exposure to pathogens as well as new virulence factors of the invading pathogen (Westermann, Gorski and Vogel, 2012). Since host-anaerobe interactions are challenging to study, previous characterisation of *C. difficile* infection has been focused on the pathogen, the host, or specific aspects of the infection, such as the pathophysiological and transcriptomic responses to *C. difficile* toxins. In this section the current knowledge of transcriptomic responses to *C. difficile* infection will be reviewed.

1.5.1 Host transcriptomic profiles

D'Auria *et al.* investigated the effect of *C. difficile* toxins Tcd1 and Tcd2 on the transcriptome of intestinal epithelial cells using microarray analysis in a murine model. Inflammation-associated genes (such as C3, CCXL1, CXCL10, DUSP1 and EGR1) played a key role in the response to Tcd1, Tcd2 and a combination of both toxins. The temporal changes in selected inflammation-associated genes

suggested that Tcd1 and Tcd2 both induce the upregulation of inflammation-associated genes, but Tcd1-induced changes were on average 2 or 3 times higher at 6 h and 16 h. Many other genes, such as GTP binding proteins or GTPases were also differentially expressed in the presence of *C. difficile* toxins (D'Auria *et al.*, 2013).

A similar study investigated whole-tissue transcriptional responses to vegetative *C. difficile* R20291 infection using a microarray-based approach in a mouse model. It was reported that the IL-1 cytokine family member, IL-33, was upregulated during *C. difficile* infection. IL-33 is transcriptionally regulated following the recognition of PAMPs by PRRs and activates of group 2 innate lymphoid cells which may provide protection against *C. difficile* infection (Polumuri *et al.*, 2012). IL-33 inhibited *C. difficile*-associated mortality via reduced activation of group 2 innate lymphoid cells (ILC2s). Finally, FMT therapy after antibiotic-mediated microbial depletion replenished IL-33 expression in the colon, suggesting that IL-33 secretion may be driven by the microbiome. This study identified IL-33 as a driver of lymph cell-mediated immune responses to *C. difficile*-associated colitis. There were no more significantly differentially expressed host genes discussed in this publication (Frisbee *et al.*, 2019).

A recent study by Fletcher *et al.* examined the host transcriptomic responses to WT and *ΔtcdR C. difficile* in a mouse model to investigate the effects of toxin-induced intestinal inflammation on host and bacterial gene expression. Gene Ontology (GO) pathways which were differentially expressed during toxin-induced inflammation included the regulation of the inflammatory response,

peptide secretion and proteolysis. Interestingly, an upregulation in numerous matrix metalloproteinases (MMPs) transcripts was also observed during toxin-induced intestinal inflammation, suggesting that the host ECM may be altered during *C. difficile* infection. It was demonstrated that collagen was degraded during incubation of fibroblast cells with TcdA and TcdB, and toxin-mediated degradation of collagen enhanced *C. difficile* growth *in vitro*. This study demonstrated how toxin-induced inflammation may alter the host gastrointestinal environment to support *C. difficile* pathogenesis (Fletcher *et al.*, 2021).

1.5.2 Bacterial transcriptomic profiles

There have been several studies focused on the transcriptomic profile of *C. difficile* during infection to identify genes which facilitate colonisation and pathogenesis, as well as achieve a better understanding of the molecular events involved in infection. Scaria *et al.* used a pig ligated-loop model to perform an *in vivo* microarray-based analysis of *C. difficile* responses to infection. Pigs were fasted overnight before injection of *C. difficile* vegetative cells into the ligated loops. RNA was extracted at three timepoints, 4 h, 8 h and 12 h to track the temporal changes in gene expression, as well as from log-phase bacterial cultures which were used as uninfected controls. An upregulation of the toxin A gene, *tcdA*, was observed at 12 h post infection. Many other colonisation and virulence-associated genes, such as CD2830 (*PPEP-1*), CD2592 (fibronectin binding protein), CD2793 (*slpA*), CD1546, CD1208 (hemolysins), and genes involved in the sporulation cascade were also upregulated early in infection. The majority of differentially expressed genes were involved in processes such as transcription,

signal transduction, amino acid transportation and/or metabolism, and carbohydrate transportation and/or metabolism. Interestingly, *codY*, a global regulator of gene expression and suppressor of several virulence factors such a toxin production and sporulation was found to be upregulated at 4 h and 8 h post infection(Dineen *et al.*, 2007; Nawrocki *et al.*, 2016). ABC transporters, flagella genes and several sigma factors were also found to be upregulated (Scaria *et al.*, 2011). This study identified several proteins and pathways important for *C. difficile* colonisation and virulence in pigs.

Janoir *et al.* (2013) investigated the adaptive strategies *C. difficile* 630 employs over the course of infection using a microarray-based approach in a germfree mouse model. The majority of gene expression changes were involved in processes such as metabolism (fermentation and amino acid and lipid metabolism), cell processes, stress response, pathogenicity, and sporulation. Several genes with potential roles in *C. difficile* colonisation, such as *slpA*, and its paralogs *cwp19*, CD0440 and CD2796 were significantly upregulated at the early stage of infection (14 h) and downregulated at the late stage of infection (38 h), suggesting that *C. difficile* may modulate its cell-surface at different points over the course of the infection. The toxin genes *tcdA* and *tcdB* were both moderately upregulated at the later stages of infection. These findings suggest that genes involved in colonisation may be modulated in the initial stages of infection before gene expression shifts to favour the upregulation of virulence associated genes at the later stages. Interestingly, almost no modulation of the three flagella assembly operons was observed in this analysis, apart from a moderate downregulation of *fliC*. This major flagellin subunit has been previously reported

to have adhesive properties and may play a role in *C. difficile* colonisation (Tasteyre *et al.*, 2001; Dingle, Mulvey and Armstrong, 2011). This result is in accordance with previous findings suggesting that motility may not be required for *C. difficile* virulence (Stabler *et al.*, 2009; Dingle, Mulvey and Armstrong, 2011). There was also a rapid induction of the sporulation cascade in the early stages of infection. However, there was no observed differential expression of *spo0A*, a major regulator of the initiation of sporulation. It was confirmed with RT-qPCR that *spo0A* expression was stable over the measured conditions, which may explain why it was not significantly differentially expressed during the differential gene expression analysis. The most highly expressed gene throughout the duration of the experiment was CD1581, a gene which codes for a protein of unknown function. Interestingly, genetic deletion of CD1581 resulted in a strain with reduced fitness and colonisation capacity. Therefore, this study also identified CD1581 as a potential new colonisation factor of *C. difficile* (Janoir *et al.*, 2013).

A study analysing the nutrients which *C. difficile* VPI 10463 requires for colonisation and pathogenesis was performed using an RNA-seq approach on bacterial RNA isolated from infected mice ceca at early and late timepoints after infection. Carbohydrate, amino acid and fatty acid uptake and metabolism were upregulated early in the process of susceptible host colonisation. As well as genes involved in the fatty acid metabolism and butyrate production pathways, the top upregulated genes at later stages of infection also included genes such as *feoB*, a ferrous iron importer. This implies that iron uptake may play a role in *C. difficile* pathogenesis and iron scarcity may be a nutritional signal for *C. difficile* during

infection. Differential expression of several proteases and peptidases was also observed, including the upregulation of Clp, aminopeptidase and peptidase S41, as well as the downregulation of PPEP-1 (Fletcher *et al.*, 2018). This study highlighted the nutrients utilised by *C. difficile* during early and late stages of colonisation.

A recent study by Fletcher *et al.* investigated the *C. difficile* responses to toxin-mediated intestinal inflammation with a microarray-based approach. This was conducted by comparing the transcriptomic profiles of WT and *ΔtcdR*, which evoked significantly less inflammation during infection of a mouse model. The results showed that toxin-mediated intestinal inflammation significantly altered the *C. difficile* transcriptomic profile during infection. Inflammation of the gut stimulated the upregulation of many metabolic pathways, including butyrate metabolism, branched-chain amino acid biosynthesis, amino acid metabolism, and lipid metabolism were significantly modulated during intestinal inflammation. Other modulated processes which may have a role in the pathogenesis of *C. difficile* including the flagella and multidrug/antibiotic ABC transport were also modulated in the presence of intestinal inflammation (Fletcher *et al.*, 2021). This study demonstrated how the host inflammatory responses may provide an environment which favours *C. difficile* growth during infection.

1.5.3 Dual transcriptomics of *C. difficile* infection

Janvilisri *et al.* investigated the transcriptional responses of Caco-2 cells and *C. difficile* simultaneously during infection using a microarray-based analysis. Caco-

2 cells were incubated with *C. difficile* for 30, 60, and 120 mins in anaerobic conditions. 271 Caco-2 genes and 207 *C. difficile* genes were identified as significantly differentially expressed compared to uninfected controls in at least one time point of infection. Differentially expressed human genes were involved in metabolic processes such as nucleic acid, protein and lipid metabolism, cell organisation and biosynthesis, transport, cell communication, signal transduction, and transcription. Differentially expressed *C. difficile* genes included those involved in nucleic acid metabolism, transcription, protein synthesis and protein modification. The results of this analysis also suggested that a bacterial stress response occurs during Caco-2 infection. Interestingly, there was an upregulation in *codY*, a repressor of toxin transcription, and no differential expression of the toxin genes in this analysis. A notable flaw of the experimental setup in this investigation is the relatively short timepoints which may not be representative of *C. difficile* infection in humans. Later timepoints could not be investigated, probably due to the reducing viability of human cells in anaerobic conditions as the time period extended (Janvilisri, Scaria and Chang, 2015).

The majority of these studies have relied on probe-based methods (such as microarrays) to investigate host and pathogen transcriptomic responses to infection. These methods use specialised host and/or pathogen probes and produce an array of fluorescent signals which relate to the expression levels of each selected gene. However, there are several limitations associated with these techniques for the study of host and pathogen transcriptional profiles, including the requirement of host and pathogen RNA to be physically separated prior to

the analysis, risks of cross-hybridisation and it is limited to include only known, coding transcripts (Westermann, Gorski and Vogel, 2012). The development of RNA sequencing (RNA-seq) methodologies overcomes many of the drawbacks of probe-based methods for the study of host-pathogen transcriptomics.

1.5.4 Dual RNA sequencing for host-pathogen transcriptomics

RNA-seq is a powerful tool for understanding variations in gene expression profiles and can be useful for studying host-pathogen interactions. RNA-seq can greatly expand our knowledge of cellular functions and gene expression profiling can reveal vast information about the physiological status of a cell. The majority of previous transcriptional analyses of infectious diseases have predominantly been focused on either the pathogen or the host. Dual RNA-seq is a sequencing methodology where the transcriptome of a pathogen and infected host cells are sequenced simultaneously during a host-microbe interaction. This enables the resolution of the global transcriptomic profiles of both host and pathogen during infection and can identify genes and pathways which are relevant for interspecies interactions. The host and pathogen RNA are typically separated *in silico* rather than physically, which removes unwanted biases which may occur during separation procedures (Westermann, Gorski and Vogel, 2012; Westermann, Barquist and Vogel, 2017).

Unlike probe-based methods, dual RNA-seq sequences both coding and non-coding transcripts which allows the detection of novel and hypothetical genes. Therefore, this technique has the potential to greatly expand our understanding of biological processes which facilitate infections (Westermann, Barquist and

Vogel, 2017). Dual RNA-seq has previously been applied to a variety of parasitic, commensal, and mutualistic relationships using *in vitro* models as well as more complex systems (Westermann, Barquist and Vogel, 2017). Transcriptomic responses of numerous bacterial pathogens and infected eukaryotic cells have been resolved with dual RNA-seq, including uropathogenic *Escherichia coli* (UPEC) (Mavromatis *et al.*, 2015), *Mycobacterium tuberculosis* Bacillus Calmette–Guérin (BCG) (Rienksma *et al.*, 2015) and *Salmonella enterica* serovar Typhimurium (Avraham *et al.*, 2016; Westermann *et al.*, 2016).

One of the main challenges associated with performing a dual RNA-seq experiment is caused by the large differences in genome size between eukaryotic and prokaryotic cells which translates to the quantity of cellular RNA. The human haploid genome is 3,000 Mb in size, whereas the genomes of typical pathogens such as *Escherichia coli* or *Salmonella enterica* are ~5 Mb. To achieve accurate representation of both species there are several steps which are recommended prior to sequencing. Firstly, to achieve a uniform population of infected cells, it is recommended to use a sorting technique such as fluorescence-activated cell sorting (FACS) to remove uninfected cells from the analysis. Secondly, ubiquitous rRNA represents the most abundant class of RNA transcripts but contains little useful information for transcriptomic studies. Depletion of rRNA transcripts is strongly recommended for dual RNA-seq experiments to increase the concentration of informational content within the samples. Thirdly, in cases where the large abundance of host RNA overwhelms the mixed RNA samples, poly(A)-based enrichment is recommended to deplete eukaryotic RNA and

enrich bacterial transcripts (Westermann, Gorski and Vogel, 2012; Westermann, Barquist and Vogel, 2017).

The first dual RNA-seq experiment was performed by Westermann *et al.* to investigate how intracellular *Salmonella* uses sRNAs to regulate bacterial gene expression, as well as the host cell responses to infection. A dual RNA-seq of *Salmonella enterica* serovar Typhimurium-infected HeLa cells was performed to profile the host and pathogen global transcriptomic responses to infection simultaneously. This study established the molecular impact of bacterial riboregulators on both the bacterial and host cells. Specifically, a small non-coding RNA, PinT, was activated upon bacterial internalisation and temporally controlled the expression of virulence genes and invasion-associated effector proteins required for intracellular survival. PinT also had an impact on the host cell transcriptional response to infection and a model was proposed in which *Salmonella* manipulates key host cell pathways to favour its replication using PinT-mediated control of virulence genes (Westermann *et al.*, 2016).

A dual RNA-seq analysis of *Plasmodium*-infected hepatocytes was conducted, using fluorescence-activated cell sorting to isolate infected host cells. The most highly downregulated host gene was energy homeostasis associated gene (*ENHO*), a positive regulator in the Notch signalling pathway. A high level of downregulation was also observed in the chemokine ligand, *CXCL10*, an important protein involved in Janus Kinase/Signal Transducer and Activator of Transcription (JAK-STAT), as well as the interferon alpha and gamma signalling pathways. Upregulated pathways include energy homeostasis, adipocytokine

signalling and gluconeogenesis. The most highly upregulated human gene was the human mucosal innate immunity gene, *MUC13* (LaMonte *et al.*, 2019). MUC13 is a transmembrane mucin glycoprotein which is highly expressed on the surface of mucosal epithelial cells in the small and large intestines (McGuckin *et al.*, 2011). As well as providing a physical barrier to potential pathogens, MUC13 also plays an important role in the regulation of inflammatory cytokines and has been implemented in the immune response against gastrointestinal pathogens such as *Helicobacter pylori* (Moehle *et al.*, 2006; Sheng *et al.*, 2013; Liu *et al.*, 2014). This gene was also upregulated during dextran sodium sulfate (DDS)-induced colitis and has been associated with colon cancer (Maher *et al.*, 2011; Sheng *et al.*, 2011, 2013). This dual RNA-seq analysis identified MUC13 as a new host factor involved in the immune response against parasites, and could be a biomarker candidate for hepatic parasite infections (LaMonte *et al.*, 2019).

Another recent application of dual RNA-seq was in *Mycobacterium tuberculosis*, an intracellular pathogen which employs strategies to withstand the harsh environment of phagosomes, as well as actively manipulates macrophages and dendritic cells to prolong intracellular survival (Gengenbacher and Kaufmann, 2012). A better understanding of the crosstalk between the host and *Mycobacterium tuberculosis* is vital for the development of novel strategies against the infection. The attenuated TB vaccine strain *Mycobacterium bovis* BCG is commonly used as a model for *M. tuberculosis* due to the high degree of similarities in the genome. Rienksma *et al.* performed a dual RNA-seq experiment to investigate the transcriptional adaptation of *M. bovis* BCG and THP-1 human macrophage-like cells 24 h after infection. The bacterial transcripts were

enriched using differential lysis with guanidine thiocyanate to overcome the unfavourable bacterium-to-host RNA ratio due to the relatively low numbers of bacteria per host cell. *M. bovis* exhibited a clear transcriptional response to THP-1 infection with 367 genes being differentially expressed, of which 216 were upregulated and 151 were downregulated. Bacterial pathways which were upregulated included cholesterol degradation, mycobactin siderophore biosynthesis and aspartite utilisation. The THP-1 cell transcriptome exhibited increased expression of several glycolytic enzymes and the induction of ketogenesis pathway genes. The host cell also induced the cholesterol biosynthesis pathway and iron uptake and transport pathways, possibly to compensate for the loss of cholesterol and iron to the pathogen. This study successfully used dual RNA-seq to elucidate the complex interactions between *M. bovis* BCG and THP-1 macrophage cells (Rienksma, Suarez-Diez *et al.* 2015). An overview of selected published dual RNA-seq studies is illustrated in Table 1.

Table 1. Overview of selected published dual RNA-seq studies.

	Westerman *et al.* (2016)	**Rienksma *et al.* (2015)**	**Mavromatis *et al.* (2015)**	**Marsh *et al.* (2019)**
Bacterial species	*Salmonella Typhimurium*	*Mycobacterium tuberculosis Bacillus Calmette-Guerin*	*Uropathogenic Escherichia coli* (UPEC)	Chlamydia trachomatis serovae E

Host model	Mouse bone marrow-derived macrophages	Human monocytic cells	Mouse bone marrow-derived macrophages	Human epithelial cells (Hep-2)
Enrichment of bacterial transcripts	With and without polyA-depletion of host transcripts	With or without differential lysis	MICROB*Enrich*	With and without polyA-depletion of host transcripts
rRNA depletion	RiboZero (epidemiology)	RiboZero (epidemiology)	RiboZero (gram negative bacteria; human/ mouse/ rat)	RiboZero (gram negative bacteria; human/ mouse/ rat)
Sequencing platform	HiSeq; NextSeq 500 (single end)	HiSeq 2500 (paired end)	HiSeq 2000 (paired-end)	GA IIx (paired-end)
Fraction of bacterial reads	~1%-10%	~2-4%	0.03%-58%	~5%

1.6 Experimental aims

C. difficile interactions with the colonic mucus layer and the gut tissue ECM remain largely unexplored. The majority of previous studies investigating the

transcriptomic responses to *C. difficile* infection have relied on probe-based methods and have been focused on the pathogen, the host, or specific aspects of the infection such as the pathophysiological and transcriptomic response to purified *C. difficile* toxins. At present, there has not been an investigation into the host or bacterial differential gene expression profiles during *C. difficile* colonisation which has used RNA sequencing and the molecular interplay between *C. difficile* and the intestinal epithelium remains unclear. The main aim of this project was to investigate the molecular mechanisms of *C. difficile* colonisation and pathogenesis using dual RNA-seq to resolve host and bacterial global transcriptomic responses to *C. difficile* infection. I also set out to confirm the expression levels of selected identified differentially expressed genes using molecular biology techniques, including RT-qPCR, western blotting and confocal microscopy. Finally, I aimed to investigate the role of bacterial and host factors modulated during infection to identify novel proteins and pathways important for *C. difficile* colonisation and pathogenesis. A better understanding of *C. difficile* interactions with the gut interface is crucial for developing novel therapeutics, effective preventative measures and improved diagnostic tools for *C. difficile* infection.

Chapter 2: Methods

2.1 Bacterial strains and growth conditions

C. difficile 630 and B1/NAP1/027 R20291 (isolated from the Stoke Mandeville outbreak in 2004 and 2005) were primarily used in this study. *ΔPPEP-1* mutant was provided by Novarits, Italy. *Clostridioides difficile* strains were streaked from glycerol stocks and cultured using brain-heart infusion (BHI) agar or broth (Sigma Aldrich, USA) supplemented with 1g/L L-cysteine (Sigma Aldrich, USA) and 5g/L yeast extract (Sigma Aldrich, USA) (BHIS) under anaerobic conditions (80% N_2, 10% CO_2, 10% H_2) in a Don Whitley workstation (Yorkshire, UK).

2.2 Cell culture, media and conditions

Intestinal epithelial cell line, Caco-2 (P6-P21) from American Type Culture Collection, mucus producing cell line, HT29-MTX (P45-P60), a gift from Nathalie Juge, Quadram Institute, Norwich, and intestinal myofibroblast cells, CCD-18co (P10- P20) were used in this investigation. Caco-2 cells were grown in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% FBS (DMEM-10) (Labtech, UK), and 1% penicillin-streptomycin (10,000 units/mL penicillin, 10 mg/mL streptomycin, Sigma Aldrich, USA). HT29-MTX were grown in DMEM-10 and CCD-18co in Eagle's Minimum Essential Medium media, both supplemented with 10% FBS, 1% penicillin-streptomycin, 2 mM glutamine, and 1% non-essential amino acids (Sigma Aldrich, USA). All cell lines were maintained in 5% CO_2 in a humidified incubator at 37°C and free from mycoplasma contamination as determined on a regular basis by the EZ-PCR Mycoplasma kit (Biological

Industries, USA). Cell lines were frozen in media, FBS and Dimethyl sulfoxide (DMSO) in a 5:4:1 ratio and stored in liquid nitrogen.

Snapwell inserts (Scientific Laboratory Supplies Ltd, UK) were prepared as reported by Anonye *et al.* (2019). Briefly, prior to seeding cells, the Snapwell inserts were coated with a 1:1 ratio of rat tail collagen (Sigma Aldrich, USA) and ethanol (VWR Chemicals, USA). Caco-2 and HT29-MTX were mixed in a 9:1 ratio and 2 x 10^5 cells total was seeded on 12 mm Snapwell inserts (tissue culture treated polyester membrane, Corning, USA) for ~2 weeks to form a polarised monolayer. Caco-2 and HT29-MTX were seeded on the apical side of the Snapwell insert for 14 days. Prior to infection experiments, the cell culture medium in the Snapwell inserts was replaced with antibiotic-free medium 24 h or 48 h before the start of infection.

2.3 Vertical diffusion chamber setup

The Snapwell inserts containing the polarised cell layers were placed between the two half chambers of the VDC (Harvard Apparatus, Cambridge, UK) and sealed with the provided clamps. 3 mL DMEM-10 was added to both sides to fill the chamber.

2.4 Infection of intestinal epithelial cells in the VDC model

A single bacterial colony was inoculated in prereduced BHIS broth and incubated at 37°C for 16 h in anaerobic conditions. The culture was centrifuged at 5,000 rpm for 5 min (Eppendorf 5810R, Eppendorf, Germany) and bacterial pellet was resuspended in DMEM-10, before being diluted to an OD_{600} of 1.0 and incubated

at 37°C in anaerobic conditions for 1 hour. Intestinal epithelial cells were infected with bacterial cultures at an MOI of 100:1 in the apical chamber. This chamber was supplied with anaerobic gas mixture (10% CO_2, 10% H_2, 80% N_2, BOC, UK) and the basolateral compartment with 5% CO_2 and 95% air (BOC, UK) at a rate of ~1 bubble every 5 seconds. At 3 h post infection, the apical media containing the *C. difficile* was removed, the chamber was washed once with PBS and 3 mL fresh prereduced DMEM-10 was added. Chambers were incubated for a further 3 - 45 h. The intestinal epithelial cell layers were washed thrice with prereduced PBS before being lysed with 1 mL sterile water. The number of cell-associated bacteria were quantified by performing serial dilutions and colony forming unit (CFU) counts from the intestinal epithelial cell lysate, prepared using serial dilutions plated on BHIS agar.

2.5 RNA isolation and treatments

Bacterial and human cells were treated with RNAprotect cell or bacterial reagent (Qiagen, Germany) and stored at -80°C for up to 1 week. To extract RNA, samples were thawed and centrifuged at 4°C (Sigma Zentrifugen 1-14K, Germany). Cell pellets were resuspended in 1 mL buffer RLT from the RNeasy mini kit (Qiagen, Germany) with a 1:200 dilution of β-mercaptoethanol (Sigma Aldrich, USA). Cells were homogenised using lysing matrix B tubes with 0.1 mm silica beads (MP Biomedicals, USA) in a FastPrep-24 5G (6.5 m/s for 20 seconds with 180 seconds rest on ice for 6 cycles) (MP Biomedicals, USA). RNA was extracted using the RNeasy RNA isolation kit, (Qiagen, Germany), according to the manufacturer's protocol. A rigorous treatment with TURBO DNase (Thermo Fisher Scientific, USA) was used to remove genomic DNA contamination from RNA samples and

clean up was performed using the RNeasy mini kit following the manufacturer's protocol (Qiagen, Germany). RNA concentrations were quantified using the Qubit RNA BR assay kit (Thermo Fisher Scientific, USA). The RNA quality was examined using the Bioanalyzer RNA 6000 pico kit (Agilent, USA).

2.6 Ribosomal RNA depletion, library preparation and RNA sequencing

For the pilot study, the Ribo-Zero Gold (epidemiology) kit was used to deplete rRNA from RNA samples, following the manufacturer's protocol (Illumina, USA). For the uninfected controls and later study, the RNase H method for rRNA depletion was used as described by Adiconis *et al.* (2013). The human rRNA single-stranded DNA probes described in this thesis were purchased as oligonucleotides (Integrated DNA technologies, USA). *C. difficile* probes complementary to the 16S, 23S and 5S rRNA sequences were designed as ~50 bp non- overlapping primers (Integrated DNA Technologies, USA) (Appendix table 1). RNase H rRNA depletion protocol was used as described by SciLifeLab (Sweden). Briefly, oligonucleotide probes were hybridised to RNA samples by incubation with hybridisation buffer (1 M NaCl, 0.5 M Tris-HCl pH 7.4) in a Bio-Rad T100™ thermocycler (95°C for 2 min, temperature decreased -0.1 C/s to 45°C and hold at 45°C). Thermostable RNase H (New England Biolabs, USA) was added, and the reaction was incubated at 45°C for 30 min. Agencourt® AMPure® XP magnetic beads were used to purify RNA. DNase I (New England Biolabs, USA) was used to remove leftover DNA oligonucleotides and Agencourt® AMPure® XP magnetic beads (Beckman Coulter Life Sciences, USA) were used to clean up rRNA depleted samples. Bioanalyzer RNA pico analysis was used to assess the RNA quality and effectiveness of the rRNA depletion (Agilent, USA). The

NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina® with the NEBNext® Multiplex Oligos for Illumina® (Index Primers Set 1) were used to generate DNA libraries for sequencing (New England Biolabs, USA) as per the manufacturers protocol. Agencourt® AMPure XP magnetic beads (Beckman Coulter Life Sciences, USA) were used for bead-based DNA clean up. Quality of DNA libraries were assessed using a Bioanalyzer High Sensitivity DNA kit (Agilent, USA). Single-end sequencing was performed using a NextSeq® 500/550 High Output Kit v2 (75 cycles) on an Illumina NextSeq™500 system (Illumina, USA). Bacterial controls were sequenced using single-end sequencing with a MiSeq™ v3 cartridge (150 cycles) on an Illumina MiSeq™ system.

2.7 RNA-seq analysis

Raw sequencing files were converted to FASTQ files using the bcl2fastq linux software package. The quality of the FASTQ files was assessed using the FastQC Linux package. FASTQ files were mapped to the appropriate reference genome (FN545816.1 for *C. difficile* R20291 and GRCh38 for human) using HISAT2 (version 2.1.0). SAM files were converted to sorted BAM files and sorted using the Samtools (version 1.3.1) 'view' and 'sort' functions. Abundance of each genomic feature was analysed using HTseq-count (version 0.11.2). The data were filtered to only include genes which had > 10 reads in > 10 samples. DESeq2 was used to calculate the differential gene expression profile using a negative binomial distribution model. Differences between groups were considered to be significant, when the adjusted *p*-value was below 0.05. Log_2(fold change) > 1 or < -1 was used as a cutoff for human and bacterial samples. The ggplot2 R package was used to generate graphs, including volcano plots, principal component

analysis (PCA) plots, Venn diagrams, etc. Heatmaps were generated using the online tool ClustVis (available at https://biit.cs.ut.ee/clustvis) which is based on the pheatmap R package. Gene expression data was scaled and centred using the z-score. Row and column clustering distance was calculated with the Euclidean distance and clustered with the average linkage method. The dendrogram demonstrates the hierarchical clustering with the tightest cluster first. Gene ontology (GO) term analysis was conducted using the TopGO R package. The STRING online database (https://string-db.org) was used to investigate KEGG pathways, protein-protein interaction networks and functional enrichment analyses.

2.8 Genomic DNA extraction

DNA was extracted using phenol:chloroform:isoamyl alcohol (25:24:1) (pH 8.0) (Sigma Aldrich, USA). 1- 10 mL of bacterial cultures were centrifuged at 5000 g for 10 min and the cell pellet was re-suspended in 180 μL lysis buffer (PBS with 10 mg/mL lysozyme). Samples were incubated at 37°C for 30 min. 4 μL RNase was added and samples were incubated for 15 min. After incubation, 25 μL proteinase K solution (Qiagen, Germany), 85 μL ddH2O and 110 μL 10% w/v SDS solution were added, and samples were incubated at 65°C for 30 min. An equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) (pH 8) (Sigma Aldrich, UK) was added and samples were mixed by inversion. Samples were centrifuged at 15,000 x g for 3 min. The top layer was transferred into a fresh Eppendorf tube and phenol/chloroform extraction was repeated a further 2 times. DNA was precipitated through the addition of 40 μL 3M sodium acetate and 800 μL ice-cold 100% ethanol. Samples were mixed and incubated at -20°C overnight.

Samples were centrifuged at 15,000 x g for 15 min at 4°C. Supernatants were removed and pellets were washed with 1 mL 70% ethanol followed by centrifugation at 15,000 x g for 3 min at 4°C. Supernatants were removed and pellets were allowed to air dry for 30 min at RT. Finally, the DNA pellets were re-dissolved in 50 μL ddH_2O. DNA concentration of samples was determined by use of Nanodrop or Qubit 2.0 Fluorometer (Life Technologies, UK) following the manufacturer's protocol.

2.9 Quantitative reverse transcription PCR

cDNA was synthesised from RNA using Superscript IV (Thermo Fisher Scientific, USA) and random hexamers (Integrated DNA Technologies, USA) according to the manufacturer's protocols. Primers were designed to produce an amplicon product of 100-300 bases with an annealing temperature of 59-62°C using the Primer3 online tool (https://primer3.ut.ee, version 4.1.0) (Table 2). The qPCR reaction was conducted with the Agilent Mx3005P qPCR System (Agilent, USA) and the Luna® Universal qPCR Master Mix (New England Biolabs, USA). The temperature profile used throughout was the recommended standard for the master mix. To convert from a qPCR output cycle threshold value (Ct) to a predicted messenger RNA (mRNA) expression level, Ct values of amplified target genes from uninfected controls and infected samples were normalised to the Ct of the housekeeping gene, glyceraldehyde 3-phosphate dehydrogenase (GAPDH), a value defined as ΔCt. The difference between the ΔCt of uninfected controls and infected samples were calculated for each timepoint, a value defined as ΔΔCt. Fold changes was analysed using $2^{-\Delta\Delta Ct}$ calculation for each gene of interest.

Table 2. Primers used in RT-qPCR of differentially expressed genes.

Target gene	Forward primer (5'-3')	Reverse primer (5'-3')
GAPDH	GTCTCCTCTGACTTCAACAGCG	ACCACCCTGTTGCTGTAGCCAA
MUC13	ATGGCTGTAACCAGACTGCG	CTTGAGACTGGAAGCAACGC
FGG	TTATTGTCCAACTACCTGTGGC	GACTTCAAAGTAGCAGCGTCTAT
ANXA6	CTGGACATAATCACCTCACG	TTGGCATCACAATAGGCAGG
C. difficile gyrA	GGTTGAAAGAATAGCAGAGTTAGTT	GCATTAGCATCCCTCTTTAATTCTA

2.10 Immunofluorescent staining and confocal microscopy analysis

Epithelial cell layers infected with *C. difficile* as described above were washed thrice with PBS to remove unadhered bacteria and fixed with 4% paraformaldehyde (PFA) (Alfa Aesar, USA) for 15 min at RT. Cells were permeabilised with 1% saponin (Sigma Aldrich, USA) in 0.3% triton X- 100 (Sigma Aldrich, USA) in PBS (Thermo Fisher Scientific, USA) and then blocked with 3% BSA (Sigma Aldrich, USA) in PBS. Rabbit anti-*C. difficile* sera was used as a primary antibody to stain *C. difficile* on infected intestinal epithelial cells (1:500 dilution in 1% BSA in PBS for 1 h at room temperature), followed by goat anti-rabbit IgG AlexaFluor 488 conjugate antibody (1:200 dilution in 1% BSA in PBS for 1 h at room temperature in the dark) (Cell Signalling Technology, USA) as the secondary antibody. Cellular fibronectin was stained using mouse anti-fibronectin primary antibody (1:100 dilution and incubated overnight at 4°C) (Sigma Aldrich, USA) and goat anti-Mouse IgG3 Cross-Adsorbed Secondary Antibody, AlexaFluor 488 (1:200 dilution in 1% BSA in PBS for 1 h at room

temperature in the dark) (Life Technologies, USA). Anti-MUC13 primary antibody (ab65109, Abcam, UK) was diluted 1:200 and incubated for at 4°C overnight. Goat anti-rabbit IgG AlexaFluor488 conjugate antibody (Cell Signalling Technology, USA) as a secondary antibody (1:200 dilution in 1% BSA in PBS for 1 h at room temperature in the dark). AlexaFluor 647 phalloidin (Cell Signalling Technology, USA) was used at a 1:100 dilution in PBS to stain the actin cytoskeleton. ProLong Gold Antifade Reagent with DAPI (Cell Signalling Technology, USA) was used to stain cell nuclei and seal coverslips. Slides were imaged using a confocal spinning-disk microscope (VOX UltraView, PerkinElmer, USA) with a 40X oil objective and two Hamamatsu ORCA-R2 cameras, by Volocity 6.0 (PerkinElmer, USA). Image analysis was performed using the Fiji software package (version 2.0.0).

2.11 Biofilm formation assay and biomass quantification

Overnight *C. difficile* cultures were diluted to OD_{600} 0.02 (measured in spectrophotometer from Biochrom, UK) in prereduced BHIS with 0.1M glucose (BHIS+G) liquid media. 1 mL aliquots of diluted cultures were pipetted into preincubated 24-well tissue culture treated polystyrene plates (Costar, USA), and incubated under anaerobic conditions at 37°C, for 6–72 h. Plates were sealed with parafilm (Bemis, USA) to prevent liquid evaporation. After incubation for the required times, biofilms were washed thrice with PBS and allowed to dry for at least 30 minutes. To stain the biofilms, 1 mL of filter-sterilised 0.2% crystal violet (CV) (Sigma Aldrich, USA) was added to the biofilms and incubated for 30 min at 37°C in anaerobic conditions. CV was removed from the wells and the biofilms were washed thrice with PBS. The dye was extracted using methanol

(VWR International, USA) incubated for 30 min at room temperature (RT) in aerobic conditions. The methanol-extracted dye was diluted 1:10 and the OD_{570} was measured in a 96-well plate (Falcon, USA) with a spectrophotometer (SPECTOstar Nano, BMG Labtech, Germany). To measure bacterial CFU counts from biofilms, the planktonic phase and media were removed, and biofilms were washed thrice with PBS. The adherent phase was disrupted using a pipette tip and resuspended in 1 mL PBS. Serial dilutions were made and plated onto BHIS agar plates and incubated overnight at 37°C in anaerobic conditions. Colonies were counted to determine the CFU present in the biofilm community.

2.12 Aggregation assay

To quantify bacterial aggregation in liquid cultures, overnight cultures grown in BHIS were diluted to an OD_{600} of 0.1 and 4 mL was aliquoted into borosilicate tubes. The OD_{600} of the top 1 mL fraction was quantified before (OD_{pre}) and after (OD_{post}) vortexing every hour for 12 hours. The percentage of cells at the bottom of culture tubes was inferred using the following formula: $[(OD_{post} - OD_{pre})/ OD_{post}] \times 100$.

2.13 *ΔPPEP-1* genetic complementation

The functional PPEP-1 gene (CD2830) was amplified using the "PPEP-1 forward" and "PPEP-1 reverse" primers (Table 3) from WT630 gDNA using Phusion™ high fidelity DNA polymerase (Thermo Fisher Scientific, USA). PPEP-1 was cloned into a pRPF185 vector (Addgene, USA) using the BamHI-HF and SacI-HF restriction enzymes with CutSmart® Buffer (New England Biolabs, USA) and T4 DNA ligase (New England Biolabs, USA). Mix & Go competent *E. coli* DH5-*α* (Zymo Research,

USA) were transformed with the ligation mix and LB agar plates supplemented with 15 μg/mL chloramphenicol were used to select for transformants. Colony PCR with GoTaq® G2 Flexi DNA Polymerase (Promega, USA) and GATC Sanger sequencing (Eurofins, Germany) using the 'pRPF185 sequencing' and 'PPEP-1 reverse' primers were used to confirm that the insert was successfully ligated into the vector. Mix & Go competent *E. coli* CA434 (Zymo Research, USA) were produced following the manufacturer's protocol and used as a conjugative donor strain for the transformation of *ΔPPEP-1* with pRPF185-PPEP-1. 1 mL stationary overnight culture of *E. coli* CA434 pRPF185-PPEP-1 was pelleted by centrifugation at 5000 x g for 10 min and washed with 0.5 mL PBS. The *E. coli* donor pellet was then re-suspended in 200 μL stationary overnight culture of *ΔPPEP-1* and the conjugation mixture were pipetted onto a BHIS non-selective plate in discrete drops and incubated for 24 h under anaerobic conditions to allow conjugal transfer of the pRPF185-PPEP-1 vector from the *E. coli* donor to the *C. difficile* recipient. Following incubation, all cells were harvested and re-streaked on *C. difficile* selective BHIS plates (supplemented with D-cycloserine 250 μg/mL, Cefoxitin 8 μg/mL and thiamphenicol 15 μg/mL). The plates were incubated under anaerobic conditions at 37°C for 24 - 72 h. Single colonies were isolated and re- streaked and cultured for storage in 25% glycerol at -80°C. The plasmid was extracted and checked by GATC Sanger sequencing as described above, to confirm uptake of pRPF185-PPEP-1.

2.14 Collagen binding assay

The OD_{600} of WT and *ΔPPEP-1* overnight cultures grown in BHIS+G were normalised to an OD_{600} of 0.5 before 1 mL of each culture was added to a collagen-

coated 24-well plate. Cultures were incubated for 3 hours in anaerobic conditions before the supernatant was removed and the wells were washed thrice with PBS. In one set of samples, adherent bacteria were released through the addition of 100 μL trypsin and incubation for 5 mins and quantified with CFU counts. In the other sets of samples, 1 mL BHIS+G or BHIS was added and plated were incubated for a further 5 hours. Adherent bacteria were then released and quantified as described above.

2.15 Western blotting

Cell lysates were collected using cOmplete™ mini protease inhibitor cocktail tablets (Sigma Aldrich, USA) dissolved in ice cold RIPA buffer (Sigma Aldrich, USA). Cells were scraped off the Snapwell insert using a sterile pipette tip and samples were centrifuged at 13,000 rpm at 4°C for 10 minutes (Sigma Zentrifugen 1-14K, Germany). The lysed cell supernatant is aliquoted and stored at -80°C until use. For *C. difficile* culture supernatants, cultures were filtered with 0.2 micron syringe filters (GE Healthcare Life Sciences, Whatman™, USA) and treated with Halt™ protease and phosphatase inhibitor cocktail (Thermo Fisher Scientific, USA). Trichloroacetic acid (TCA) precipitation was used to concentrate proteins from culture supernatants. 1 volume of 100% TCA was added to 4 volumes of protein sample and incubated at 4°C for 10 mins. Samples were centrifuged at 14,000 rpm for 5 mins and the pellet was washed with cold acetone twice. Pellet was dried at 95°C before being resuspended in 60 μL dH_2O. Protein concentrations were quantified using Qubit protein assay kit (Thermo Fisher Scientific, USA) or Pierce™ detergent compatible Bradford assay (Thermo Fisher Scientific, USA). 4X SDS buffer (50 mM Tris-HCl pH 6.8, 2% SDS, 10%

glycerol, 1% β-mercaptoethanol, 12.5 mM EDTA, 0.02% bromophenol blue) was added and samples were boiled at 95°C for 10 min. Protein sample concentrations were normalised and run on a 10% gel Mini-PROTEAN® TGX™ Precast Protein Gel (Bio-Rad, USA). Gel was blotted using Trans-Blot® Turbo™ Mini PVDF Transfer Packs (Bio-Rad, USA) and the Trans-blot turbo Bio-Rad transfer machine. Membranes were blocked with 5% BSA (Sigma Aldrich, USA) with shaking at room temperature for 1 hour. Recombinant anti-fibrinogen α-chain (ab92572, Abcam, UK), anti-MUC13 (ab65109, Abcam, UK) and anti-PPEP-1 rabbit sera (gifted from Jeroen Corver, Leiden University) primary antibodies were diluted 1:500 and incubated with the membrane at 4°C shaking overnight. Anti-rabbit HRP-linked secondary antibody (Cell Signalling Technology, USA) was diluted 1:3000 and incubated with shaking for 1 h at room temperature. Membranes were incubated in the detection reagents from the Pierce™ Fast Western Blot Kit for 5 minutes. Membranes were imaged using the G-box system (Syngene, India) and exposed for 6-13 minutes. Membranes were stripped with mild stripping buffer (200 mM glycine, 0.1% SDS, 1% Tween20) and re-probed with monoclonal anti-*β*-actin primary antibody (Sigma Aldrich, USA) and anti-mouse HRP-linked secondary antibody (Cell Signalling Technology, USA) both diluted 1:3000 before being exposed for 1-4 min and imaged in the G-box system.

2.16 ATP release assay

The OD_{600} of *C. difficile* R20291 overnight cultures grown in BHIS were resuspended in pre-reduced DMEM-10, diluted to an OD_{600} of 1 and cultures were incubated for 1 h. 1 mL of each culture was added to intestinal epithelial cells cultured as described above in a 24-well plate. Inoculum tubes kept as

uninfected bacterial controls and 1 mL of pre-reduced DMEM-10 was added to half of the intestinal epithelial cells to be uninfected cell controls. After the required amount of time, supernatants from infected cells or bacterial cultures were filter sterilised through a 0.2 micron syringe filter and extracellular ATP (eATP) was quantified using the CellTiter-Glo® Luminescent Cell Viability Assay (Promega, USA) according to the manufacturer's protocol. This assay can detect ATP from <10 cells (~1 μM). Luminescence was measured by a FLUOstar Omega Multimode Microplate Reader (BMG Labtech, Germany).

2.17 CRISPR/Cas9 knockout of MUC13

MUC13-specific sgRNA sequences were designed using the CHOPCHOP online tool (available online at http://chopchop.cbu.uib.no) (Table 3). Oligonucleotide sgRNA sequences (Integrated DNA technologies, USA) were annealed with T4 polynucleotide kinase (New England Biolabs, USA). Annealed oligonucleotide sequences were cloned into the pSpCas9(BB)-2A-GFP (PX458) vector (Addgene, USA) at the Bbs1 restriction site using FastDigest (BPil/Bbs1) (Thermo Fisher Scientific, USA) and T4 DNA ligase (New England Biolabs, USA). Chemically competent *E. coli* DH5-α (New England Biolabs, USA) were transformed with the ligation reactions and transformants were selected for on LB agar plates supplemented with 100 μg/mL ampicillin. Plasmids were extracted from selected clones with the QIAprep Spin Miniprep Kit (Qiagen, Germany) and GATC Sanger sequencing (Eurofins Genomics, Germany) to confirm the sgRNA inserts had been successfully ligated into the vector using the 'U6 sequencing primer' (Table 3). The QIAGEN Plasmid Plus Midi Kit was used to extract the vectors containing the inserts. For transfection, Caco-2 and HT29-MTX cells were grown

in a 10 cm dish to 80% confluency. Cells were then transfected with 37 μg plasmid DNA in opti-MEM® (Gibco™, Thermo Fisher Scientific, USA), using FuGENE® HD (Promega, USA) in a 1:3 ratio. The mixture was added dropwise to cells in 12 mL fresh media. Cells were grown for 24 hr, when media was replaced with fresh DMEM. After a further 24 hrs, cells were prepared for FACS sorting by trypsinisation and centrifugation. Cells were then filtered to remove clumps and debris, and held in starving media (DMEM, 1% FBS, 1% PenStrep and 2nM L-glutamine) on ice. GFP positive cells were isolated by fluorescence activated cell sorting (FACS) with the BD FACSMelody™ Cell Sorter (BD Biosciences, USA), and single clones were generated by dilution and plating of 50 cells into a 10 cm dish. Clones were grown for 3 weeks before being selected using 4.8 mm cloning discs (Scienceware®, Sigma Aldrich, USA) and trypsin-EDTA to be expanded in a 24-well plate. Once confluent, cells were split into two wells of a 6-well plate. Once confluent, one well for each clone was frozen as described above. The other well underwent a gDNA extraction with the DNeasy blood and tissue kit (QIAGEN, Germany). PCR was used to amplify the sgRNA-targeted regions of MUC13 (primers detailed in Table 3) and knockouts were confirmed with GATC Sanger sequencing (Eurofins, Germany).

Table 3. Oligonucleotide sequences for sgRNA sequences used for CRISPR/Cas9 knockout of MUC13 and *ΔPPEP-1* genetic complementation.

Oligonucleotide sequence	**Sequence (5'-3')**
MUC13 sgRNA 1	CATCTTGGCAAGGATTGCTGGG
MUC13 sgRNA 2	GACTCACCTAATAGTCAGGGCGG
PPEP-1 forward	CTGAGCTCCTGCAGTAAAGGAGAAAATTTTATG AGACCAAGTAAAAAATT
PPEP-1 reverse	TAGGATCCGGCTATTTAGCTAAATTTTGCA
pRPF185 sequencing	CGATGCCCTGGACTTCATG
U6 sequencing	AGGCTGTTAGAGAGATAATTGG
MUC13 PCR 1 forward	GTCATTTCTCTGGATTTCTAATCTCT
MUC13 PCR 1 reverse	ACTGCACCCTGGTGATCT
MUC13 PCR 2 forward	CTAAGGTCAGGATAACGGCC
MUC13 PCR 2 reverse	CGTTGCCTTTACACAGCC

Chapter 3: Dual RNA-seq for differential gene expression analysis during infection

3.1 Introduction

C. difficile is an opportunistic anaerobic bacterium which colonises the gut mucosa of humans and animals. The anaerobic nature of *C. difficile* makes the study of host-pathogen interactions challenging and there is a lack of appropriate infection models which can facilitate the co-incubation of *C. difficile* with human cells for extended periods of time. As a result, the host and pathogen responses to *C. difficile* infection have been particularly understudied. Animal models are commonly used to model *C. difficile* infection for the study of host-pathogen interactions. However, there are several pathophysiological differences in *C. difficile* infection of animal models compared to the disease in humans (Best, Freeman and Wilcox, 2012). Therefore, there are currently large gaps in our knowledge regarding the interactions between *C. difficile* and the intestinal epithelium. Investigation of the transcriptomic responses to *C. difficile* infection is important to gain a better understanding of the molecular processes and interactions which facilitate infections. *C. difficile* is a particularly challenging infection to diagnose and treat (Abt, McKenney and Pamer, 2016; Kociolek, Polage and Riley, 2016). The findings from this study can be potentially be used to identify novel diagnostic markers for the disease, as well as contribute towards the development of new strategies for the treatment of the infection.

One way of determining the molecular processes which are occurring within a cell is with transcriptomic profiling, including the use of methods such as microarrays and RNA-seq. These tools quantify the relative expression levels of genes within a cell, which is collectively defined as the transcriptome. RNA-seq has several advantages over microarrays, such as increased sensitivity and the inclusion of all transcripts, rather than being limited to include only known, coding transcripts (Westermann, Gorski and Vogel, 2012). To investigate host-pathogen interactions during infection, dual RNA-seq is a technique which can be applied to resolve host and pathogen transcriptomic profiles simultaneously during an interaction (Westermann, Gorski and Vogel, 2012). This technique has been applied to a variety of pathogens, including bacteria, fungi and protists in a range of parasitic, commensal, and mutualistic relationships. This technique has provided an improved understanding of the molecular processes which are altered in response to these host-microbe interactions (Rienksma *et al.*, 2015; Westermann *et al.*, 2016; Westermann, Barquist and Vogel, 2017; LaMonte *et al.*, 2019).

In this study, a dual RNA-seq analysis was performed to study host and bacterial transcriptomic responses to *C. difficile* infection. An *in vitro* human gut model recently developed in the laboratory was employed to facilitate *C. difficile* infection of intestinal epithelial cells (Anonye *et al.*, 2019). Total RNA transcripts from host cells and adherent bacteria were extracted from infected intestinal epithelial cells. Samples were prepared for Illumina sequencing by depleting rRNA transcripts and preparation of DNA libraries. After sequencing, the quality of the data was assessed to ensure the obtained reads are of good quality and can

be used for the transcriptomic analyses. A bioinformatics pipeline optimised for dual RNA-seq was developed to analyse and interpret the large-scale sequencing data, as well as perform differential gene expression analyses. Many computational techniques were used to aid in the visualisation and interpretation of the gene expression data. This chapter summarises the dual RNA-seq experimental setup, assessment of the RNA samples prior to sequencing, the sequencing itself and subsequent data visualisation and interpretation.

3.2 Results

3.2.1 Experimental design

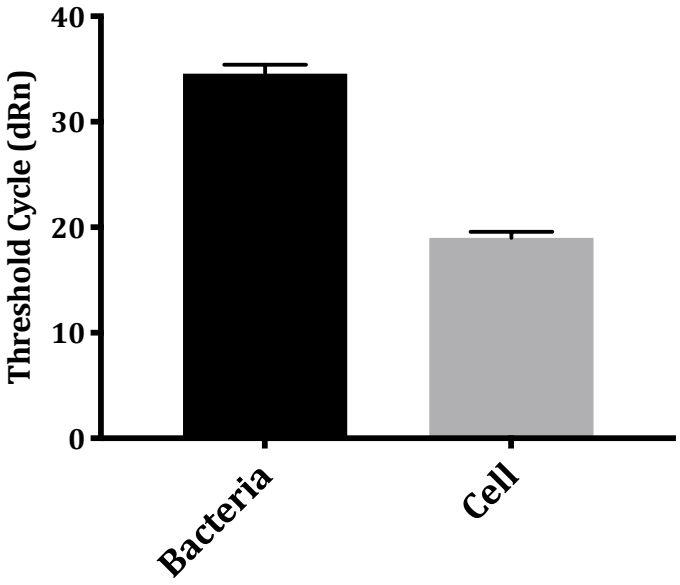
The VDC is an *in vitro* system which can facilitate the co-incubation of host cells with anaerobic bacteria (Schüller and Phillips, 2010; Jafari *et al.*, 2016; Anonye *et al.*, 2019). In this study, Caco-2 colonocytes and HT29-MTX mucus producing goblet cells were seeded in a 9:1 ratio to generate a simple model the intestinal epithelial layer. These cell layers were then infected with *C. difficile* in the VDC system. After incubation in the VDC over several timepoints, cell layers were washed with PBS to remove unadhered bacteria. Total RNA was extracted from the infected cells to be processed for RNA-seq.

3.2.2 RT-qPCR amplification of human and bacterial transcripts in mixed RNA samples

In dual RNA-seq experiments, it is possible that the bacterial RNA can be overwhelmed by host RNA in samples from infected cells due to large differences

in the abundance of host and pathogen transcripts. In numerous dual RNA-seq studies, enrichment of bacterial transcripts was required prior to sequencing to allow them to be detectable and sequenced (Westermann, Barquist and Vogel, 2017). To investigate whether *C. difficile* transcripts could be detected in an RNA sample from infected intestinal epithelial cells, RT-qPCR was used to quantify the relative abundance of transcripts from the host and bacteria. RNA was converted to cDNA with reverse transcription and primers for the human housekeeping gene, GAPDH, and bacterial housekeeping gene, DNA gyrase subunit A, *gyrA,* were used to amplify and quantify cDNA fragments with qPCR. This step was important to ensure the quantity of bacterial RNA present in the samples was sufficient to be converted to cDNA and amplified to be prepared for sequencing (Figure 6).

A.



B.

(average bacterial Ct – average host Ct) x 2000

(34.5 - 19) x 2000

= 3100 fold more cellular RNA than bacterial RNA

Figure 6. RT-qPCR to quantify host and bacterial housekeeping transcripts.

A. The threshold cycle of RNA from *C. difficile*-infected intestinal epithelial cells. Host cDNA was amplified using primers for GAPDH, and bacterial cDNA was amplified using primers for *gyrA*. B. Calculation of the fold difference between the quantity of cellular RNA and bacterial RNA. Error bars are representative standard deviation.

The *C. difficile* transcripts for the *gyrA* housekeeping gene were successfully amplified from a mixed RNA sample, suggesting that it is likely to be possible to generate bacterial cDNA fragments during library preparation and which can be sequenced (Figure 6A). The fold difference between the quantity of host and bacterial cellular RNA was calculated with the equation ((average bacterial Ct –

average host Ct) x 2000). It was calculated that the host cell contained approximately 3100 fold more cellular RNA than the bacteria (Figure 6B).

3.2.3 Pilot study

Prior to performing a large-scale dual RNA-seq experiment, it was important to ensure that enough bacterial mRNA transcripts could be sequenced and detected to obtain bacterial transcriptomic data from infected samples. Therefore, a small-scale pilot study was conducted to avoid the waste of resources and to allow optimisation, such as the enrichment of bacterial transcripts, to be performed if required before performing the experiment on a larger scale. Intestinal epithelial cells infected with *C. difficile* R20291 were incubated in the VDC for 3 h and 24 h.

To prepare the samples for sequencing, contaminating gDNA was removed from the RNA samples with a DNase treatment. Illumina Ribo-Zero Gold epidemiology kit was used for ribosomal RNA depletion of the infected samples as it depletes both human and bacterial rRNA simultaneously. Reducing the quantity of host and bacterial rRNA transcripts is an important step to increase the number of sequencing adaptors available for mRNA transcripts which contain valuable transcriptional information. This step is particularly necessary during dual RNA-seq workflows as there are a relatively small quantity of bacterial transcripts present in the RNA samples from infected cells. Therefore, without rRNA depletion, the transcriptional profile of the bacteria is likely to be overwhelmed by rRNA transcripts. During library preparation, RNA is converted into cDNA by reverse transcription and ligated to adaptors required for sequencing. These adaptors contain specific sequences designed to interact with the surface of the

flow-cell in the Illumina sequencing platform. Indexes are also ligated to the DNA fragments during library preparation to facilitate demultiplexing of the RNA samples *in silico,* after the sequencing process is completed. The RNA integrity was analysed with the Bioanalyzer by chip-based electrophoresis after each step to ensure the RNA was of adequate quality and was not degraded before continuing to the next process in the workflow (see Methods section 2.5, Figure 7).

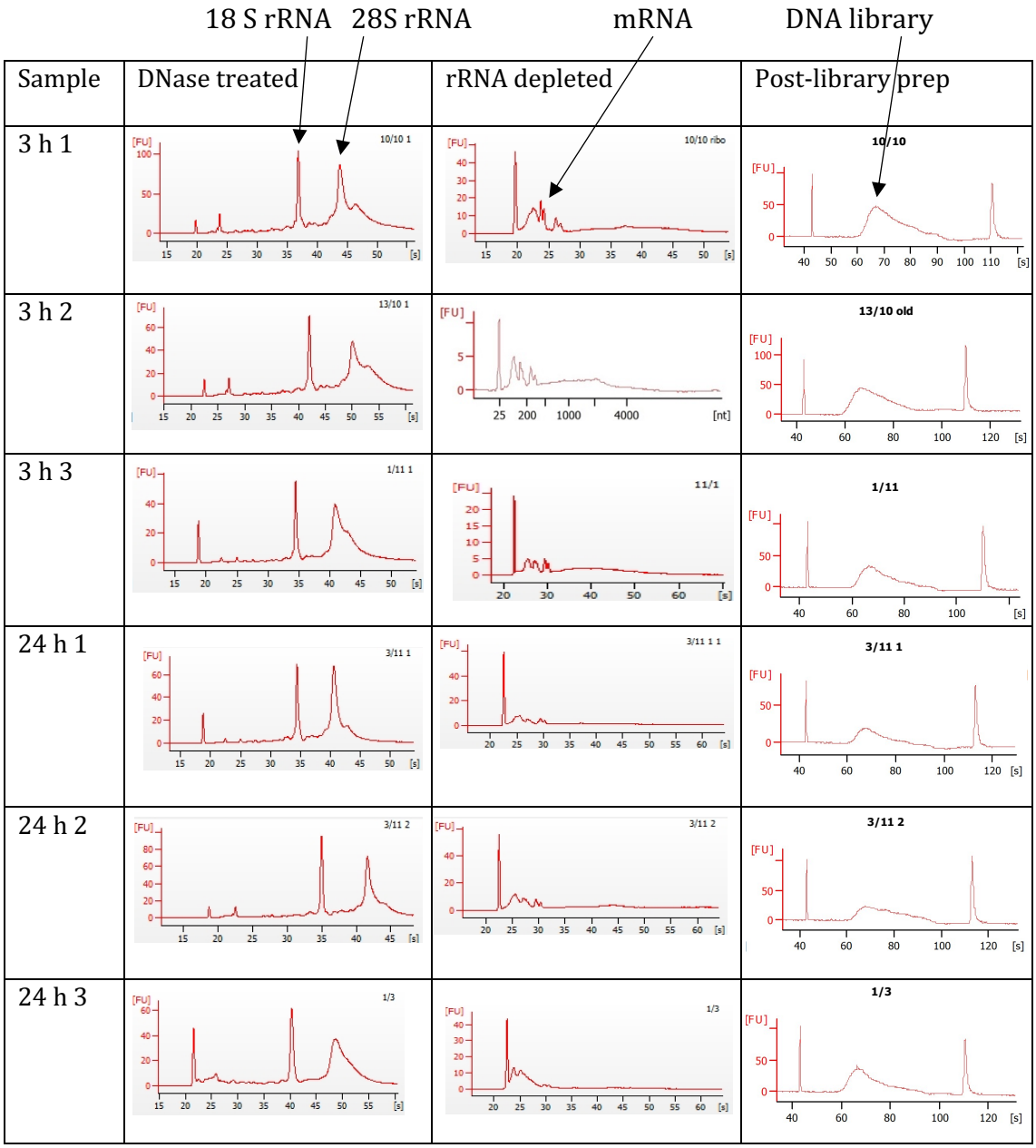


Figure 7. Bioanalyzer traces from the pilot study.

RNA bioanalyzer traces of total RNA and rRNA depleted RNA from infected samples at 3 h and 24 h, as well as DNA bioanalyzer traces of library prepped samples from the pilot study. Key peaks are annotated for the first row of the figure.

The bioanalyzer traces show that there were two distinct peaks after the DNase treatment which are representative of the 18S and 28S human rRNA transcripts. This suggests that the total RNA was not degraded during the extraction

processes and the integrity of the RNA was adequate at this stage. In some of the samples there was a minor second peak next to the 28S rRNA peak which may be from the 23S bacterial rRNA transcripts. The post-rRNA depletion bioanalyzer traces showed that there was a degradation of the two peaks associated with rRNA transcripts, indicating that the rRNA depletion kit was successful at removing rRNA transcripts from the samples. The remaining RNA transcripts in the sample were shorter fragments which are likely to be the highly informational mRNA transcripts. Finally, after the library preparation, the bioanalyzer traces show that diverse DNA libraries were produced with an average fragment size of approximately 300 bp (Figure 7). The DNA libraries were of an acceptable quality to be used to perform sequencing. Samples were sequenced on the Illumina NextSeq™ 500 and FASTQ files were generated.

FastQC is a tool which can be used to assess the quality of the sequencing reads, to ensure the data is reliable. FastQC considers a number of statistics including total number of reads, per sequence quality score, per sequence GC content, sequence duplication levels, overrepresented sequences and kmer content. This software was developed to analyse genomic DNA sequencing data. Many of the samples failed on statistics such as per base sequence content, per base GC content, per sequence GC content, but this is common for transcriptomic data (Appendix table 4).

The percentages of reads aligned to the human and *C. difficile* R20291 genomes were calculated. The percentage alignments of the infected samples to the human reference genome ranged from 61.83-95.99% (Figure 8A). From this, the

coverage of the human and *C. difficile* genomes could be calculated using the formula included in Figure 8B, which approximates the number of times each gene in the genome has been sequenced. Two of the 3 hour samples (samples 1 and 2) had a low percentage of genes aligned to the human reference genome which resulted in low genome coverage. All other samples had a genome coverage of least 1.59. For *C. difficile* genome coverage, 24 h 1 sample had a very low percentage alignment of 0.05%, and therefore also had a low genome coverage of 0.669 (Figure 8B). Using DESeq2 to perform differential gene expression analysis, 475 human genes were identified as being differentially expressed between 3 h and 24 h (Appendix table 10). However, no bacterial differentially expressed genes could be identified, potentially because the gene expression profiles may be similar at both timepoints. Since bacterial transcripts were successfully amplified and sequenced, the enrichment of bacterial transcripts was likely to not be required in the subsequent dual RNA-seq analyses.

A.

Sample	Number of reads	% alignment	
		Human	*C. difficile*
3 h 1	65785600	61.83	1.15
3 h 2	17802602	66.17	1.09
3 h 3	93426878	95.99	0.24
24 h 1	78450420	94.06	0.05
24 h 2	83397248	76.15	0.60
24 h 3	93053258	94.02	0.10

B.

$$C = LN / G$$

C = coverage
G = haploid genome length
L = read length
N = number of reads

C.

Sample	Coverage (LN / G)	
	Human	*C. difficile*
3 h 1	1.02	12.90
3 h 2	0.29	3.31
3 h 3	2.24	3.82
24 h 1	1.85	0.67
24 h 2	1.59	8.53
24 h 3	2.19	1.59

Figure 8. Quantification of genome alignment and coverage.

A. Percentage of genes aligned to the human and *C. difficile* reference genomes using HISAT2 in the pilot study data. B. The genome coverage calculation equation. C. Human and *C. difficile* genome coverages of 3 h and 24 h infected samples.

3.2.4 Large-scale dual RNA-seq study

Since the small-scale study successfully amplified and sequenced bacterial transcripts in RNA samples from infected cells, a second dual RNA-seq study was carried out on a larger scale. There were a few areas in the experimental design and setup in which optimisation was conducted. For the pilot study, three bacterial inoculums were added to the VDC system and RNA was extracted from one selected timepoint to produce three biological replicates. In the larger scale study, one set of intestinal epithelial cells were incubated in the VDC system, and RNA was extracted at each timepoint, so each biological replicate was produced on a different day. This was done to account for variability between different experiments and improve the robustness of the data obtained. The intestinal epithelial cell layers were infected with *C. difficile* in the VDC system over 4 timepoints after infection; 3, 6, 12 and 24 h. To generate host uninfected control samples, cell layers were incubated in the VDC system for the required timepoints without the addition of *C. difficile*. As a control for the bacterial samples, *C. difficile* culture was grown to log-phase in pre-reduced DMEM-10, the same media used in the infected samples. Total RNA was extracted from all infected samples and uninfected controls (Figure 9).

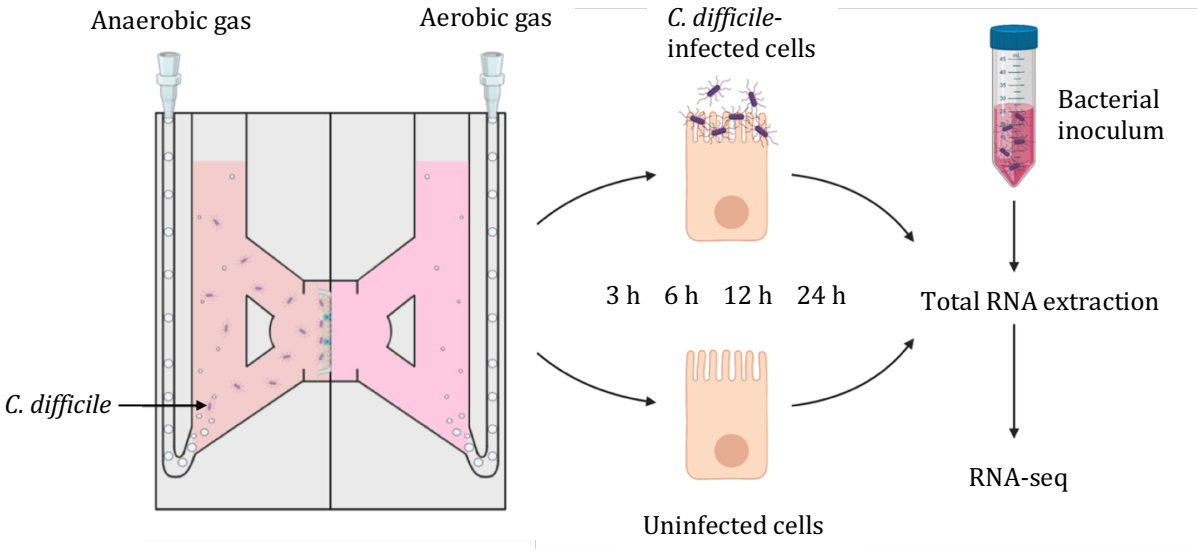


Figure 9. A schematic of the dual RNA-seq experimental setup.

The VDC system was used as an *in vitro* human gut model to facilitate *C. difficile* infection of intestinal epithelial cells. Total RNA was extracted from intestinal epithelial cells incubated in the VDC system for 3 h, 6 h, 12 h, or 24 h, with or without *C. difficile*. A bacterial inoculum grown to logarithmic phase in prereduced DMEM-10 was used as an uninfected bacterial control.

3.2.4.1 RNase H rRNA depletion

Following total RNA extraction and DNase treatment, rRNA transcripts were depleted from the samples. As the RiboZero epidemiology rRNA depletion kit, which was used in the pilot study was discontinued by Illumina, and no suitable alternatives were available on the market which were capable of depleting both human and bacterial rRNA simultaneously at that point in time. For the this set of samples, the RNase H method for rRNA depletion was used as described by Adiconis *et al.* (2013). This method involved designing 50 bp single stranded DNA (ssDNA) probes which are complementary to ribosomal RNA transcripts

followed by an annealing reaction (Appendix table 1). The RNase H endonuclease is then used to cleave DNA-RNA hybrids, removing the rRNA from the sample. DNase is added to the reaction to digest leftover DNA oligos from the RNA sample, generating an rRNA depleted RNA sample (illustrated in Figure 10).

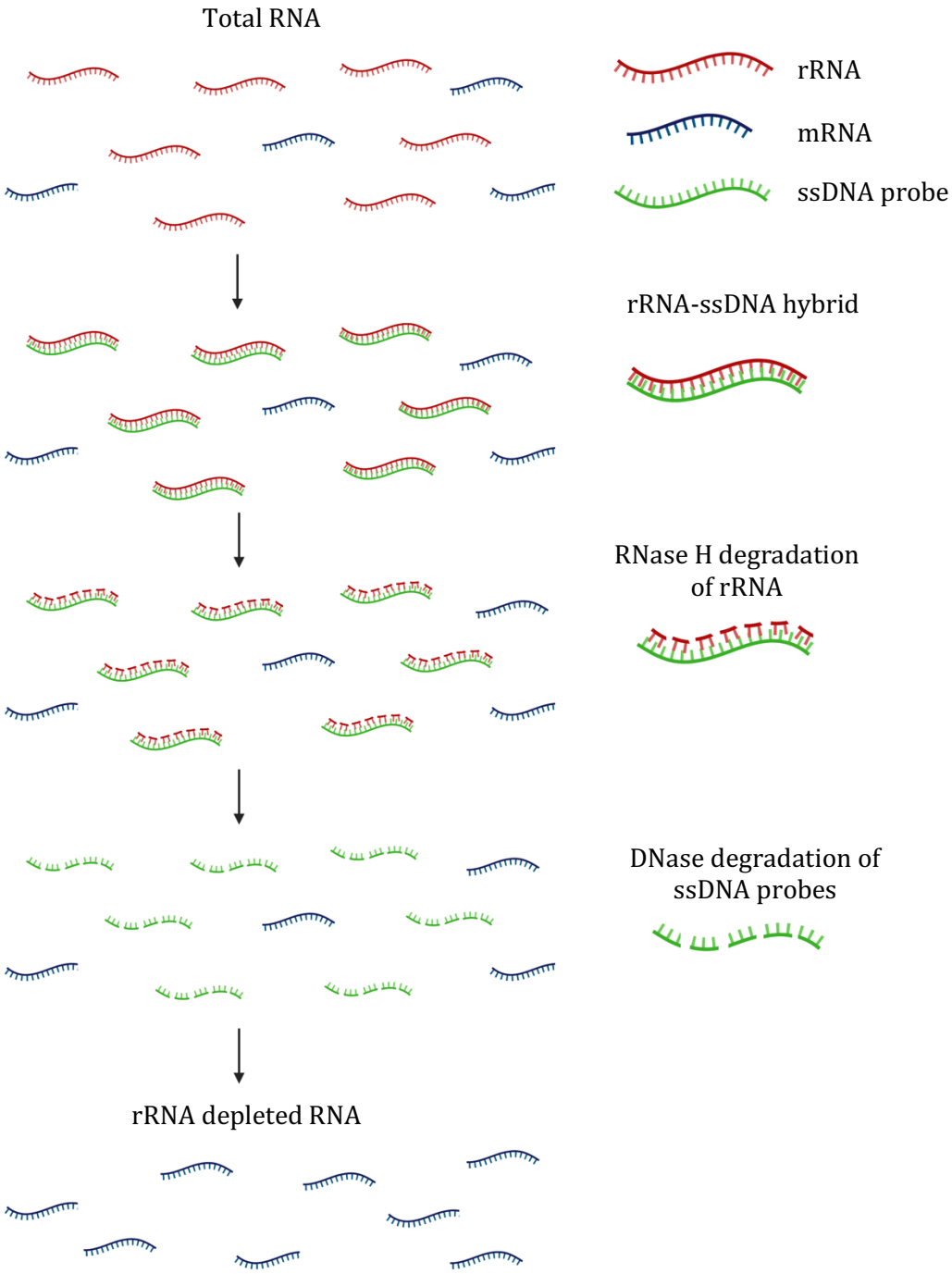
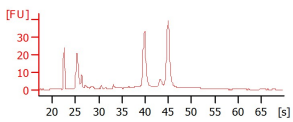
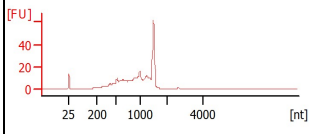
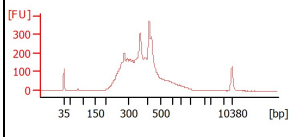
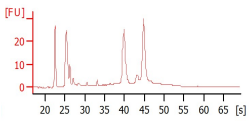
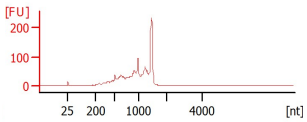
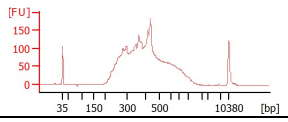
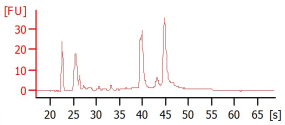
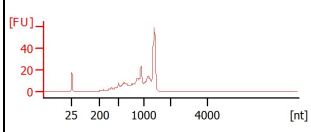
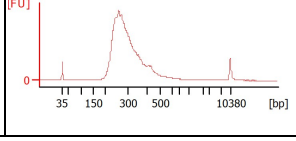
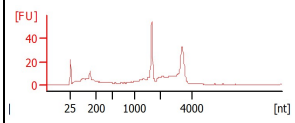
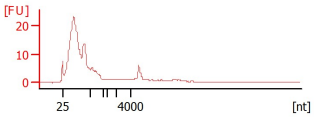
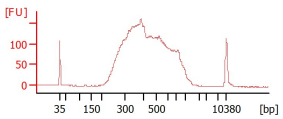
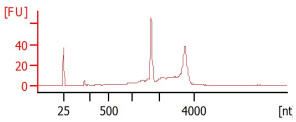
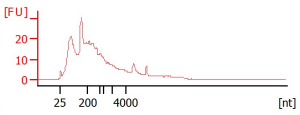
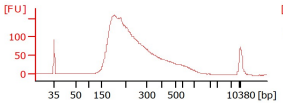
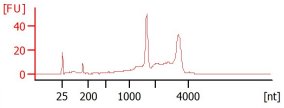
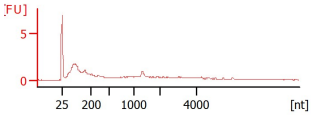
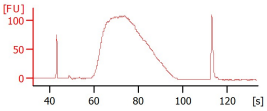
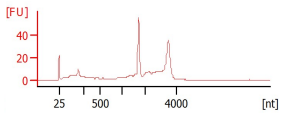
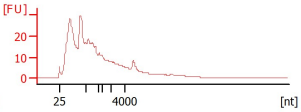
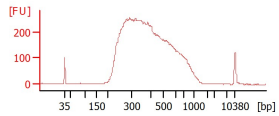
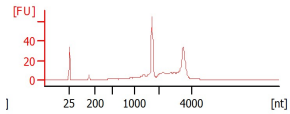
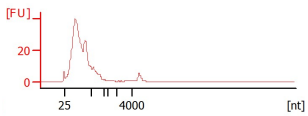
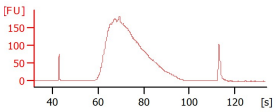
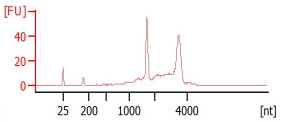
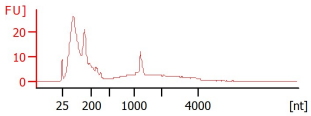
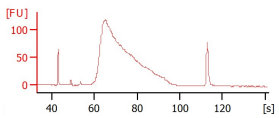
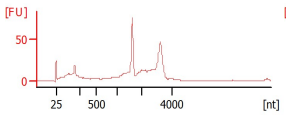
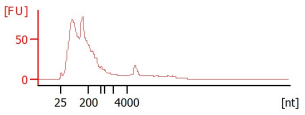
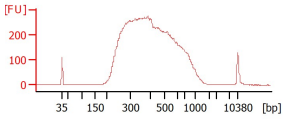
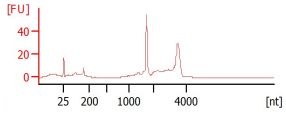
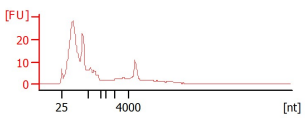
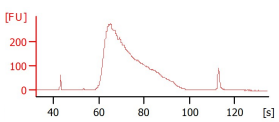


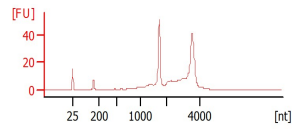
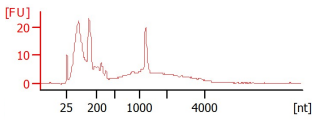
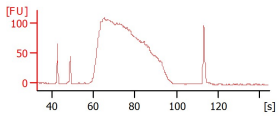
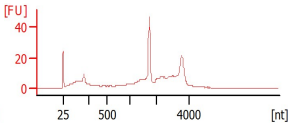
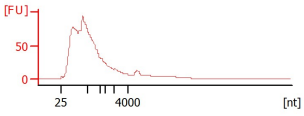
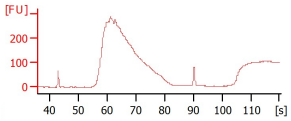
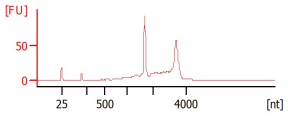
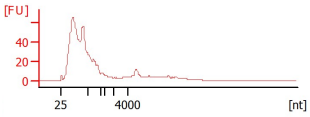
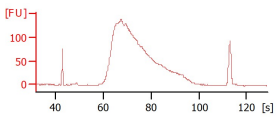
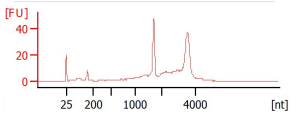
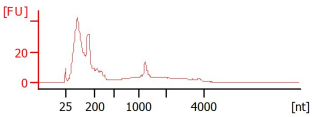
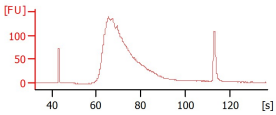
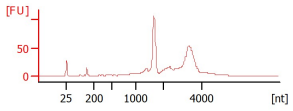
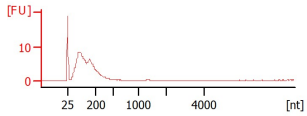
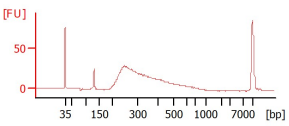
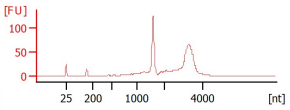
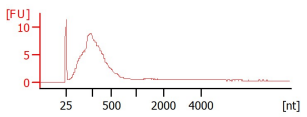
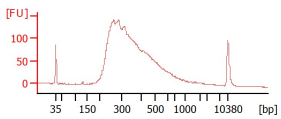
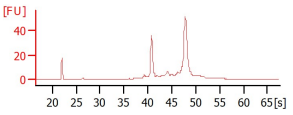
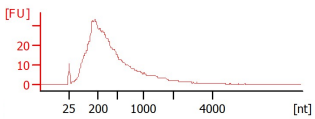
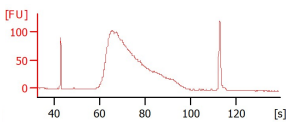
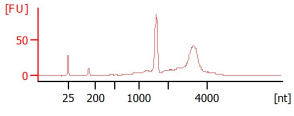
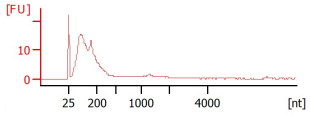
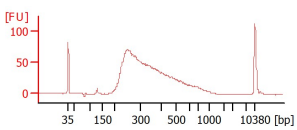
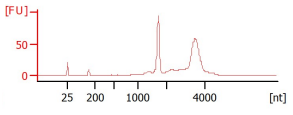
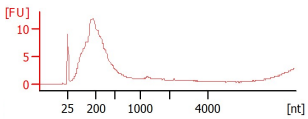
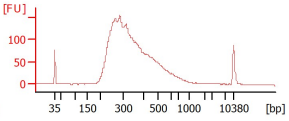
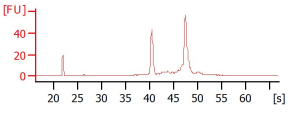
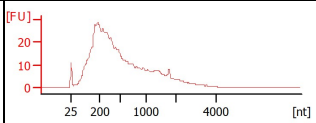
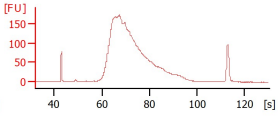
Figure 10. Schematic of the RNase H method for rRNA depletion.

RNA samples are treated with ssDNA probes which hybridise to rRNA transcripts to form rRNA-ssDNA hybrids. RNase H treatment degrades the rRNA transcripts in the rRNA-ssDNA hybrids before the leftover ssDNA probes are degraded with DNase treatment, generating a rRNA depleted sample.

The bioanalyzer was used to assess the quality of the extracted RNA and the effectiveness of RNase H rRNA depletion. In the bacterial control samples, the RNase H method for rRNA depletion was successful at removing the majority of the 16S rRNA peak, determined by analysing rRNA depleted bioanalyzer traces (Figure 11). However, analysis of the bioanalyzer traces shows that rRNA from the 23S subunit may not have been fully removed, as a peak can still be observed at ~1500 nucleotides. The library preparation procedure was completed, and some abnormal peaks can be observed, likely from a larger quantity of DNA fragments which were generated and amplified from the leftover rRNA from the 23S subunit. For the human control and infected samples, the bioanalyzer traces showed that the rRNA subunit peaks were successfully removed. There was an increase in the quantity of fragments approximately 200 bp in size. These peaks had relatively low fluorescent units (FU), and were assumed to be mRNA transcripts. The DNA library preparations for the human uninfected control and the infected samples exhibited the optimal shape for a standard library, suggesting the libraries were of good quality and diversity (Figure 11).

Sample	DNase treated	rRNA depleted	Post-library prep
Bacterial control 1			
Bacterial control 2			
Bacterial control 3			

Human control 3 h 1			
Human control 3 h 2			
Human control 3 h 3			
Human control 6 h 1			
Human control 6 h 2			
Human control 6 h 3			
Human control 12 h 1			
Human control 12 h 2			

Human control 12 h 3			
Human control 24 h 1			
Human control 24 h 2			
Human control 24 h 3			
Infected 3 h 1			
Infected 3 h 2			
Infected 3 h 3			
Infected 6 h 1			
Infected 6 h 2			
Infected 6 h 3			

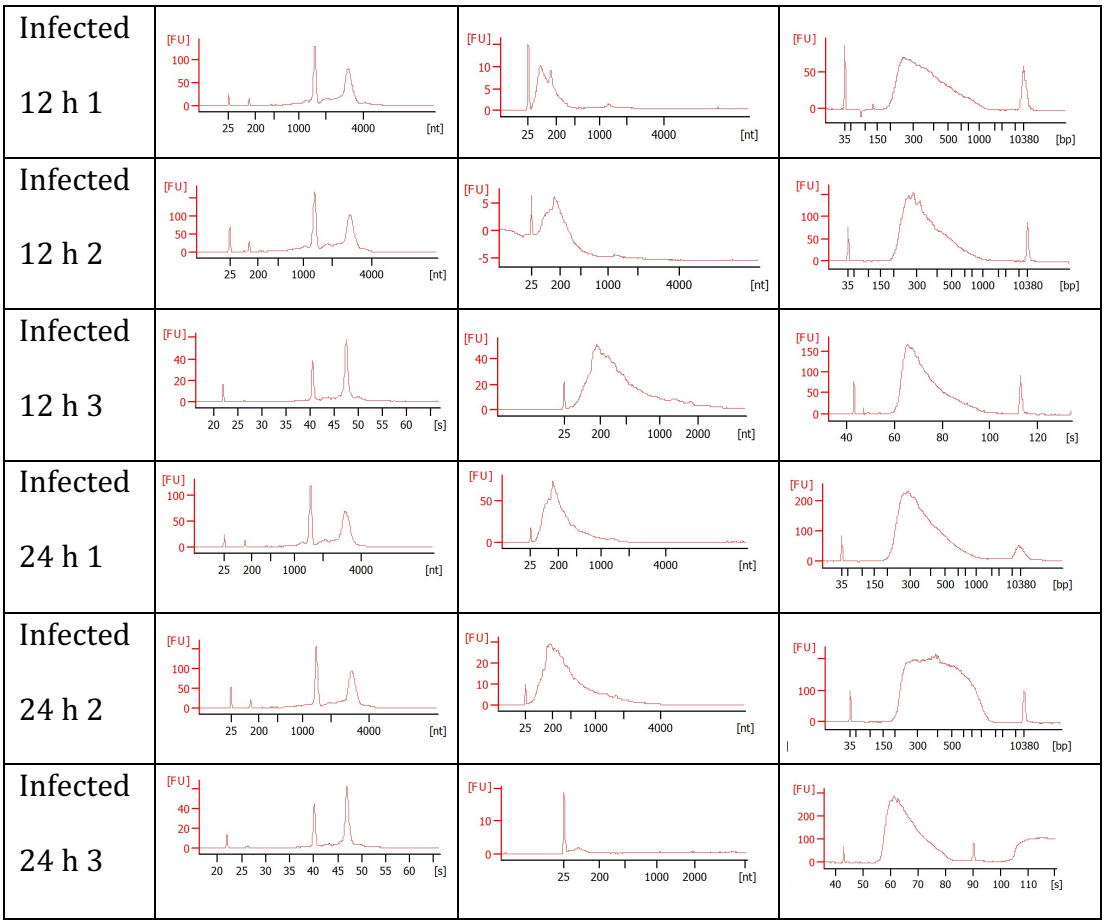


Figure 11. Bioanalyzer traces from the large-scale transcriptomic study.

RNA bioanalyzer traces of total RNA and rRNA depleted RNA from infected samples and uninfected controls, as well as DNA bioanalyzer traces of library prepped samples from the large-scale transcriptomic study. Peaks at 20 nucleotides or smaller are assumed to be markers which act as internal standards used to align the ladder data with sample data.

3.2.4.2 Next-generation sequencing and bioinformatics

Bacterial control samples were sequenced using the Illumina MiSeq™, since prokaryotic genomes are relatively small and the output from the MiSeq™ is sufficient to obtain adequate genome coverages. Human uninfected controls and infected samples were sequenced using the Illumina NextSeq™ 500 as a significantly larger sequencing output is required to obtain a sufficient coverage

of the human genome. During Illumina sequencing, cDNA transcripts bind to adaptors on the flow cell, generating clusters which produce one single read, or one paired-end read. If a low number of clusters are generated, underclustering may occur which reduces the data output. If too many clusters form, overclustering may occur which negatively affects the sequencing data in a number of ways, including a reduction in sequence quality and less accurate demultiplexing after sequencing. The cluster density of the bacterial control sequencing run was within the recommended optimal range. The human control sample run was slightly overclustered which may have reduced the sequencing quality due to overloaded signal intensities. The infected sample runs were slightly underclustered, so the number of reads in the data output is smaller than if the cluster density was within the optimal range (Table 4). The Phred quality score (Q-score) is a measure of base calling accuracy and is a commonly used method for assessing the accuracy of Next-generation sequencing. Overall, all sequencing runs had a at least 92% of reads with a Q-score greater than or equal to 30 (Figure 12) and acceptable numbers of reads were obtained for RNA-seq analysis (Table 4).

Table 4. Cluster densities and numbers of reads from dual RNA-seq sequencing runs.

Sequencing run	Platform	Cluster density (K/mm²)	Optimal cluster density (K/mm²)	Total number of reads
Bacterial controls	MiSeq™	1385	1200-1400	91,785,959
Human controls	NextSeq™ 500	228	170-220	371,452,755
3 h and 24 h infected	NextSeq™ 500	158	170-220	344,735,881
6 h and 12 h infected	NextSeq™ 500	155	170-220	337,612,949

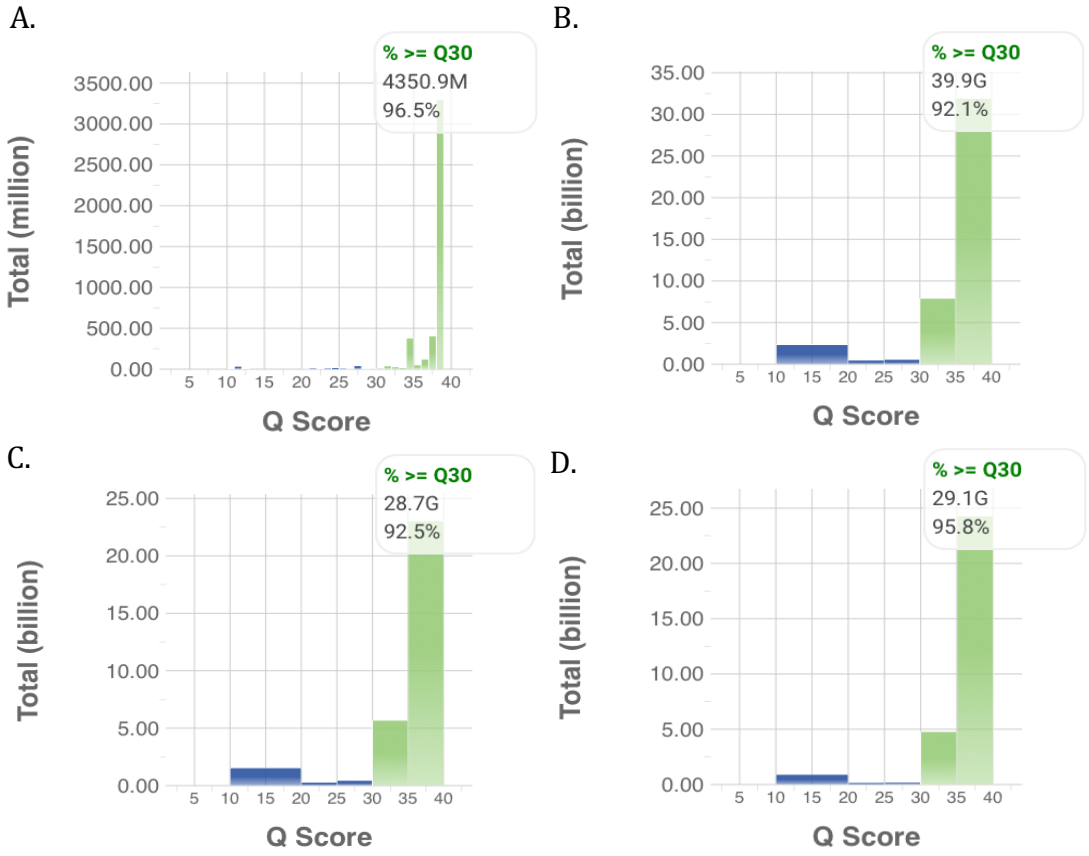


Figure 12. Q-score distribution of Illumina sequencing runs.

A. bacterial control samples, B. Human control samples, C. 3 h and 24 h infected samples, D. 6 h and 12 h infected samples.

Following FASTQ generation from the raw sequencing files, trimming was used to remove sequencing adaptors and reads with a Q-score less than 30. The FastQC software was used to assess the quality of the sequencing reads. As discussed earlier, FastQC was developed to be used with genomic DNA sequencing data, so it is common for transcriptomic data to fail on metrics such as per base sequence content, per base GC content and sequence duplication levels. The majority of samples also contained overrepresented sequences, which were identified as rRNA transcripts using a BLASTn search. This suggested that the rRNA depletion technique was not effective at removing all rRNA transcripts from the samples (Appendix table 3).

Another process which was optimised after the small-scale pilot study was that all samples, including uninfected controls, were aligned to a concatenated reference genome containing both the human (GRCh38) and *C. difficile* R20291 genomes. Aligning to a concatenated genome allowed all samples to be aligned to the same reference genome file and allows the removal of any ambiguous reads which align to both the human and *C. difficile* genomes. The removal of ambiguous reads is particularly important in dual RNA-seq experiments as it would not be possible to determine which species the reads aligned to regions in both genomes originated from. Without the removal of these ambiguous reads, there would be a potential source of bias in the data. All samples had over 95% of the reads aligned to the concatenated genome. The ~5% of reads which did not align to the concatenated genome may have originated from the PhiX sequencing control, leftover sequencing adaptors or reads which were

sequenced with a low Q-score potentially altering the sequence of the read. Samples were also aligned to the *C. difficile* R20291 genome separately to determine how many bacterial reads were present in the sample. There was a range of alignments to the *C. difficile* genome, from 0.06% to 2.06% (Table 5).

Table 5. Total number of reads, percentage of reads aligned to a concatenated dual reference genome and percentage of reads aligned to *C. difficile* genome of all sequenced samples.

Sample type	Number of reads	Aligned to concatenated genome reference (%)	Aligned to *C. difficile* genome (%)
Bacterial control 1	8936289	99.11	99.21
Bacterial control 2	12094137	99.02	99.40
Bacterial control 3	70755533	99.33	99.55
Human control 3 h 1	19115353	94.38	0.00
Human control 3 h 2	25470128	90.68	0.00
Human control 3 h 3	27025069	95.67	0.00
Human control 6 h 1	21468597	91.70	0.00
Human control 6 h 2	24593073	92.11	0.00
Human control 6 h 3	57975499	96.20	0.00
Human control 12 h 1	22390842	91.97	0.00
Human control 12 h 2	23882910	94.54	0.00
Human control 12 h 3	58915160	96.13	0.00
Human control 24 h 1	23835563	90.63	0.00
Human control 24 h 2	21564602	93.29	0.00
Human control 24 h 3	45215959	94.42	0.00
3 h infected 1	49226860	95.45	1.32
3 h infected 2	58246564	97.60	0.19

3 h infected 3	55316199	96.99	1.68
6 h infected 1	57700752	97.21	0.14
6 h infected 2	58987014	98.14	0.06
6 h infected 3	57515435	97.75	0.32
12 h infected 1	56507425	97.25	0.33
12 h infected 2	53660909	98.25	0.11
12 h infected 3	53241414	97.93	0.42
24 h infected 1	62622140	97.76	2.06
24 h infected 2	64673740	97.50	0.06
24 h infected 3	54650378	96.98	0.44

From the number of reads in each sample and the percentage of reads aligned to each genome, the genome coverage can be calculated to estimate the number of times each gene in the genome has been sequenced (Figure 8B, Table 6). Some of the human control samples had a relatively low coverage of less than 1 despite a high percentage of alignment, due to a low number of reads present in the samples. All infected samples had coverage values of greater than 1, suggesting that acceptable coverage of the human genome was obtained. There were 2 samples with a coverage of the *C. difficile* genome less than 1: 6 h sample 2 and 24 h sample 2. This suggests that some *C. difficile* genes may not have been sequenced which may introduce biases during the transcriptomic data analysis. All other samples had sufficient coverage of the *C. difficile* genome (Table 6).

Table 6. Human and *C. difficile* genome coverages of infected samples and uninfected controls.

Sample type	Human genome coverage	*C. difficile* genome coverage
Bacterial control 1	0	151.120
Bacterial control 2	0	204.913
Bacterial control 3	0	1200.633
Human control 3 h 1	0.451	0
Human control 3 h 2	0.577	0
Human control 3 h 3	0.646	0
Human control 6 h 1	0.492	0
Human control 6 h 2	0.566	0
Human control 6 h 3	1.394	0
Human control 12 h 1	0.514	0
Human control 12 h 2	0.564	0
Human control 12 h 3	1.416	0
Human control 24 h 1	0.540	0
Human control 24 h 2	0.503	0
Human control 24 h 3	1.067	0
3 h infected 1	1.175	11.076
3 h infected 2	1.421	1.886
3 h infected 3	1.341	15.841
6 h infected 1	1.402	1.377
6 h infected 2	1.447	0.603
6 h infected 3	1.406	3.137
12 h infected 1	1.374	3.179

12 h infected 2	1.318	1.006
12 h infected 3	1.303	3.812
24 h infected 1	1.530	21.989
24 h infected 2	1.576	0.661
24 h infected 3	1.325	4.099

The Samtools software package was used to convert the .sam aligned sequencing files to binary .bam files and sort them. HTSeq-count was used to count the number of reads per gene feature. The HTSeq-count output provides details of the reads which were not aligned to the annotated reference file. There was a staggeringly large proportion of the total reads which were not aligned to the reference file due to the alignment not being unique, suggesting that it aligned to more than one location on the reference genome. This suggested that there were rRNA transcripts present in the samples as there are repeating regions of ribosomal genes in both the human and *C. difficile* genomes. The FastQC report previously showed that there were overrepresented sequences present in the sample which aligned to rRNA transcripts. To investigate the quantity of rRNA reads present in the sample, the bbduk.sh software was used to align the FASTQ file for each sample to a rRNA reference file composed of concatenated host and *C. difficile* rRNA sequences. The software provides a percentage of reads which aligned to the rRNA reference file which enables assessment of the quantity of rRNA contamination. The program then splits the file into two new files, one containing the reads which aligned to the rRNA reference file and the other containing reads that did not (Table 7).

Table 7. The percentage of reads aligned to a rRNA reference file containing human and bacterial rRNA sequences.

Sample type	% alignment to rRNA
Bacterial control 1	21.10
Bacterial control 2	2.88
Bacterial control 3	89.88
Human control 3 h 1	31.23
Human control 3 h 2	68.51
Human control 3 h 3	59.24
Human control 6 h 1	67.25
Human control 6 h 2	70.74
Human control 6 h 3	55.15
Human control 12 h 1	37.42
Human control 12 h 2	44.77
Human control 12 h 3	14.24
Human control 24 h 1	16.38
Human control 24 h 2	13.57
Human control 24 h 3	21.81
3 h infected 1	47.76
3 h infected 2	84.63
3 h infected 3	57.99
6 h infected 1	36.92
6 h infected 2	83.82
6 h infected 3	71.28
12 h infected 1	31.71
12 h infected 2	80.67

12 h infected 3	59.80
24 h infected 1	83.52
24 h infected 2	82.69
24 h infected 3	54.62

There were a large range of percentages of alignments to the rRNA reference sequence. In general, a large proportion of the reads mapping to rRNA sequences were removed from the analysis during the counting of reads per gene feature because rRNA reads map to more than one location of the genome, leaving a smaller number of unambiguous reads available for differential gene expression analysis (Table 7). After removal of ribosomal reads, the reads per gene feature were quantified again and the gene lists were separated *in silico* into human and bacterial gene count matrices. In R Studio, the gene count lists were restricted to include only genes which had at least 10 reads present in at least 10 samples (out of 24 for the human analysis and 15 for the bacterial analysis) to reduce biases and remove genes which were expressed at very low levels in multiple samples. DESeq2 was used to normalise the data and perform differential gene expression analysis.

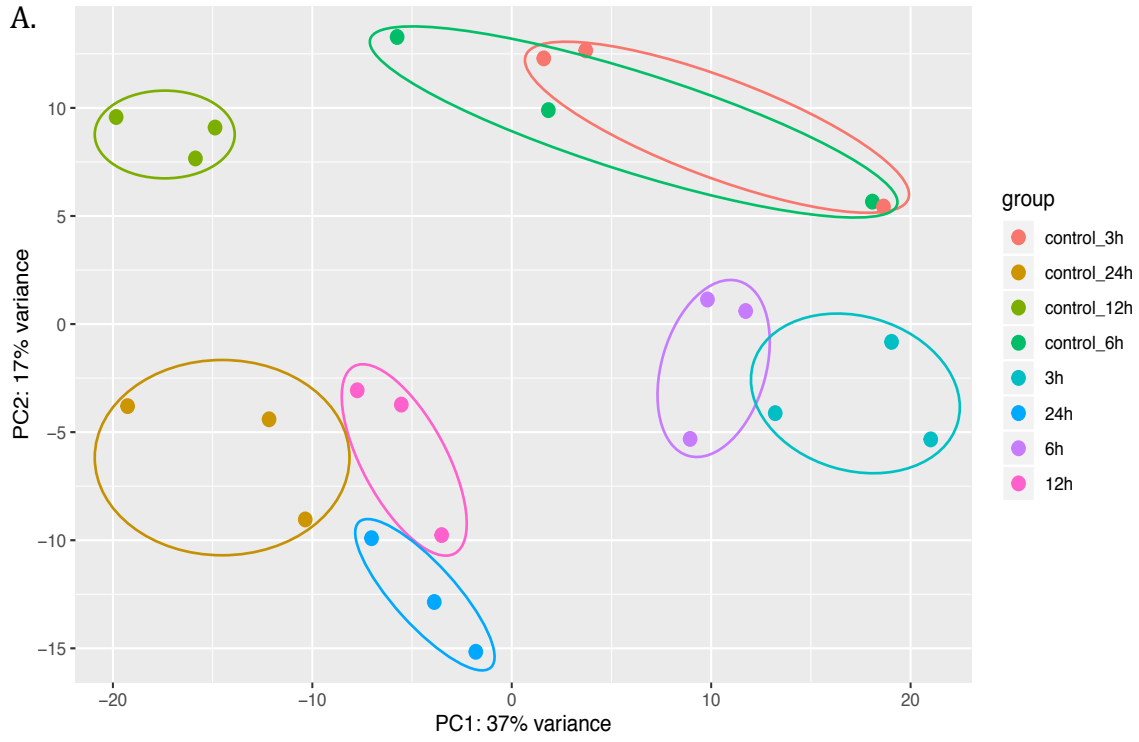
3.2.4.3 Visualising variations between samples and replicates

To investigate the amount of variation between samples and their biological replicates a, principal component analysis (PCA) was conducted. A PCA plot is a technique for visualising the variation between samples and groups. There are multiple principal components which measure the variation in the data, but only the first two (PC1 and PC2) can be represented on a PCA plot. Percentages

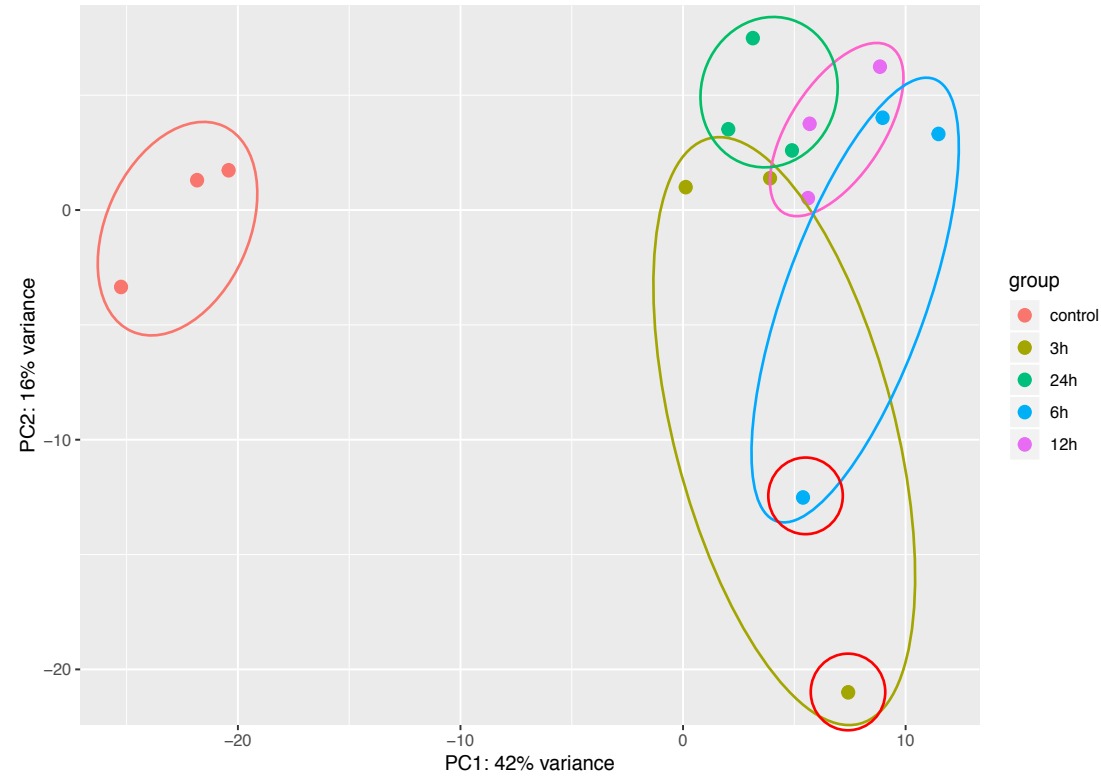
indicate how much of the total variation in the samples can be represented on that principal component. The lower the percentages of PC1 and PC2 the more variation there is present in the samples which can't be represented on the PCA plot.

The human principal component analysis displayed sample-sample distances between each sample and group at all timepoints. The majority of samples were moderately clustered together in their relevant groups and a clear transcriptional shift was observed between uninfected controls and infected samples at all timepoints. There were also changes in the gene expression profiles over the timecourse in both the uninfected controls and infected samples. This suggests that incubation in the VDC system may alter the gene expression profiles of the intestinal epithelial cells, even in the absence of *C. difficile* infection. Interestingly, the uninfected 24 h controls were clustered more closely to the late-stage infected groups than the other uninfected control groups. This suggests that incubation in the VDC system may elicit a stress response in the intestinal epithelial cells which is exacerbated over time. The percentages on the principal component axes were acceptable at 37% for PC1 and 17% for PC2 considering the high numbers of principal components in the analysis. This suggests that 54% of the variation present in the data set can be represented on the plot, but there is a remaining 46% which cannot be represented on the plot (Figure 13A).

A.


10
5
0
−5
−10
−15
PC2: 17% variance
−20
−10
0
10
20
PC1: 37% variance
group
control_3h
control_24h
control_12h
control_6h
3h
24h
6h
12h

B.


0
−10
−20
PC2: 16% variance
−20
−10
0
10
PC1: 42% variance
group
control
3h
24h
6h
12h

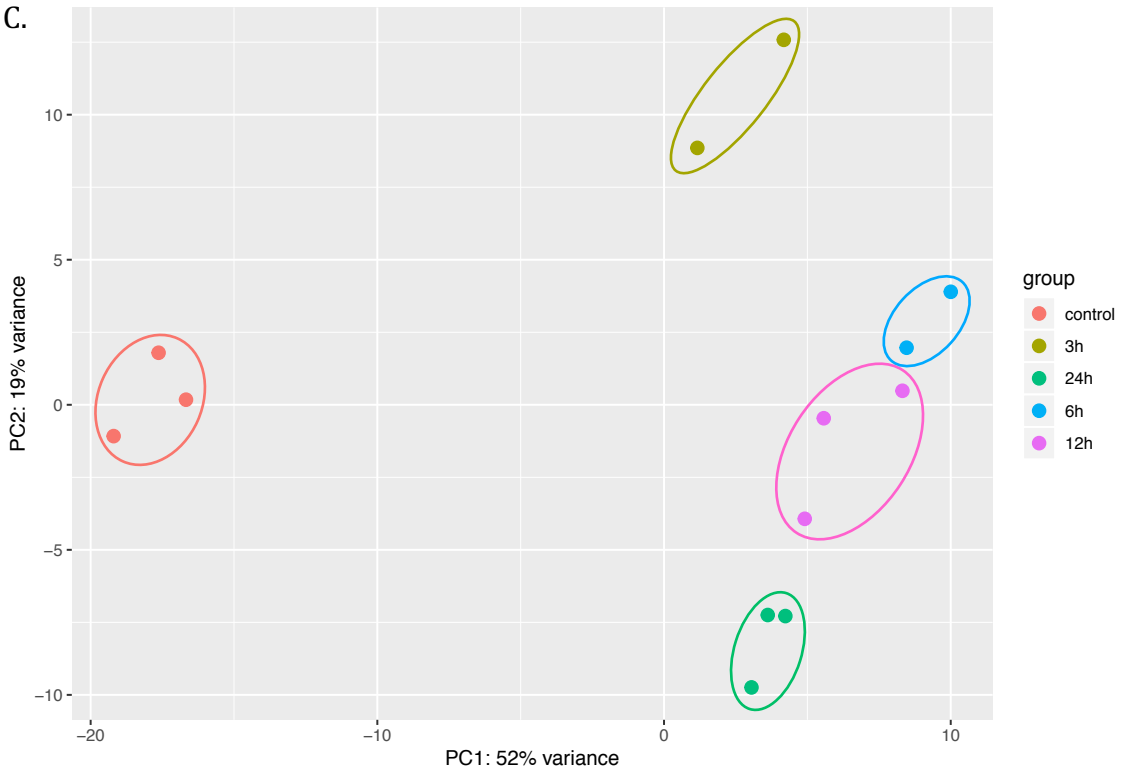


Figure 13. Principal component analyses of transcriptomic data.

A. PCA plot of host RNA-seq data to show sample-sample distances of infected samples and uninfected controls at 3 h, 6 h, 12 h and 24 h. B. PCA plot of bacterial RNA-seq data to show sample-sample distances of infected samples at 3 h, 6 h, 12 h and 24 h compared to an uninfected control culture. The red circles indicate samples which were removed from the analysis due to high variation from the other replicates. C. PCA plot of bacterial RNA-seq data after the removal of two highly distant samples.

In the bacterial PCA plot there were biological replicates in the 3 h and 6 h infected samples which had large genetic variations compared to the other replicates. For this reason, these outliers were completely removed from the analysis (Figure 13B). Following their removal, all replicates within samples were clustered together and a clear transcriptional shift from the bacterial

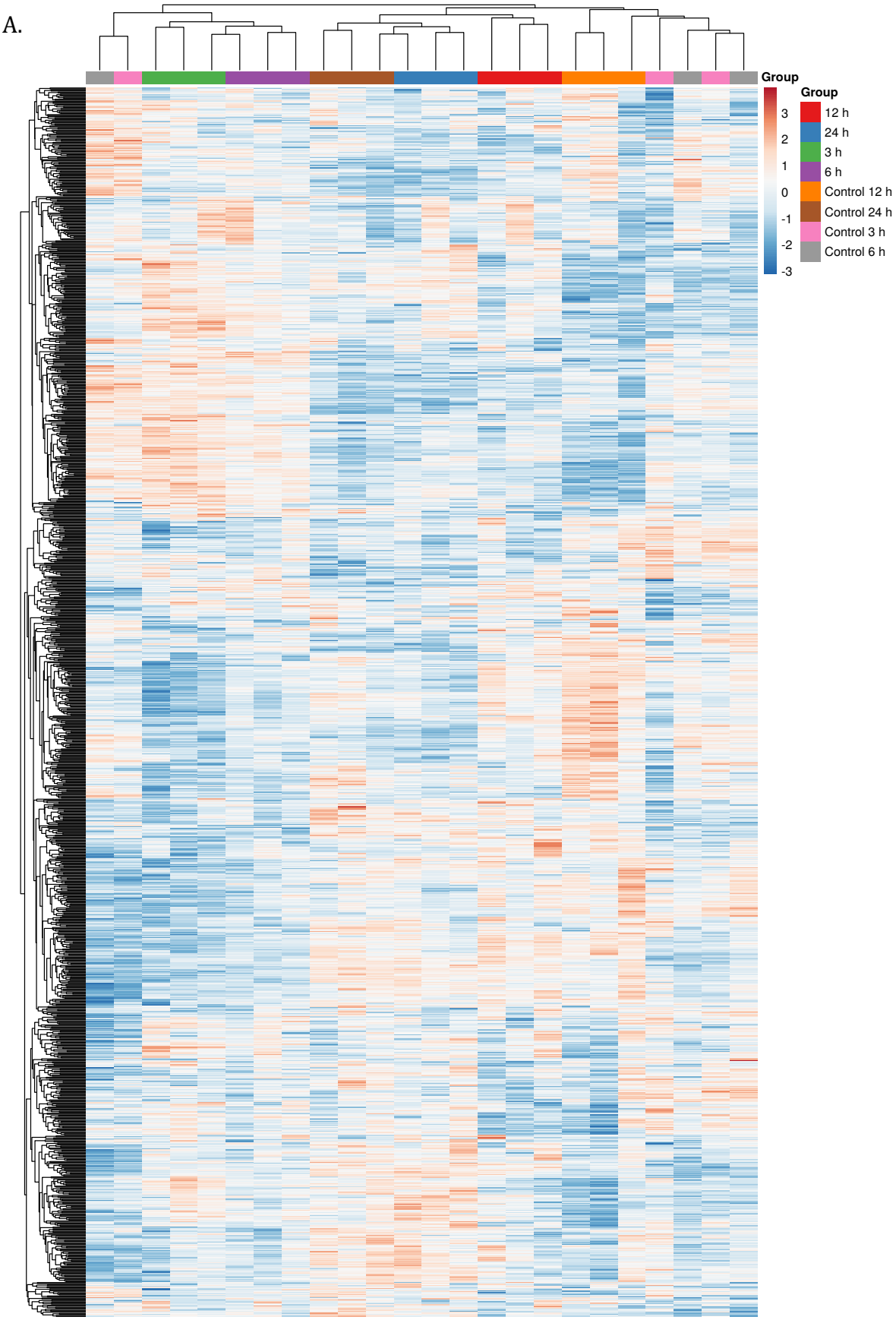
culture control was observed for all timepoints. Alterations in *C. difficile* gene expression over the course of the infection could be determined from the PCA plot. After removal of the outliers, the principal components were relatively high at 52% for PC1 and 19% for PC2. This indicates that 72% of the variation in the data can be represented in the graph which is the majority of the variation (Figure 13C).

Heatmaps are another technique for visualising the genetic distances between samples and groups. A regularised logarithmic transformation was applied to the DESeq dataset to obtain normalised expression levels for each gene. Heatmaps were generated using the online tool ClustVis which is based on the pheatmap R package. Heatmaps illustrate the gene expression profiles using the z-score, a normalised measure of transcript abundance which is calculated by subtracting the overall average gene abundance from the raw expression for each gene and dividing that result by the standard deviation of all of the measured counts across all samples. A high z-score (blue) indicates a high level of transcript abundance and a low z-score (red) indicates a low level of transcript abundance. Row and column clustering distances were calculated with the Euclidean distance and clustered with the average linkage method. The dendrograms illustrate the hierarchical clustering which groups similar samples together with the tightest cluster first.

Analysis of the heatmap from the human transcriptomic analysis showed that there were a moderate number of variations in gene expression profiles between

the replicates at each timepoint. Hierarchical clustering between the samples showed that the majority of the replicates were clustered together, indicating that they hold a high degree of similarity. The hierarchical clustering is consistent with what was shown in the PCA plot, where one replicate from the 3 h uninfected controls and one from the 6 h uninfected controls were genetically distant from the other replicates in these groups. Another finding which is consistent with the results from the PCA plot, is that the uninfected 24 h controls were clustered more closely to the infected sample groups than the other uninfected control groups (Figure 14A). Analysis of the heatmap of the bacterial transcriptomic analysis, demonstrated that a clear transcriptional shift can be observed between the uninfected controls and infected samples, through analysis of the contrasting colour scale and genetic distances shown on the hierarchical clustering dendrogram (Figure 14B).

A.


Group
3
2
1
0
-1
-2
-3
Group
12 h
24 h
3 h
6 h
Control 12 h
Control 24 h
Control 3 h
Control 6 h

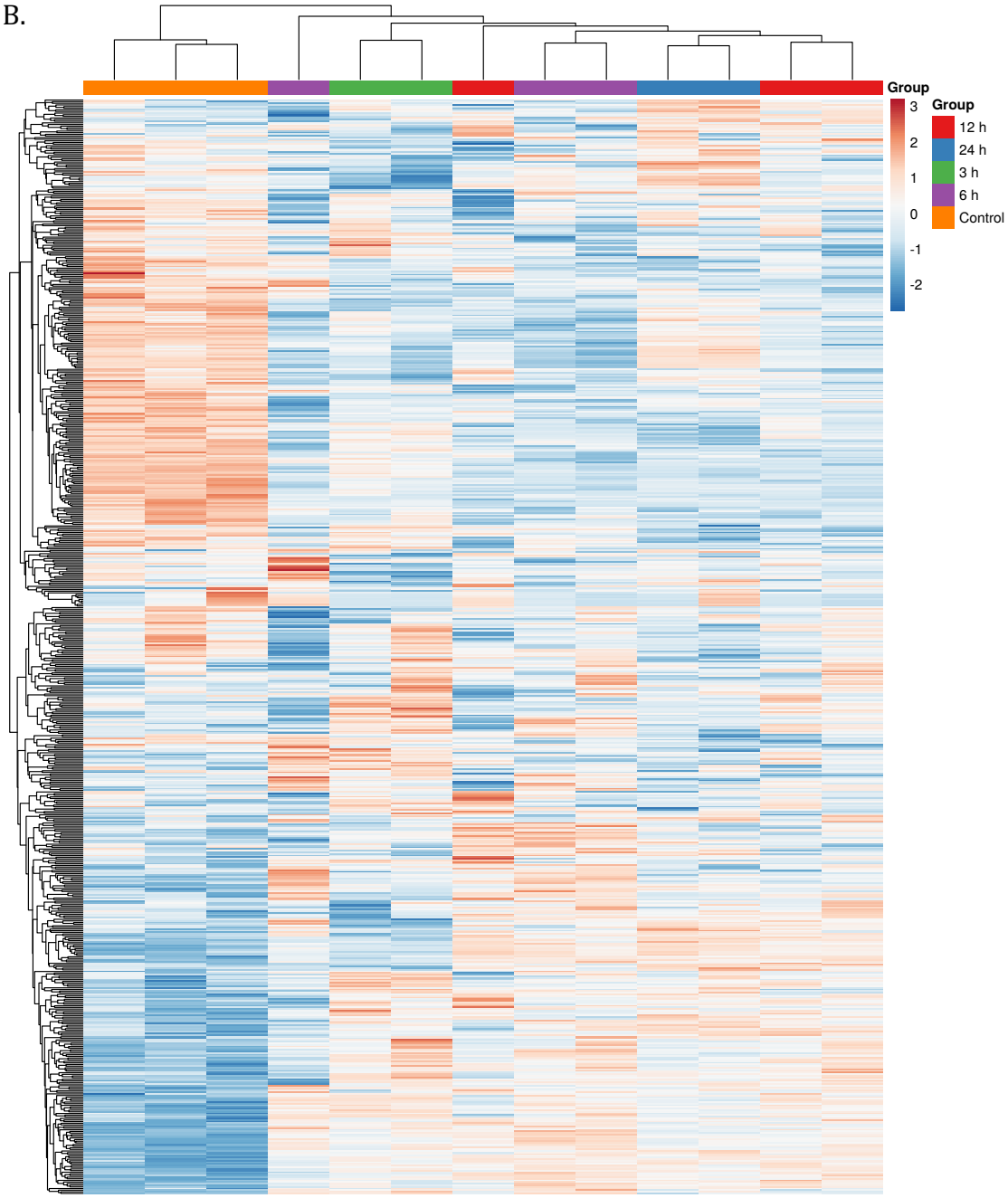


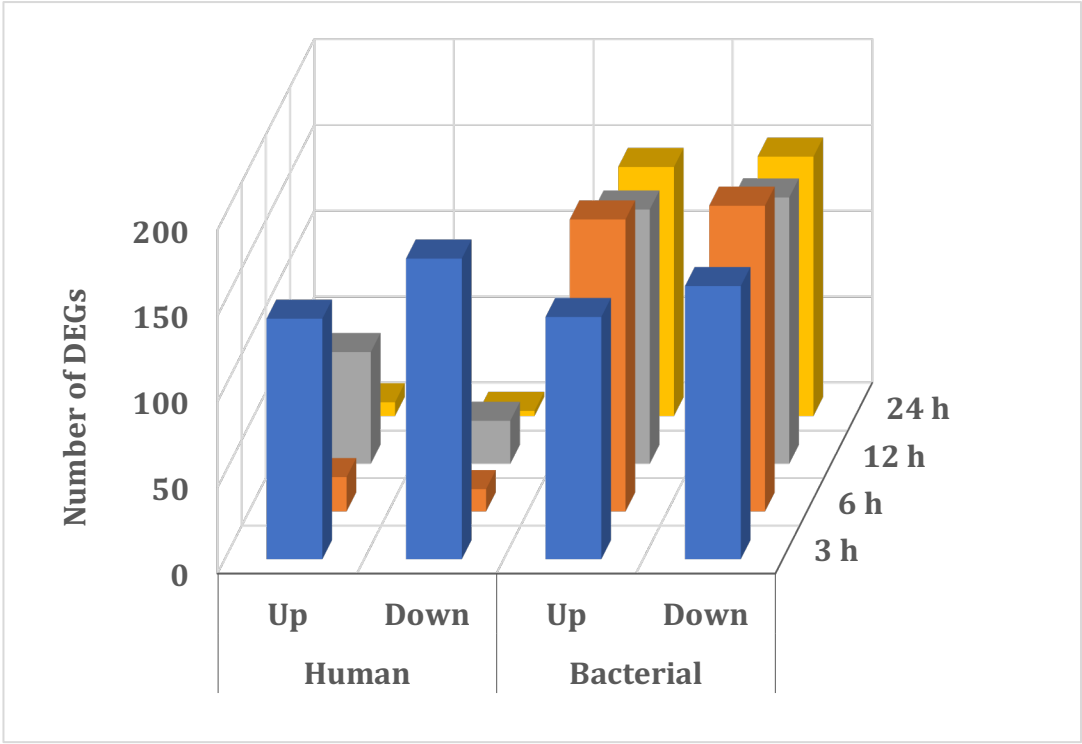
Figure 14. Heatmaps of host and bacterial transcriptomic data.

A. Heatmap of host transcriptomic data to demonstrate the variations in gene expression between the infected samples and uninfected controls at each timepoint. B. Heatmap of bacterial transcriptomic data to demonstrate the variations in gene expression between the infected samples and uninfected controls at each timepoint. Differential gene expression values calculated using the z-score which is represented with a colour scale, where red indicates a positive z-score, and blue indicates a negative z-score.

3.2.4.4 Analysis of the distribution of differentially expressed genes

To determine the number of genes which were differentially expressed between the uninfected controls and infected samples at each timepoint, a cut-off was applied to include only genes with an adjusted *p*-value less than 0.05, and a $\log_2$(fold change) greater than 1 or less than -1. The number of up- or downregulated human genes varied between the different timepoints. At 3 h, 140 genes were upregulated, and 175 genes were downregulated. At the later timepoints there was a considerable reduction in the number of differentially expressed genes, with just 20 upregulated genes and 13 downregulated genes identified at 6 h, 65 upregulated genes and 25 downregulated genes identified at 12 h and just 8 upregulated genes and 3 downregulated genes identified at 24 h (Figure 15A, full lists available in Appendix tables 11-14). There were a higher total number of significantly differentially expressed bacterial genes identified than human genes, despite the size of the bacterial genome being very small compared to the human genome and there only being a small proportion of bacterial RNA in the infected samples compared to the large amount of human RNA. In the bacterial transcriptomic analysis, there were 141 upregulated and 159 downregulated genes identified at 3 h, 170 upregulated and 178 downregulated genes at 6 h, 148 upregulated and 155 downregulated genes at 12 h and 145 upregulated and 151 downregulated genes at 24 h (Figure 15A, full lists available in Appendix tables 4-7).

A.



B.

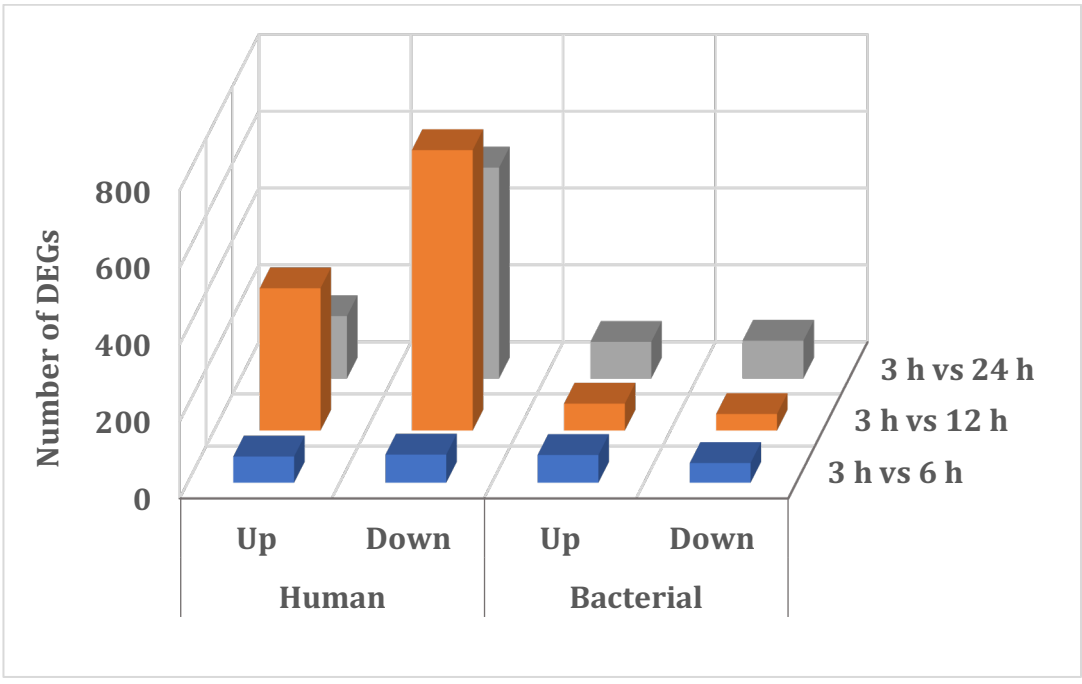


Figure 15. The numbers of up- or downregulated genes.

A. Differentially expressed genes (DEGs) at 3 h, 6 h, 12 h, and 24 h post infection, as compared to the relevant uninfected controls. B. Differentially expressed genes between

the 3 h infected samples and the other infected samples at 6 h, 12 h and 24 h post infection.

Since there was a relatively low genome coverage value for some of the human uninfected controls, the infected sample timepoints were compared to the gene expression profile at 3 h to investigate the alterations in gene expression over the course of the infection. High numbers of human genes were identified as differentially expressed between 3 h and 12 h post infection, as well as 3 h and 24 h post infection (Figure 15B). The genes which were differentially expressed when comparing the uninfected 3 h controls and the relevant uninfected timepoint were removed from the lists of differentially expressed genes between the 3 h infected samples with the other infected timepoints. The numbers of differentially expressed genes which were modulated between the 3 h infected samples and other relevant infected timepoints, but not between the relevant uninfected timepoints was greatest between 3 h and 24 h (Figure 16, full list available in Appendix table 15).

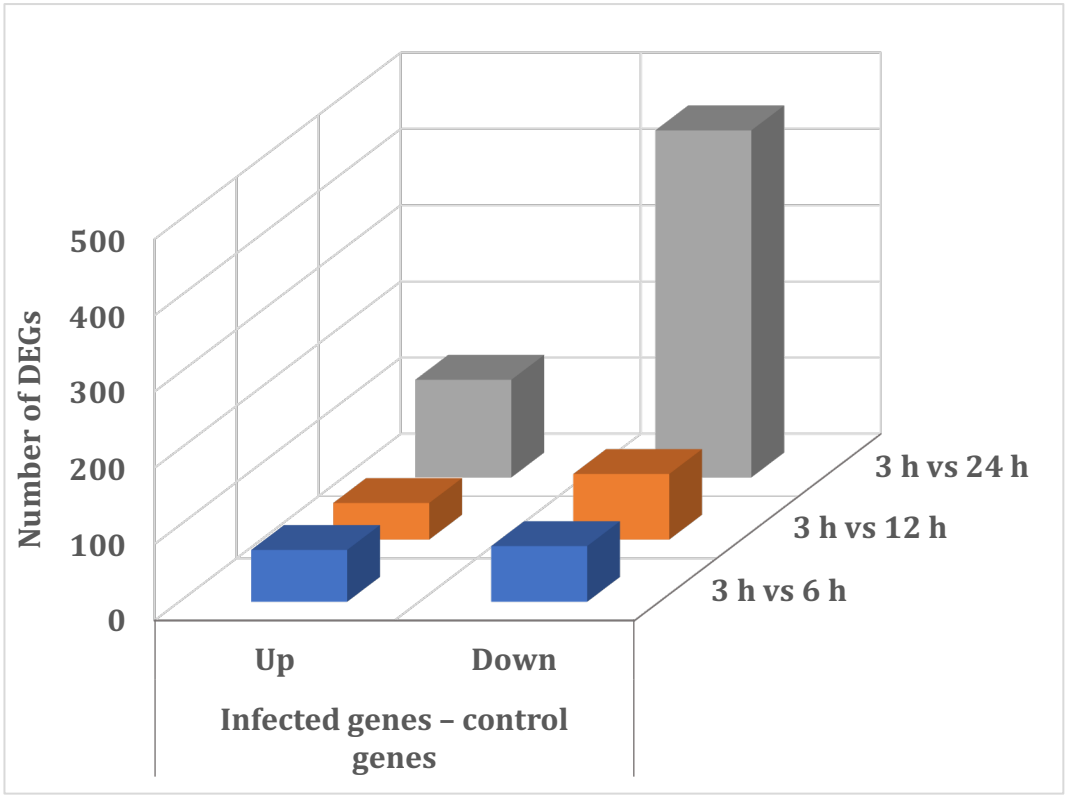


Figure 16. The numbers of differentially expressed genes between the human infected samples minus the differentially expressed genes between the uninfected control samples.

3.2.4.5 Visualising the distribution of differentially expressed genes

Volcano plots were used to visualise the distribution of human and bacterial differentially expressed genes by plotting $\log_2$(fold change) against the $\log_{10}$(adjusted *p*-value). Some of the most significantly differentially expressed genes were pointed out with arrows. In the human transcriptomic analysis at 3 h post infection, there were a larger number of downregulated genes with larger $\log_2$(fold change) values and more significant adjusted *p*-values compared to the upregulated genes. At the later timepoints, there were a reduced number of significantly up- or downregulated human genes (Figure 17). The distributions

of bacterial differentially expressed genes were similar at all timepoints, with a high number of significantly differentially expressed genes (Figure 18).

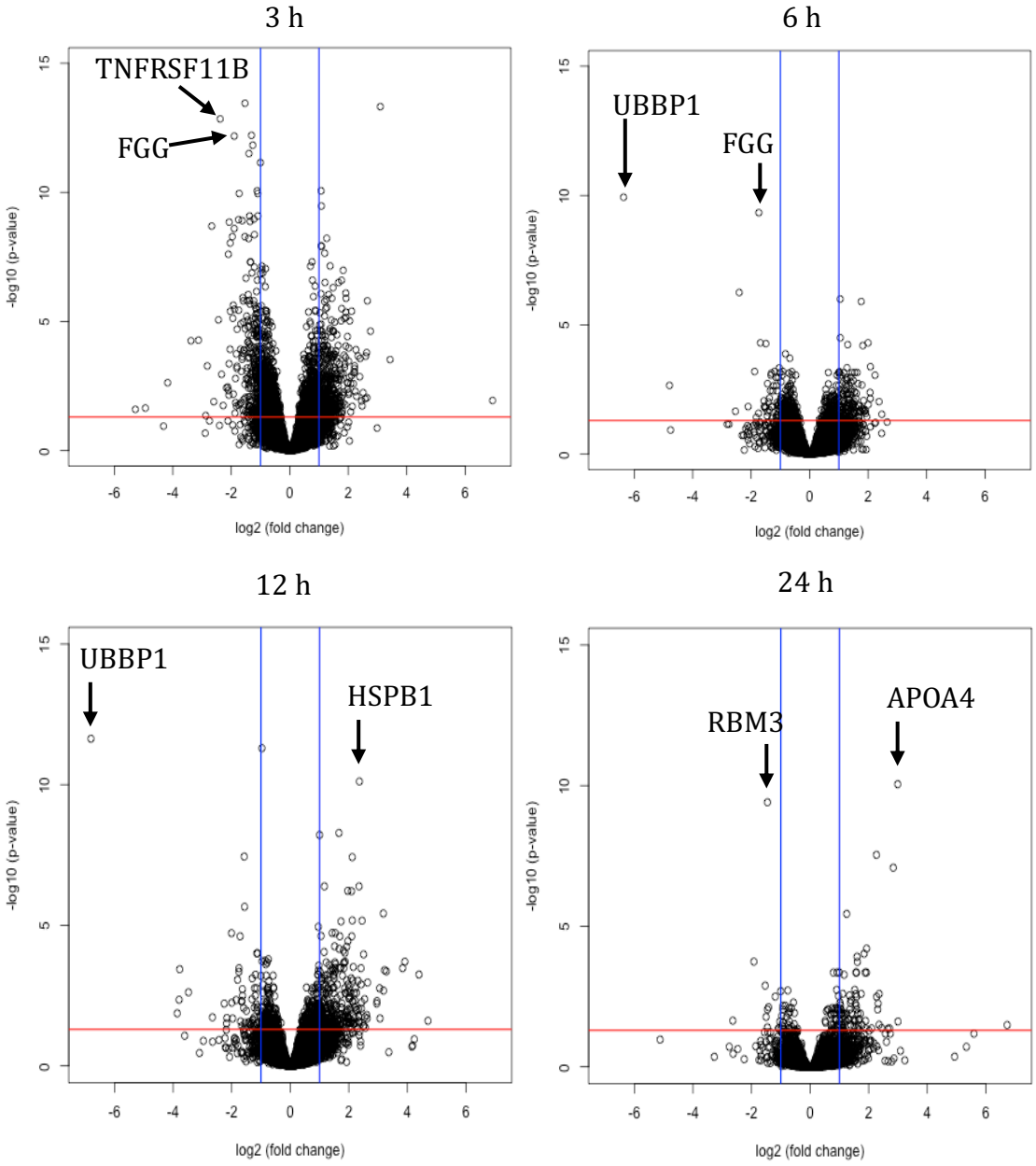


Figure 17. Volcano plots to illustrate the distribution of human genes between uninfected controls and infected samples at each timepoint.

Significantly differentially expressed genes can be visualised as having a $\log_2$(fold change) greater than 1 or less than -1 (blue lines) and adjusted *p*-value greater than 0.05 (red line). Arrows point out selected highly significantly differentially expressed genes.

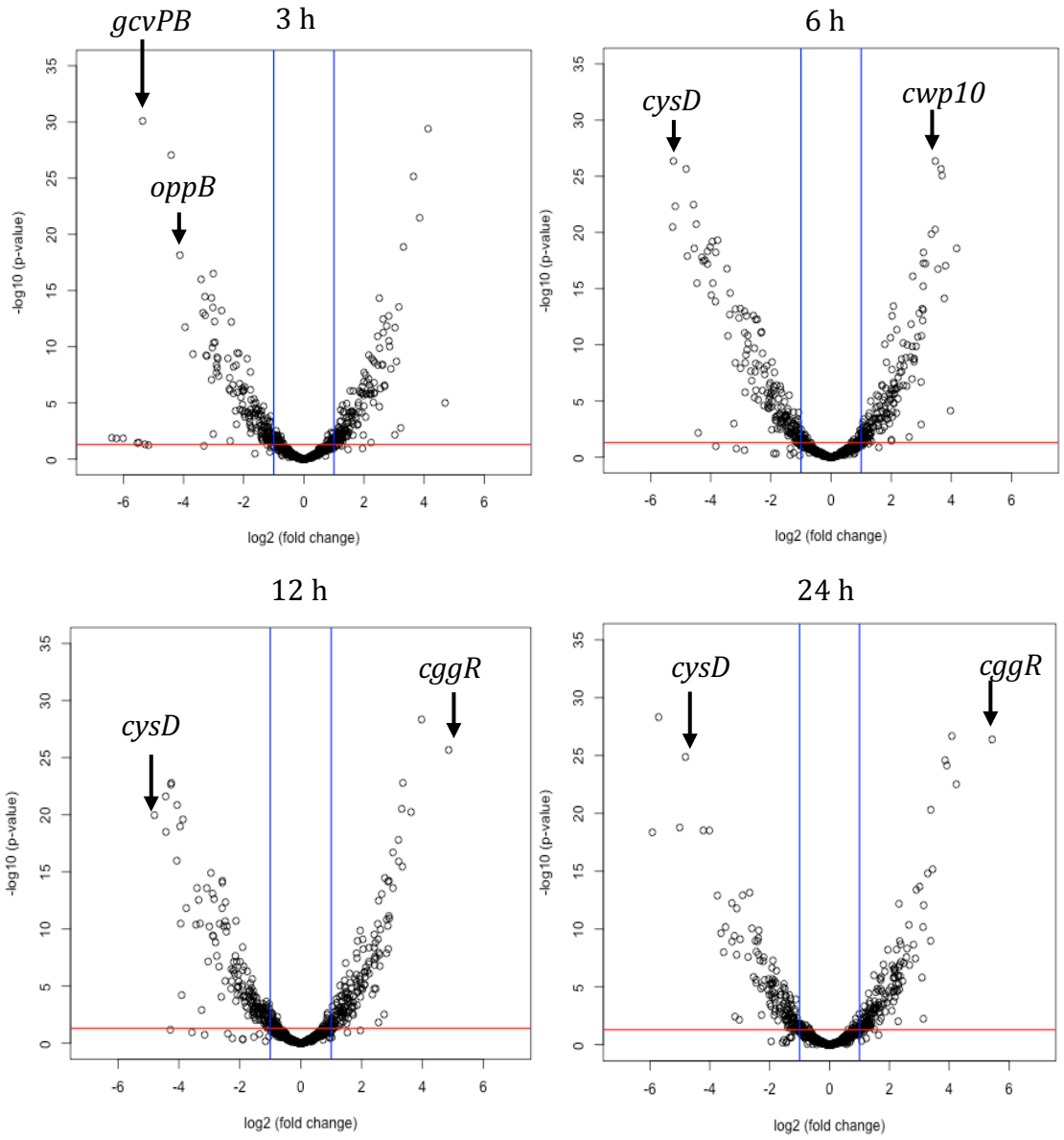


Figure 18. Volcano plots to illustrate the distribution of bacterial genes of uninfected controls vs infected samples at each timepoint.

Significantly differentially expressed genes can be visualised as having a $\log_2$(fold change) greater than 1 or less than -1 (blue lines) and adjusted p-value greater than 0.05 (red line). Arrows point out selected highly significantly differentially expressed genes.

Venn diagrams show the number of common up- or downregulated host genes between each of the timepoints. From the Venn diagrams, it can be calculated that there was a total of 705 significantly upregulated human genes, 291 of which

were only upregulated at 3 h post infection while there were only 7 upregulated at all timepoints. There were also 360 significantly downregulated human genes, 250 of which were only downregulated at 3 h post infection while there was only 1 gene downregulated at all timepoints. Therefore, the majority of the differentially expressed host genes were present at 3 h only. There was a total of 230 *C. difficile* genes which were identified as being significantly upregulated in at least one timepoint, 81 of which were upregulated at all timepoints. There were also 250 *C. difficile* genes which were identified as being significantly downregulated in at least one timepoint, 78 of which were significantly downregulated at all timepoints (Figure 19). Therefore, in contrast to the human analysis, the majority of the bacterial differentially expressed genes were differentially expressed at all four timepoints (lists available in Appendix tables 8-9).

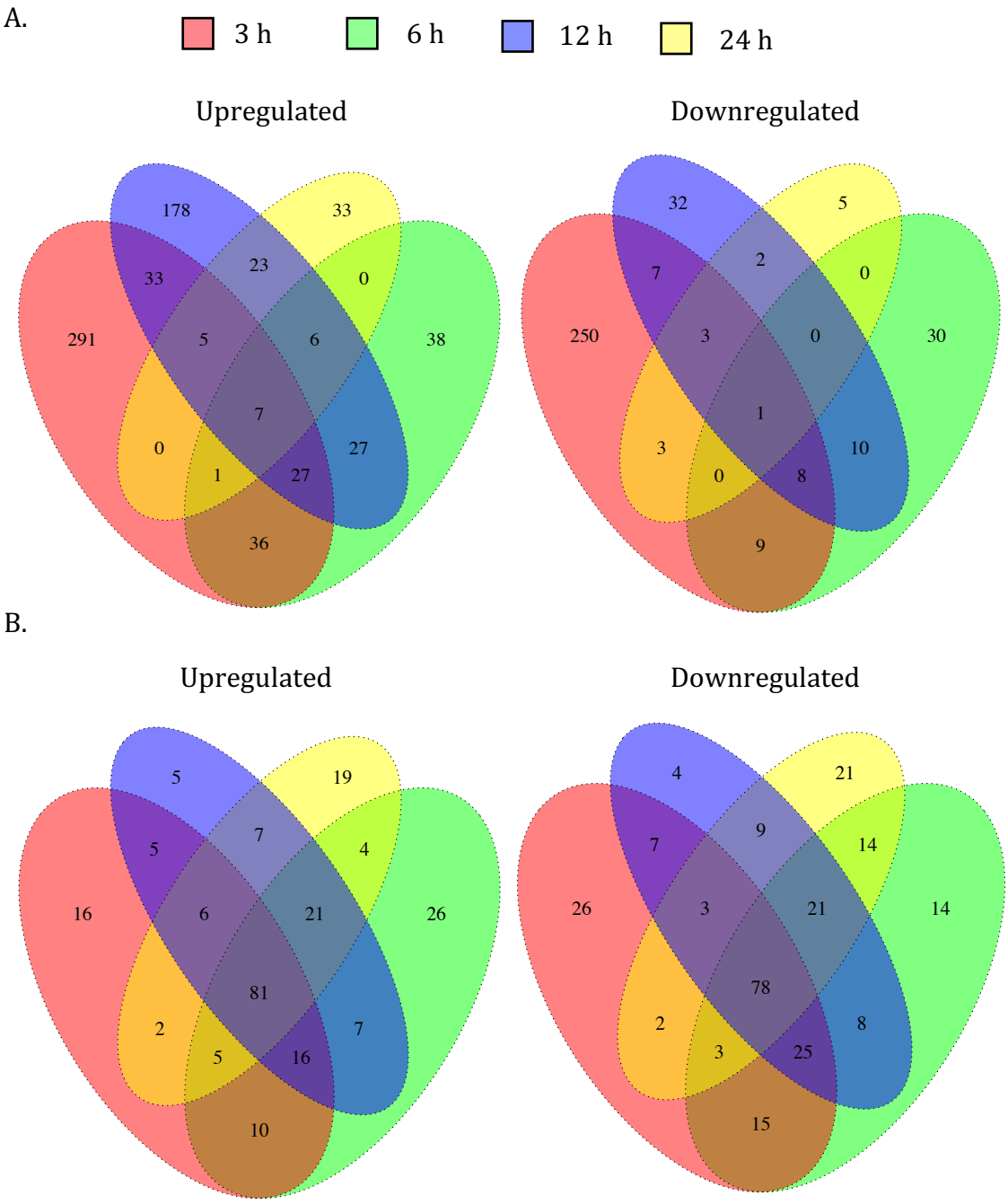


Figure 19. Venn diagrams of the differentially expressed host (A) and bacterial (B) genes to visualise genes commonly up- or downregulated between each timepoint.

3.3 Discussion

C. difficile adhesion to the gut mucosa is critical for the establishment of infection and progression of the disease. However, the specific interactions between vegetative *C. difficile* and the host have been understudied. RNA-seq is a powerful tool which has greatly expanded our understanding of biological processes which occur within cells. The resolution of transcriptomic profiles can be used to examine modifications in gene expression in response to environmental stimuli. Dual RNA-seq is a sequencing methodology used to study transcriptional responses during a host-microbe interaction. Host and bacterial total RNA are extracted from infected cells and high-throughput sequencing approaches are used to sequence and quantify the RNA transcripts. The gene expression profiles of infected samples and uninfected controls are compared to capture alterations in gene expression in response to infection. This methodology has been successfully applied to a variety of pathogens and has improved our current understanding of the molecular processes which facilitate infections (Westermann, Barquist and Vogel, 2017). In this study, the VDC system was used to facilitate *C. difficile* infection of intestinal epithelial cells and the dual RNA-seq technique was applied to investigate the transcriptional responses during *C. difficile* infection.

There are several challenges associated with the dual RNA-seq technique, including the large differences in quantities of cellular RNA between bacterial and eukaryotic cells. There are several possible ways to overcome this hurdle. Firstly, it is particularly important to efficiently remove rRNA transcripts which make up the majority of the total RNA sample but hold little value in

transcriptomic analyses. Depletion of rRNA transcripts allows more sequencing adaptors to be available for information rich mRNA transcripts during the sequencing process. Other techniques to overcome the large differences in host and bacterial RNA abundances include using a high depth of sequencing, enriching for bacterial transcripts prior to sequencing and enriching for infected cells using FACS (Westermann, Barquist and Vogel, 2017). The quantities of host and bacterial transcripts from *C. difficile*-infected cells were confirmed with RT-qPCR using primers for the GAPDH and *gyrA* housekeeping genes. It was found that there was a relatively small amount of bacterial *gyrA* transcripts, 3600 fold less than host GAPDH transcripts. However, the bacterial transcripts were successfully amplified from the mixed sample, suggesting that bacterial transcripts were likely to be able to be sequenced.

Before carrying out the dual RNA-seq experiment with a large number of samples, a small-scale pilot study was conducted to ensure bacterial transcripts could be sequenced and to allow the opportunity to optimise the experimental workflow if required. For the samples used in the pilot study, the Illumina RiboZero Gold epidemiology kit was used to remove human and bacterial rRNA simultaneously. The percentage of reads mapped to the *C. difficile* genome varied from ~0.05-1.15%. Previous dual RNA-seq studies have reported a wide range of percentages of reads aligned to the bacterial genome, between 0.002% to 16% depending on the bacterial pathogen and the processes used to prepare the samples for sequencing (Westermann, Barquist and Vogel, 2017). The *C. difficile* genome coverage values were greater than 1 in 5 out of the 6 samples, suggesting that each gene in the *C. difficile* genome has been sequenced at least once (Figure

8). Together, these results suggested that the percentages of bacterial reads and genome coverages obtained in this analysis were acceptable to be used for dual RNA-seq analyses. However, despite obtaining ample coverage of the *C. difficile* genome there were no bacterial differentially expressed genes identified in this experiment, which may be due to similar bacterial gene expression profiles at both 3 h and 24 h post infection.

Two of the human samples had a percentage of genes aligned to the human genome lower than 67% ('3 h 1' – 61.83% and '3 h 2' - 66.17%) (Figure 8). Files containing the unmapped reads were generated and BLAST was used to identify sequences that aligned to the unmapped reads to investigate their origin. It was found that the unmapped sequences aligned mostly to the *Cyprinus carpio* genome. This suggested that there was likely to be a large quantity of Illumina sequencing adaptors leftover in the samples. A small percentage of the unmapped reads also aligned to the PhiX genome which was added to the pooled samples as a sequencing control. The genome coverage value was also greater than 1 in 5 out of the 6 human samples (Figure 8). Overall, this pilot study demonstrated that the sequencing depth used in this analysis provided acceptable coverage of the human and *C. difficile* genomes.

After successful sequencing was performed in the pilot study, the dual RNA-seq experiment was then carried out on a larger scale. The RNase H method for rRNA depletion was used in the preparation of uninfected controls and new infected samples after the discontinuation of the Illumina RiboZero kit. Prior to sequencing, analysis of the bioanalyzer traces showed complete removal of the

usual peaks associated with ribosomal RNA transcripts. However, after sequencing, it was found that there was a large range of percentages of reads which mapped to the human and *C. difficile* rRNA sequences, ranging from 2.88-89.88% (Table 7). Therefore, the RNase H method was not effective at removing all rRNA transcripts, suggesting that sequencing adaptors were taken up by rRNA transcripts rather than mRNA transcripts during sequencing, reducing the output of valuable mRNA reads. It is possible that RNase H may have cleaved the rRNA transcripts into 200-400 bp fragments which appeared to make up the majority of the samples in the post rRNA depletion bioanalyzer traces (Figure 11). At this size they were unable to be separated from mRNA by the AMPure XP bead purification step in the RNase H rRNA depletion protocol. Another possibility is that the DNase step in the rRNA removal protocol did not work efficiently, which would result in DNA primers complementary to rRNA sequences being leftover in the RNA sample and subsequently sequenced. Since the DNA probes were designed to be 50 bp in length, a DNA bioanalyzer with the RNA samples could potentially be used to investigate whether there were large concentrations of 50 bp DNA fragments present in the samples to determine if this was a source for rRNA transcript sequencing.

Differential gene expression analysis was conducted between the infected samples in the pilot study as well as the large-scale study. The significantly differentially expressed genes identified when comparing the uninfected controls with the pilot study (treated with RiboZero) or large-scale study (treated with RNase H) at 3 h and 24 h post infection were very different. There were also a larger number of differentially expressed genes identified in the pilot

study. This suggests that the RNase H method for rRNA depletion may have had a negative effect on the mRNA transcripts, potentially through nonspecific binding of mRNA to the rRNA oligonucleotide probes. This may result in the formation of ssDNA-mRNA hybrids which could have been cleaved during treatment with RNase H. The annealing temperature in the rRNA removal protocol was 45˚C which is potentially low enough for nonspecific binding to occur between the ssDNA probes and mRNA transcripts. Although the RNase H rRNA removal protocol has been previously optimised for human samples, further optimisation with the samples from this study was likely required to provide more effective rRNA depletion and preservation of mRNA transcripts. To optimise this technique, the purchase of further sequencing cartridges would be required to accurately assess the quantity of rRNA transcripts leftover in the sample. This was not pursued due to the large financial costs involved.

Unlike in the pilot study, the FASTQ files from this analysis were trimmed prior to alignment to remove sequencing adaptors and reads with a Q-score <30. All samples had more than 95% of reads aligning to the concatenated reference genome file which contained both the human and *C. difficile* genome sequences in a .fasta format (Table 5). The unmapped reads were investigated using BLAST and were found to be mostly from the PhiX sequencing control. The percentage of reads mapping to the single *C. difficile* genome ranged from 0.06-2.06% (Table 5). These percentages were comparable to those obtained in other published dual RNA-seq analyses (Westermann, Barquist and Vogel, 2017). During the differential gene expression analysis by DESeq2, counts for each of the gene feature were normalised to the total number of reads in each sample. Since there

were separate human and bacterial counts matrix files at this stage of the data analysis process, the infected samples with a smaller percentage of reads aligned to the bacterial genome will be normalised and can be compared more accurately to the samples with higher percentages of alignment. Therefore, the bacterial reads obtained were acceptable to perform bacterial differential gene expression analysis.

To visualise variations in the transcriptomic profiles, principal component analyses were performed. In the human PCA plot, one 3 h and one 6 h uninfected controls were genetically distant to the other replicates of that group. The genetically distant samples were not from a particular experimental set of samples. In both cases, the removal of these outliers did not largely alter the lists of genes which were significantly differentially expressed. However, there was an increase in the adjusted *p*-values of many of the original significantly differentially expressed genes due to only 2 replicates remaining in the uninfected control groups. Therefore, the decision was made to keep these samples in the analysis to improve statistical significance and stringency of the gene lists. There were alterations in gene expression over the course of the experiment in both the infected samples and uninfected controls. The changes in gene expression identified between the uninfected controls suggests that incubation in the VDC system elicits alterations in gene expression. This may also suggest that some of the changes in gene expression over time in the infected samples may be due to incubation in the VDC system as well as *C. difficile* infection. Interestingly, the samples from the 24 h uninfected control group were clustered more closely to the late-stage infected groups than the other uninfected

groups (Figure 13A). This suggests that extended incubation in the VDC system may be causing transcriptional alterations similar to the responses observed during *C. difficile* infection. The intestinal epithelial cells were incubated with both anaerobic and anaerobic gas in the VDC system. Over time, it is possible that the cells could become starved of oxygen and elicit a stress response at a transcriptional level. For example, oxygen deprivation has been reported to stimulate IL-8 secretion, a cellular response which also occurs during *C. difficile* infection (Desbaillets *et al.*, 1997; Linevsky *et al.*, 1997a; El Feghaly *et al.*, 2013). In the bacterial principal component analysis, two bacterial replicates were genetically distant from the other replicates at the relevant timepoints (Figure 13B). These samples were both from the same experiment (experiment number '2') at the 3 h and 6 h timepoints. These outlier samples also had some of the lowest percentage of reads mapping to the *C. difficile* genome (0.19% at 3 h and 0.06% at 6 h) and had low genome coverage values (Table 6). Removal of these outlier samples resulted in alterations in the lists of differentially expressed genes, suggesting that their inclusion was causing biases in the data. Therefore, these samples were removed from the subsequent analyses (Figure 13C).

Heatmaps are another technique which can be used to investigate the variations between individual samples and groups. Analysis of the heatmap from both the human and bacterial transcriptomic data showed that the majority of the samples from each group were clustered together and the results were consistent with what was observed in the PCA plot (Figure 14A). Analysis of the heatmap from the bacterial transcriptomic data showed that there was a clear transcriptional shift between the uninfected control group and all infected

timepoint groups, demonstrated by a large distance on the dendrogram. There were some samples in the 6 h and 12 h groups which were clustered more closely to the samples from other groups (Figure 14B). This suggests that there may be a large alteration in transcription during infection, but there may be fewer changes in the gene expression profile over the course of infection.

The total number of significantly differentially expressed human genes was highest at 3 h post infection and decreased at the later timepoints (Figure 15A, Figure 19). A possible reason for this is that incubation in the VDC system may have an effect on the gene expression profiles of the uninfected control samples to make them more similar to the infected samples. To investigate this further, alterations in gene expression between the uninfected control groups were investigated to identify genes modulated by incubation in the VDC system. It was found that there were very low numbers of differentially expressed genes between 3 h and 6 h, 6 h and 12 h and 12h and 24 h. However, when comparing 3 h to 12 h and 24 h as well as 6 h to 24 h there were large numbers of differentially expressed genes (Figure 16). This analysis highlighted the impact of VDC incubation on the host cells. Since incubation in the VDC system changes the gene expression profiles of uninfected control samples over time, it is possible that the reason for the low numbers of differentially expressed genes at 24 h may be due to the VDC incubation inducing similar stress response pathways as *C. difficile* infection. As well as oxygen deprivation, another possibility which may evoke a cellular stress response in the VDC system is the accumulation of toxic metabolic by-products over time. A fluidic model may be a better option for performing sensitive transcriptomic experiments.

To summarise this chapter, a small-scale pilot study was conducted to ensure the dual RNA-seq experiment could be successfully executed before performing the experiment on a larger scale. Several methods were used to visualise and interpret the dual RNA-seq data to examine the variations between samples and groups, as well as the distributions of differentially expressed genes.

Chapter 4: Transcriptomic responses of *C. difficile* during infection

4.1 Introduction

C. difficile adhesion to the intestinal mucosa is crucial for the establishment of infection. There have been numerous bacterial factors which have been reported to be involved in *C. difficile* adhesion to host tissues. These factors include the flagella, pili, cell surface proteins and ECM-binding proteins (Calabi *et al.*, 2002; Barketi-Klai *et al.*, 2011; Tulli *et al.*, 2013; McKee *et al.*, 2018). However, we don't have a comprehensive understanding of the transcriptional alterations which occur in response to infection. A better understanding of the transcriptional changes during *C. difficile* infection may provide insights into the mechanisms used by *C. difficile* to adhere to host cells and cause disease. This knowledge could aid in the development of novel diagnostic tools and treatment strategies.

In this study, a dual RNA-seq approach was used to resolve the host and bacterial transcriptional responses during *C. difficile* infection of human intestinal epithelial cells in a VDC *in vitro* gut model. This part of the study focuses on the analysis of bacterial transcriptomic changes and the investigation of genes and pathways with potentially important roles in *C. difficile* pathogenesis.

4.2 Results

Comparisons were made between the transcriptomic profiles of a single bacterial culture grown in DMEM-10 for ~5 h to log-phase and *C. difficile* from infected

samples at different timepoints incubated in the same medium. The analysis of transcriptomic profiles can establish bacterial gene expression alterations in response to infection.

4.2.1 Single gene expression profiles

There were a large number of interesting differentially expressed *C. difficile* genes identified in this study which have roles across many cellular processes, including virulence, colonisation and metabolism. To analyse the changes in gene expression over the course of the infection, single gene expression profiles were generated by plotting the fold changes between the bacterial control and infected samples at each timepoint.

4.2.1.1 Transcriptional modulation of genes involved in *C. difficile* virulence

In this analysis, *codY* (CDR20291_1115) was significantly downregulated at 6 h, 12 h and 24 h after infection. CodY is a global transcriptional regulator found in many Gram-positive bacteria, including *Bacillus* species. CodY controls the expression of many genes involved in *C. difficile* nutrient acquisition and virulence, including the repression of toxin production and sporulation (Dineen *et al.*, 2007; Dineen, McBride and Sonenshein, 2010; Nawrocki *et al.*, 2016). Proline-Proline endopeptidase 1 (PPEP-1/CDR20291_2721) was also significantly downregulated at 3 h, 12 h and 24 h post infection. PPEP-1 is a secreted zinc metalloprotease which cleaves the CD2831 and CD3246, cell-surface associated collagen binding proteins and may have a role in *C. difficile* virulence (Hensbergen *et al.*, 2014, 2015). *fliC* (CDR20291_0240) was also

significantly downregulated at 12 h post infection (Figure 20). The flagellin protein, FliC, may play a role in *C. difficile* adhesion to host tissues, and a *fliC* mutant exhibited increased toxicity in a hamster model of infection, suggesting it may also play a role in toxin production and virulence (Tasteyre *et al.*, 2001; Dingle, Mulvey and Armstrong, 2011; Baban *et al.*, 2013).

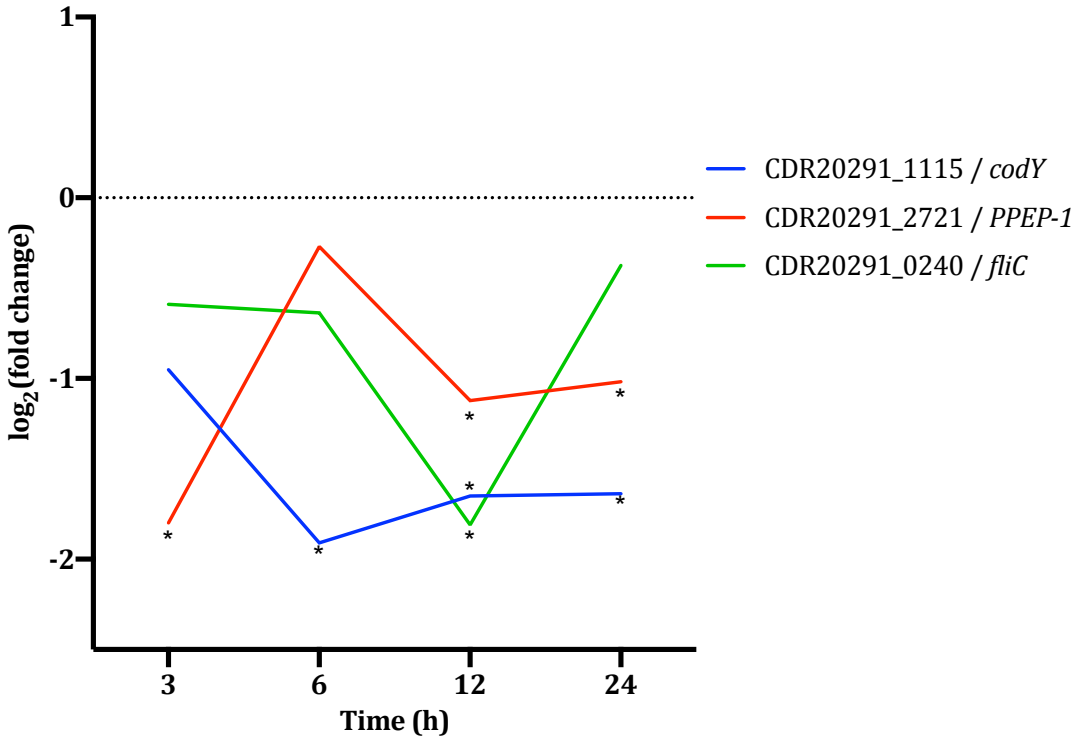


Figure 20. Single gene expression profiles of selected *C. difficile* virulence associated genes.

Statistical significance indicated with an asterisk.

4.2.1.2 Oligopeptide transporters

The *C. difficile* genome encodes two oligopeptide transport systems, Opp and App. In this analysis, the *opp* operon (CDR20291_0783-87) was significantly downregulated at 3 h, 6 h and 12 h after infection. The *app* oligonucleotide

transporter operon was also modulated during *C. difficile* infection, where *appB* (CDR20291_2561) was significantly downregulated at all timepoints after infection, and *appA* (CDR20291_2560) or *appC* (CDR20291_2562) were significantly downregulated in at least one timepoint (Figure 21). However, despite the rest of the *app* operon being downregulated, the first gene in the operon, *appF* (CDR20291_2558) was significantly upregulated at all timepoints after infection. The downregulation in oligonucleotide transport operons may have a role in *C. difficile* pathogenesis, as disruption in oligopeptide transporters resulted in an increase in sporulation frequency and virulence in a hamster model (Dineen, McBride and Sonenshein, 2010; Edwards, Nawrocki and McBride, 2014; Nawrocki *et al.*, 2016).

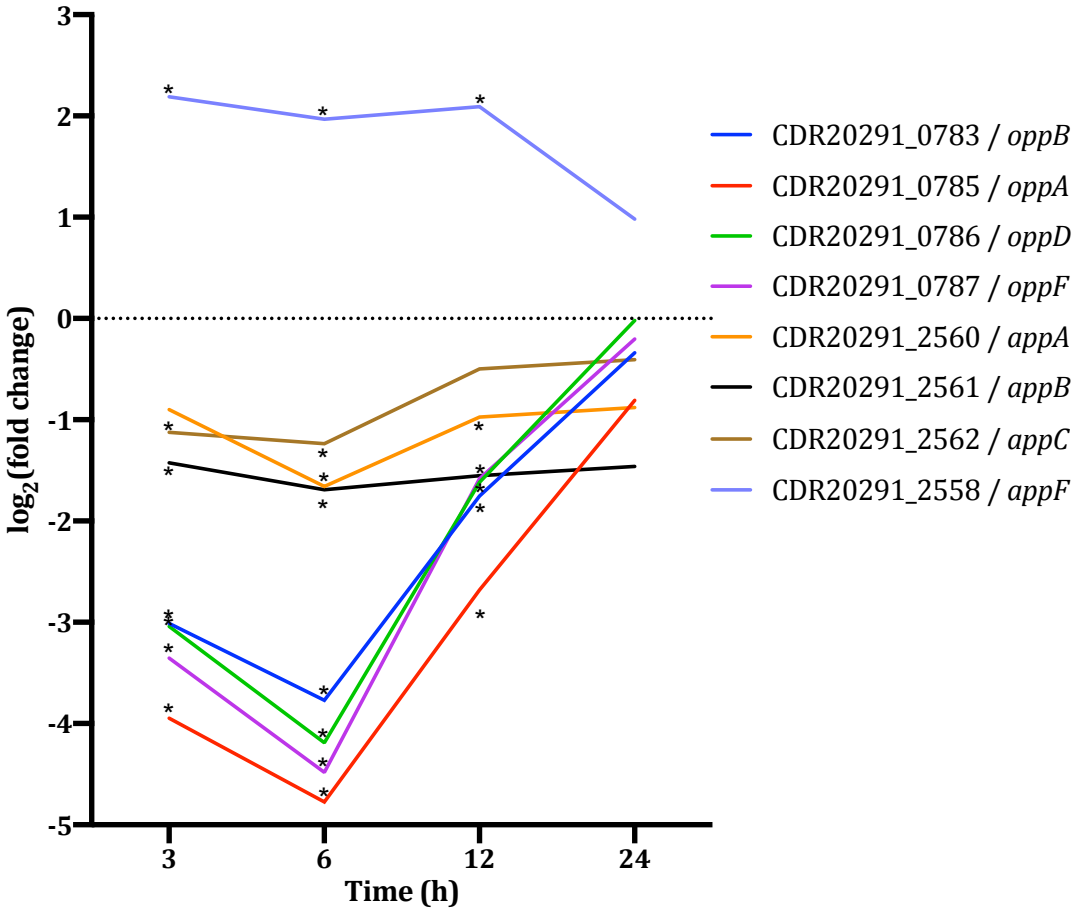


Figure 21. Single gene expression profiles of *C. difficile* oligopeptide transporters at different time points during *C. difficile* infection.

Statistical significance indicated with an asterisk.

4.2.1.3 The *C. difficile* cell surface proteins are modulated during infection

Interestingly, the *cwp84* (CDR20291_2676), a cell wall hydrolase required for the formation of the S-layer was significantly downregulated at all timepoints. *slpA* (CDR20291_2682), the precursor for the S-layer proteins, was also significantly downregulated at 6 h post infection (Figure 22). The modulation of numerous cell-well proteins was also found during *C. difficile* infection. Cwp66 (CDR20291_2678), a cell wall protein suggested to play a role in adhesion to Vero cells (Waligora *et al.*, 2001), *cwp13* (CDR20291_1645), *cwp14* (CDR20291_2624),

cwp17 (CDR20291_0892), *cwp18* (CDR20291_0903), *cwp19* (CDR20291_2655), and *cwp20* (CDR20291_1318) were significantly downregulated at all timepoints after infection. Despite the majority of the cell wall proteins being downregulated during *C. difficile* infection, *cwp10* (CDR20291_2685), CDR20291_0184 (putative cell wall hydrolase), and CDR20291_2686 were highly upregulated at all timepoints after infection (Figure 22). Little is known about these three proteins, but CDR20291_2686 is a putative exported protein which has a fibronectin-binding domain, suggesting that it may have a role in *C. difficile* colonisation and adhesion to host cells or the ECM. The upregulation observed in this experiment suggests these proteins may have an important role in *C. difficile* pathogenesis.

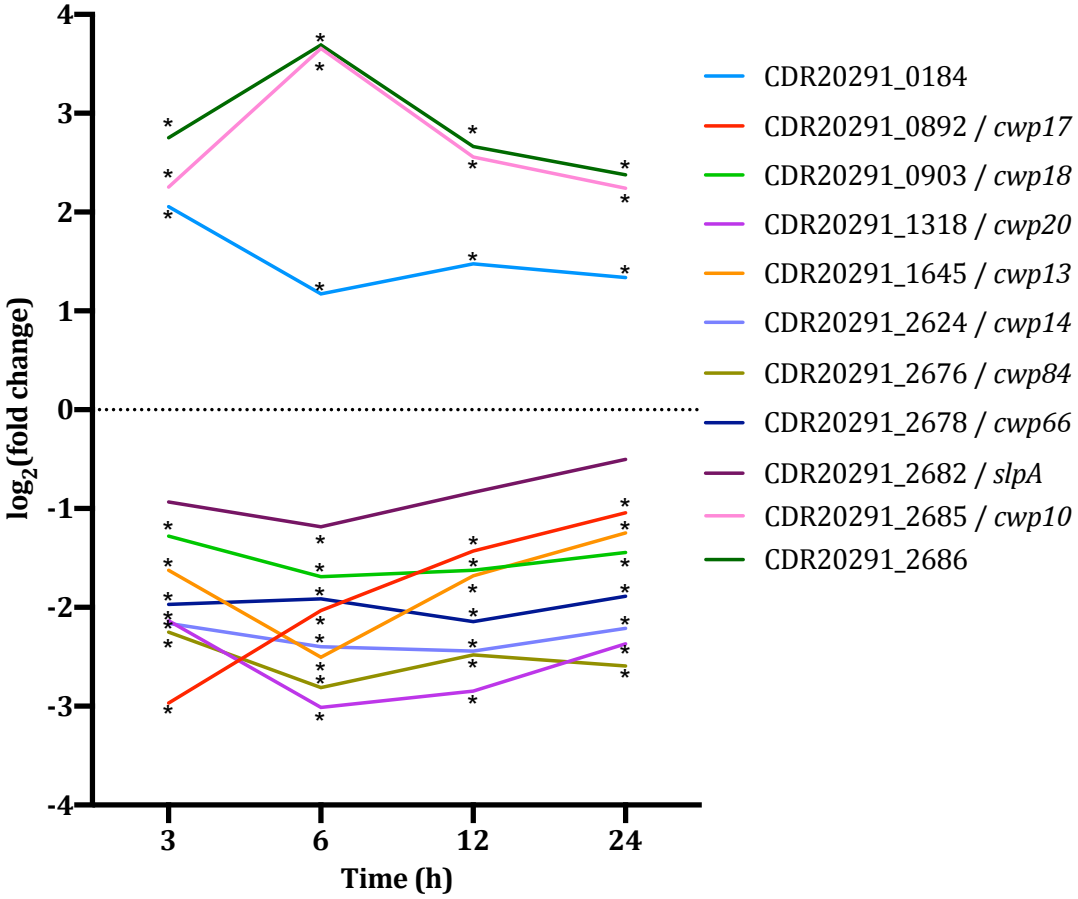


Figure 22. Single gene expression profiles of significantly differentially expressed *C. difficile* cell-surface proteins at different timepoints of infection.

Statistical significance indicated with an asterisk.

4.2.1.4 Differential expression of genes involved in iron transport during *C. difficile* infection

The modulation of genes involved in iron transport was also observed during *C. difficile* infection of intestinal epithelial cells. Fur (CDR20291_1127), the ferric uptake regulation protein, was highly upregulated at 6 h, 12 h and 24 h post infection (Figure 23). Fur is a transcriptional regulator which controls gene expression in response to iron levels by binding to target promotors in the presence of iron (Ho and Ellermeier, 2015). The deletion of the *fur* gene in *C.*

difficile resulted in the modulation of many genes, including the upregulation of the *feo1* ferrous iron transport operon (CDR20291_1327 - CDR20291_1328) as well as CDR20291_0516, a putative cation transporting ATPase, indicating that the expression of these genes are suppressed by Fur (Ho and Ellermeier, 2015). Consistent with this, the *feo1* operon and CDR20291_0516 were significantly downregulated at all timepoints in this analysis, possibly due to suppression by the increased expression levels of Fur. CDR20291_1639, putative ferrous iron transport protein A, was also highly downregulated at all timepoints after infection. Many genes involved in iron uptake were also upregulated during infection. *fhuD* (CDR20291_2774), a hydroxamate siderophore ABC transporter and substrate-binding protein, exhibited a high level of upregulation at all timepoints. CDR20291_1545, a putative iron compound ABC transporter and permease protein was also significantly upregulated at 6 h, 12 h and 24 h post infection (Figure 23). The modulation of iron uptake has been previously reported during *C. difficile* infection (Ho and Ellermeier, 2015; Fletcher *et al.*, 2018), demonstrating the importance of this nutritional requirement during pathogenesis.

In addition to iron transporters, Fur also regulates the expression of ferredoxins and flavodoxins, electron transfer proteins involved in numerous metabolic processes (Yoch and Valentine, 1972; Ho and Ellermeier, 2015). Ferredoxins, which require iron-sulphur clusters to coordinate electron transfer, are induced by Fur or high-iron conditions, while flavodoxins, which do not require iron are repressed by Fur. In this analysis, the CDR20291_3444 ferredoxin was highly upregulated at the early stages of infection (3 h and 6 h post infection), and the

flavodoxin, *fldX* (CDR20291_1925), was upregulated at 6 h, 12 h and 24 h post infection (Figure 23). These changes in gene expression may reflect the availability of iron in the VDC system.

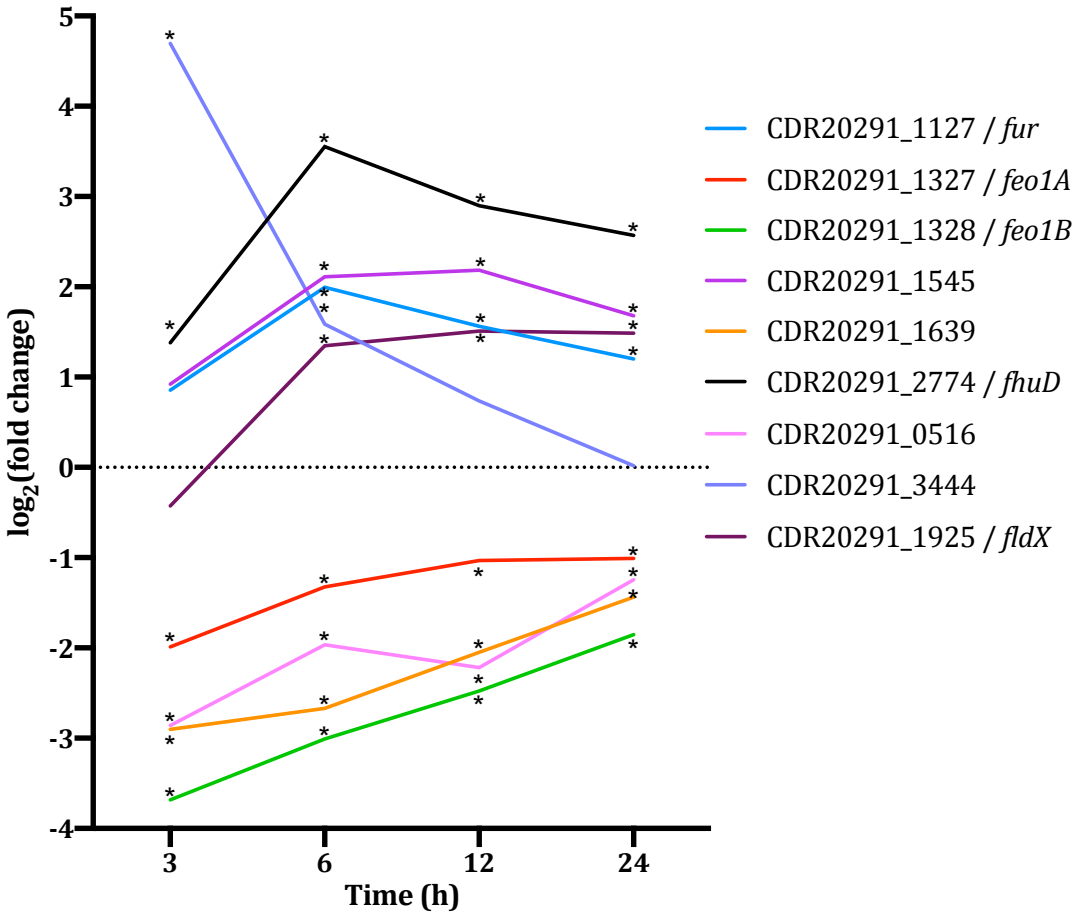


Figure 23. Single gene expression profiles of *C. difficile* genes involved in iron acquisition.

Statistical significance indicated with an asterisk.

4.2.2 Pathway enrichment analysis

To investigate the molecular processes which are modulated in response to *C. difficile* infection, a pathway enrichment analysis was conducted with the differential gene expression data from the dual RNA-seq experiment. Since the

gene annotations for *C. difficile* R20291 are not included in any pathway analysis database, sequencing reads were aligned to the *C. difficile* strain 630 reference genome as this is the most closely related strain included in pathway databases. The STRING database (available online at https://string-db.org) is a well-known database on collected associations between proteins from multiple sources, and can analyse protein-protein interactions, as well as perform functional enrichment analyses.

To investigate protein-protein interactions during *C. difficile* infection, the functional enrichment of Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathways were investigated. The up- or downregulated genes at each timepoint were entered into the STRING database online tool to identify functional pathways enriched during *C. difficile* infection. Entering up- and downregulated genes separately rather than all of the differentially expressed genes together allows for a better understanding of the positive and/or negative modulation of each pathway. The differential expression of the pathways modulated during *C. difficile* infection were visualised with heatmaps, where the colour scale represents the statistical significance ($\log_{10}$(false discovery rate (FDR)), Figure 24).

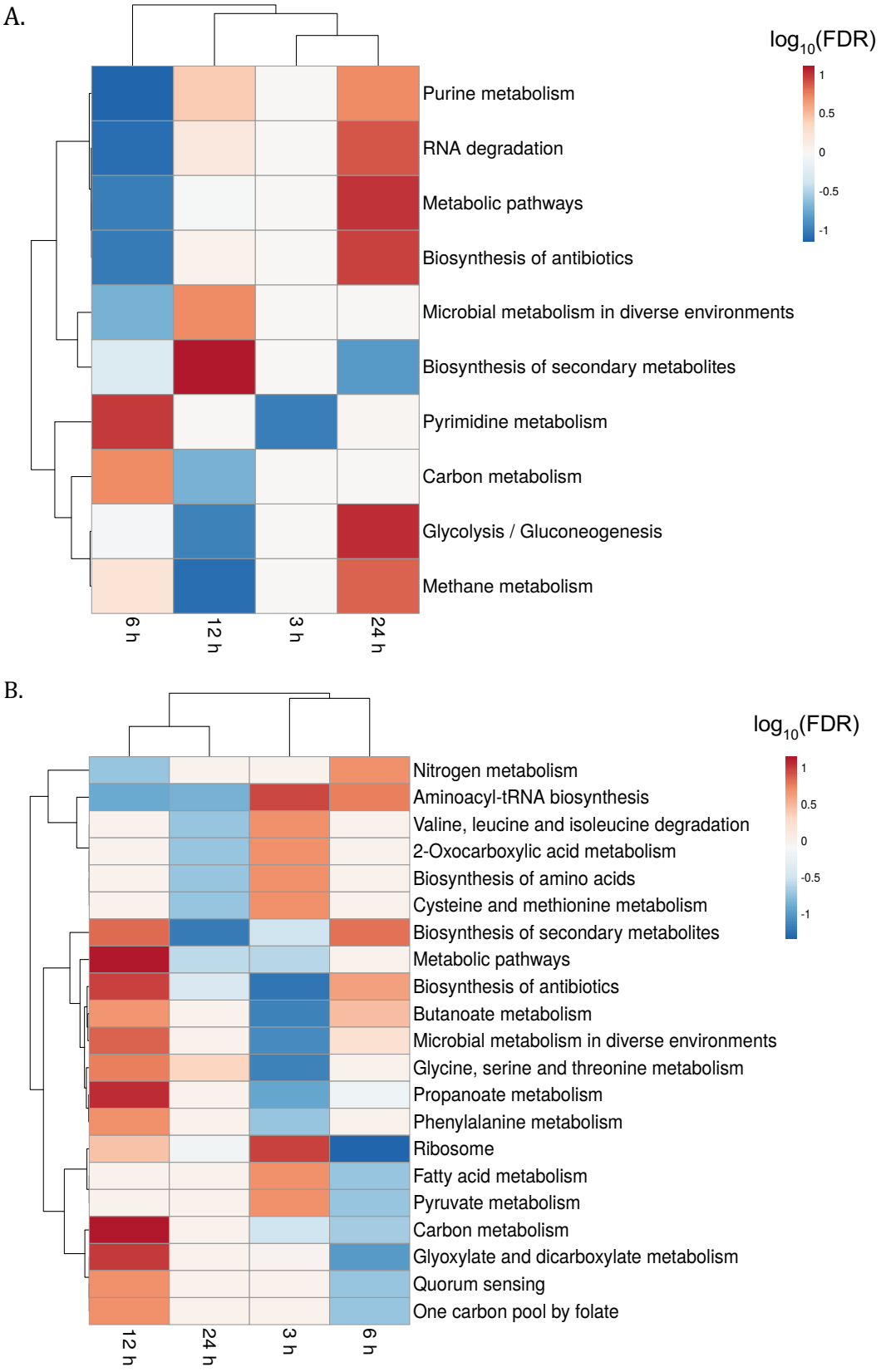


Figure 24. *C. difficile* significantly differentially expressed KEGG pathways during infection.

A. Heatmaps of the KEGG pathway functional enrichment analysis of significantly upregulated genes during infection. B. KEGG pathway functional enrichment analysis of significantly downregulated genes during infection. Colours are representative of the enrichment significance (log_{10}(FDR)) and the dendrograms illustrate the hierarchical clustering of samples and pathways.

There were many interesting significantly modulated pathways identified in this analysis and the majority of these pathways were associated with the biosynthesis or metabolism of various substrates. The metabolic pathways of purines, pyrimidines, carbon and methane were significantly upregulated during *C. difficile* infection. The biosynthesis of antibiotics, secondary metabolites and gluconeogenesis pathways were also upregulated during *C. difficile* infection (Figure 24A). The metabolic pathways of nitrogen, 2-oxocarboxylic acid, cysteine, methionine, butanoate, glycine, serine, threonine, propanoate, phenylalanine, fatty acid, pyruvate, carbon, glyoxylate and dicarboxylate were significantly downregulated during *C. difficile* infection. The biosynthesis pathways of amino acids, secondary metabolites, aminoacyl-tRNA and antibiotics were also significantly downregulated during *C. difficile* infection. Interestingly, the quorum sensing pathway was also significantly downregulated at 6 h and 12 h post infection, which has a key role in the formation of biofilms (Figure 24B). Some pathways, such as carbon metabolism and biosynthesis of secondary metabolites were identified as both up- and downregulated, suggesting that these pathways were modulated, where some genes were upregulated, and others were downregulated.

4.2.3 PPEP-1 has multiple roles in *C. difficile* pathogenesis

PPEP-1 is a secreted proline-proline endopeptidase which cleaves collagen-binding adhesins on the *C. difficile* cell surface (Hensbergen *et al.*, 2014). A *ΔPPEP-1* mutant strain exhibited an attenuation in virulence in a hamster model of infection, indicating that PPEP-1 may play a role in *C. difficile* virulence (Hensbergen *et al.*, 2015). The gene which encodes PPEP-1 was significantly downregulated at 3 h, 6 h and 24 h after infection. This downregulation may result in reduced cleavage of cell-surface associated adhesins and therefore might have a role in *C. difficile* colonisation of the intestinal epithelial layer. To investigate the role of PPEP-1 during *C. difficile* colonisation, a *ΔPPEP-1* mutant in *C. difficile* strain 630 was compared to the WT and a complemented strain.

4.2.3.1 PPEP-1 deletion does not have an effect on biofilm formation

C. difficile has been reported to form biofilms *in vitro* and during infection of murine models, which may play an important role during *C. difficile* colonisation and pathogenesis (Semenyuk *et al.*, 2015; Soavelomandroso *et al.*, 2017). The PPEP-1 protease is expressed during the motility-associated c-di-GMP ribotype, and cleaves cell-surface proteins which are expressed in the adhesion-associated ribotype, suggesting it may play a role in *C. difficile* virulence (Peltier *et al.*, 2015; Corver *et al.*, 2017). To investigate the role of PPEP-1 in biofilm formation, WT and *ΔPPEP-1 C. difficile* cultures were grown into biofilms for 24 h and the biomass was quantitated using crystal violet staining as described in Methods section 2.11. Bacterial cultures were incubated in BHIS+G media as it has been previously reported to stimulate biofilm formation in *C. difficile* strain 630 (Dapa *et al.*, 2013). There was no significant difference observed in the quantify of

biofilm biomass between WT and *ΔPPEP-1* (Figure 25). This suggests that PPEP-1 may not play a direct role in *C. difficile* biofilm formation. These results are consistent with what was recently reported by Dawson *et al.* who also found that PPEP-1 deletion did not have a significant effect on *C. difficile* biofilm formation (Dawson *et al.*, 2021). However, there was significantly more biofilm formation in the PPEP-1 complemented strain compared to the WT. This strain was incubated with antibiotics to maintain the plasmid in the cells, as well as anhydrotetracycline (ATc) which induces ectopic expression of PPEP-1 from the plasmid. These substances may have elicited a stress response in *C. difficile* which may stimulate the formation of biofilms.

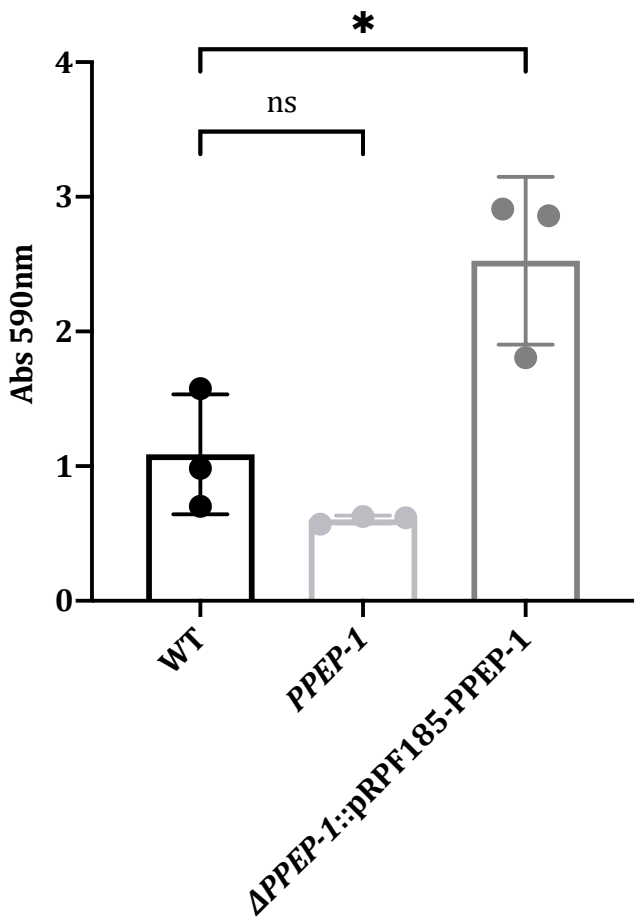


Figure 25. Biofilm formation assay for WT, *ΔPPEP-1* and *ΔPPEP-1*::pRPF185-PPEP-1.

Biofilms were developed over 24 h in BHIS+G media as quantitated by crystal violet staining. Significance tested with One-way ANOVA with multiple comparisons to the WT as a control, $p = 0.013$. Bars are representative of the mean +/- standard deviation.

4.2.3.2 PPEP-1 may play a role in *C. difficile* autoaggregation

Since PPEP-1 cleaves cell-surface adhesins, I hypothesised that it may have a role in the modulation of bacterial autoaggregation. To investigate this, an aggregation assay was conducted with WT and *ΔPPEP-1* to quantify the percentages of aggregated cells in liquid cultures over the log-phase of growth. The OD_{600} of log-phase cultures were measured before (OD_{pre}) and after (OD_{post})

vortexing, and the percentages of aggregated cells were calculated using the calculation (OD_{post} - OD_{pre}/OD_{post}) x 100. There was no significant difference in the rates of growth between WT, *ΔPPEP-1* and *ΔPPEP-1*::pRPF185-PPEP-1 (Figure 26A). However, *ΔPPEP-1* exhibited a higher percentage of aggregation compared to the WT and *ΔPPEP*-1::pRPF185-PPEP-1 complemented strain (Figure 26B). The timepoint with the largest difference between the percentage of aggregation between WT and *ΔPPEP-1* was 4 h. The values at 4 h were investigated for statistical significance with an unpaired t-test. It was found that there was a significantly higher percentage of autoaggregation in *ΔPPEP-1* compared to the WT (**p = 0.0048, (Figure 26C). This suggests that the PPEP-1 protease may have a role in the modulation of cellular aggregation. However, there was no significant difference between the percentages of aggregation between the WT and *ΔPPEP-1*::pRPF185, suggesting that further investigations may be required to determine the role of PPEP-1 in autoaggregation.

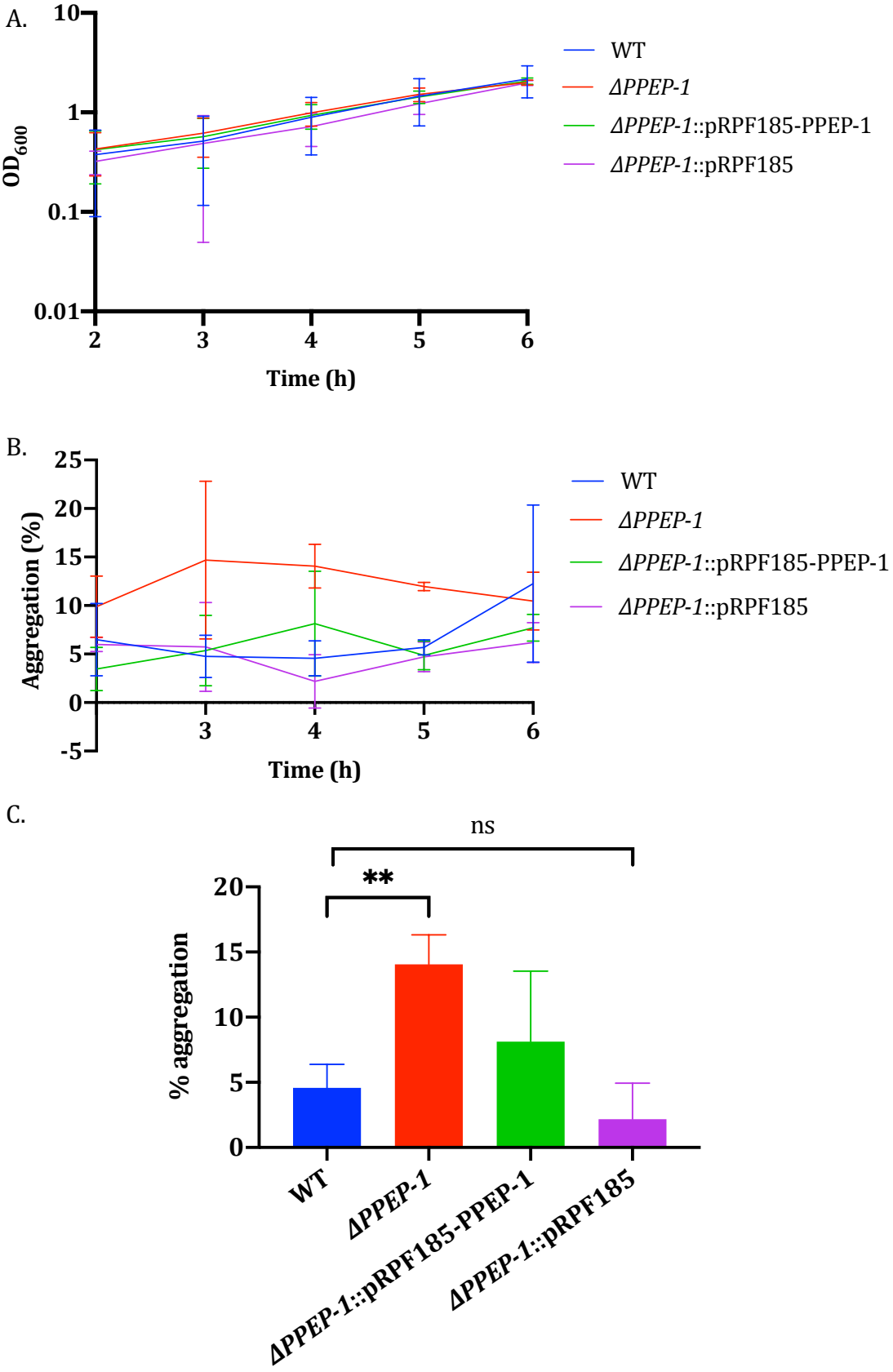


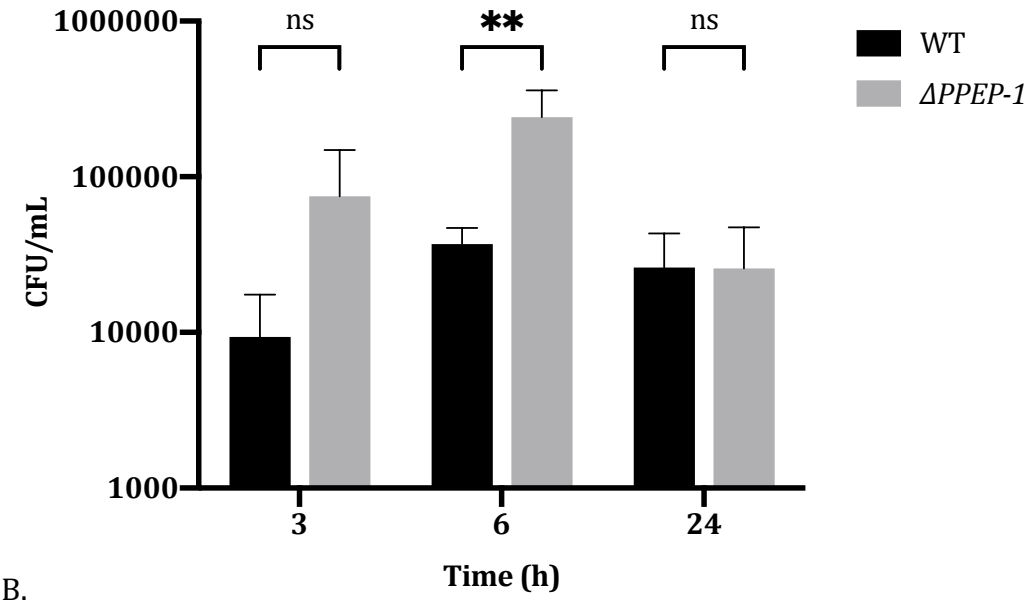
Figure 26. Quantification of *C. difficile* aggregation.

A. Growth curves of WT, *ΔPPEP-1* and *ΔPPEP-1*::pRPF185-PPEP-1 were measured with OD_{600} over time in in BHIS+G medium. B. Aggregation assay of WT, *ΔPPEP-1* and *ΔPPEP-1*::pRPF185-PPEP-1, conducted by monitoring the percentage of aggregated cells over time in BHIS+G medium, calculated with the equation (OD_{post} - OD_{pre}/OD_{post}) x 100. C. The percentage of aggregation at 4 h in each strain. Statistical significance was tested with an unpaired t-test, $^{**}p = 0.0048$. All graphs represent data from three biological replicates (n=3). Bars are representative of the mean +/- standard deviation.

4.2.3.3 *ΔPPEP-1* may play a role in *C. difficile* colonisation of an intestinal epithelial layer

PPEP-1-mediated cleavage of collagen binding proteins may have a role in *C. difficile* colonisation of the intestinal epithelium. To investigate the role of PPEP-1 in *C. difficile* colonisation, intestinal epithelial cells were infected with WT or *ΔPPEP-1* in the VDC system, and the number of bacteria attached to the epithelial cells was quantified with CFU counts. There was a significantly higher number of *ΔPPEP-1* bacterial cells adhered to the intestinal epithelial layers compared to the WT at 6 h post infection ($^{**}p = 0.0064$). At 3 h and 24 h post infection there was no significant difference in adhesion between the WT and *ΔPPEP-1* (Figure 27A). To normalise the number of adhered bacterial cells to the inoculum, the percentage of adhered bacteria was calculated using the equation (CFU of adhered cells/CFU of inoculum) x 100. Overall, a low percentage of bacterial cells from the inoculum adhered to the epithelial cell layer. At 6 h post infection, there was a significantly higher percentage of *ΔPPEP-1* bacterial cells adhered to the epithelial cell layer compared to the WT ($^{*}p = 0.0342$, Figure 27B).

A.



B.

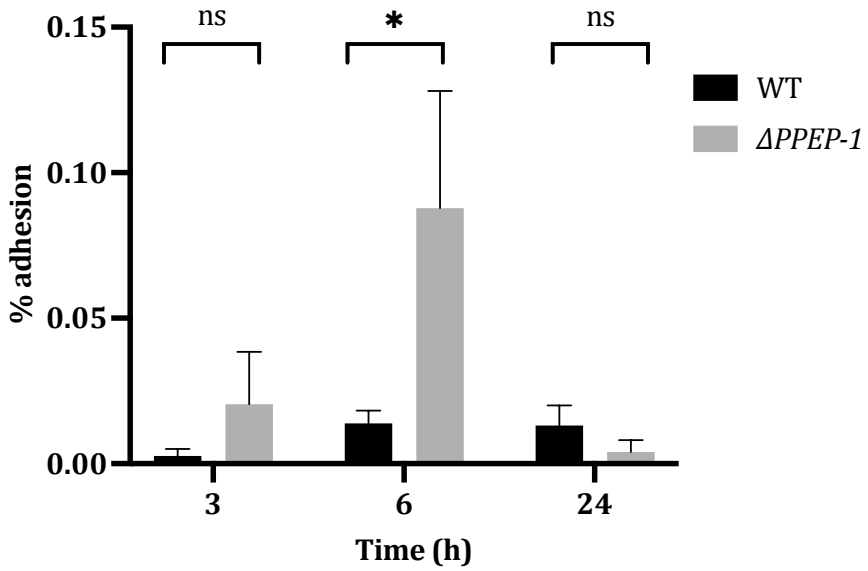


Figure 27. Quantification of WT and *ΔPPEP-1* adherence to intestinal epithelial cells.

A. The number of bacteria attached to intestinal epithelial cells in the VDC system, quantified using CFU counts (n=3) (***p* = 0.003 at 6 h, determined with two-way ANOVA with Sidak multiple comparisons test). B. The percentage of bacterial cells which adhered out of the inoculum, calculated using the equation (CFU of adhered cells/CFU of inoculum) x 100. Statistical significance was tested with an unpaired t-test, **p* = 0.0342. Bars are representative of the mean +/- standard deviation.

Intestinal epithelial layers infected with WT or *ΔPPEP-1* in the VDC system were analysed with confocal microscopy to compare levels of bacterial adhesion. *C. difficile*-infected intestinal epithelial layers were stained with an anti-*C. difficile* antibody to fluorescently label the bacterial cells before imaging with confocal microscopy. Qualitative analysis of the microscopy images showed that there were more *ΔPPEP-1* bacterial cells attached to the epithelial cell layers compared to the WT at 6 h and 24 h post infection (Figure 28A). Quantification of the numbers of *C. difficile* cells attached to the intestinal epithelial cells confirmed that there was significantly more *ΔPPEP-1* adherence to the intestinal epithelial layers at 6 h and 24 h post infection (Figure 28B).

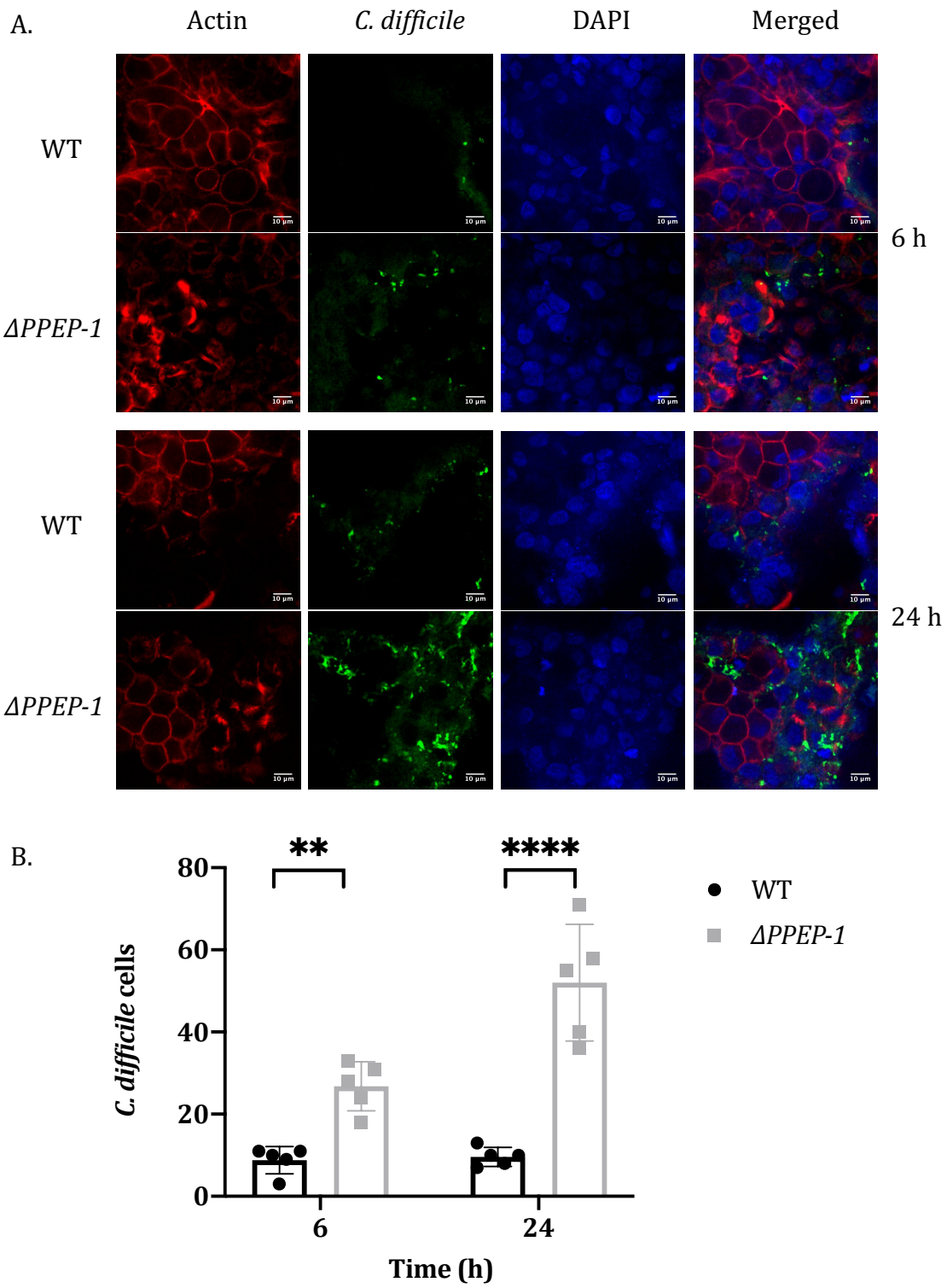


Figure 28. Quantification of WT or *ΔPPEP-1* adhesion to intestinal epithelial cells with confocal microscopy

A. Cell layers were stained with phalloidin (red), DAPI (blue) and anti-*C. difficile* antibody (green) and imaged with confocal microscopy. B. *C. difficile* cells adhered to the epithelial cell layers were counted. Statistical significance tested by two-way ANOVA, $^{**}p = 0.0051$,

****p < 0.0001. Data representative of 3 biological replicates with 5 technical replicates/fields analysed. Bars are representative of the mean +/- standard deviation.

4.2.3.4 PPEP-1 may be involved in *C. difficile* release from collagen

The PPEP-1 protease is expressed when levels of intracellular c-di-GMP are low which activates the expression of motility-associated genes, such as flagella and chemotaxis genes (McKee *et al.*, 2013). However, the CD2831 and CD3246 substrates of PPEP-1 are expressed when c-di-GMP levels are high, and the bacteria are in an adhesive state (Hensbergen *et al.*, 2014). I hypothesised that PPEP-1 may be involved in the transition between sessile and motile lifestyles by cleaving collagen-binding proteins to release the bacteria from the epithelial layer or ECM during infection. To test this hypothesis, the number of WT or *ΔPPEP-1* bacteria adhered to collagen before and after an induced phase shift were compared.

The levels of PPEP-1 secretion from cultures grown in BHIS or BHIS+G media were quantified with a western blot to determine whether it was possible to modulate PPEP-1 expression with different media. It was found that WT *C. difficile* had higher levels of PPEP-1 secretion when cultured in BHIS than BHIS+G, suggesting that incubation in different media can modulate PPEP-1 expression (Figure 29A). Next, WT and *ΔPPEP-1* bacterial cultures were normalised and incubated in a collagen-coated transwell plate with BHIS+G medium to stimulate the expression of adhesion-associated genes and facilitate bacterial adhesion to collagen (Dapa *et al.*, 2013). After 3 h, wells were washed with PBS to remove bacteria which had not adhered to the collagen before

BHIS+G or BHIS media was added to the collagen-adhered bacteria to stimulate different expression levels of PPEP-1. After 5 hours, the numbers bacteria remaining adhered to the collagen were quantified with CFU counts. Surprisingly, the collagen-binding assay showed that there was a higher adhesion to collagen in the WT than *ΔPPEP-1* after the initial and subsequent incubations in BHIS+G media (Figure 29B). After the initial incubation in BHIS+G, the number of *ΔPPEP-1* bacteria adhered to collagen was not significantly altered by the addition of either BHIS+G or BHIS media. However, the addition of BHIS media significantly reduced the number of WT bacteria remaining adhered to the collagen (*$p = 0.01$) (Figure 29B). The fold changes between the number of bacteria before and after the subsequent addition of BHIS+G or BHIS media were also calculated to compare WT and *ΔPPEP-1* release from collagen. The number of WT bacteria adhered to the collagen after the addition of BHIS was reduced by 10 fold, while the number of *ΔPPEP-1* was only reduced by 1.3 fold (Figure 29C). This suggests that the cleavage of cell-surface associated collagen-binding proteins by the PPEP-1 protease may contribute toward the release of bacteria from collagen, potentially aiding the transition between sessile and motile lifestyles during infection.

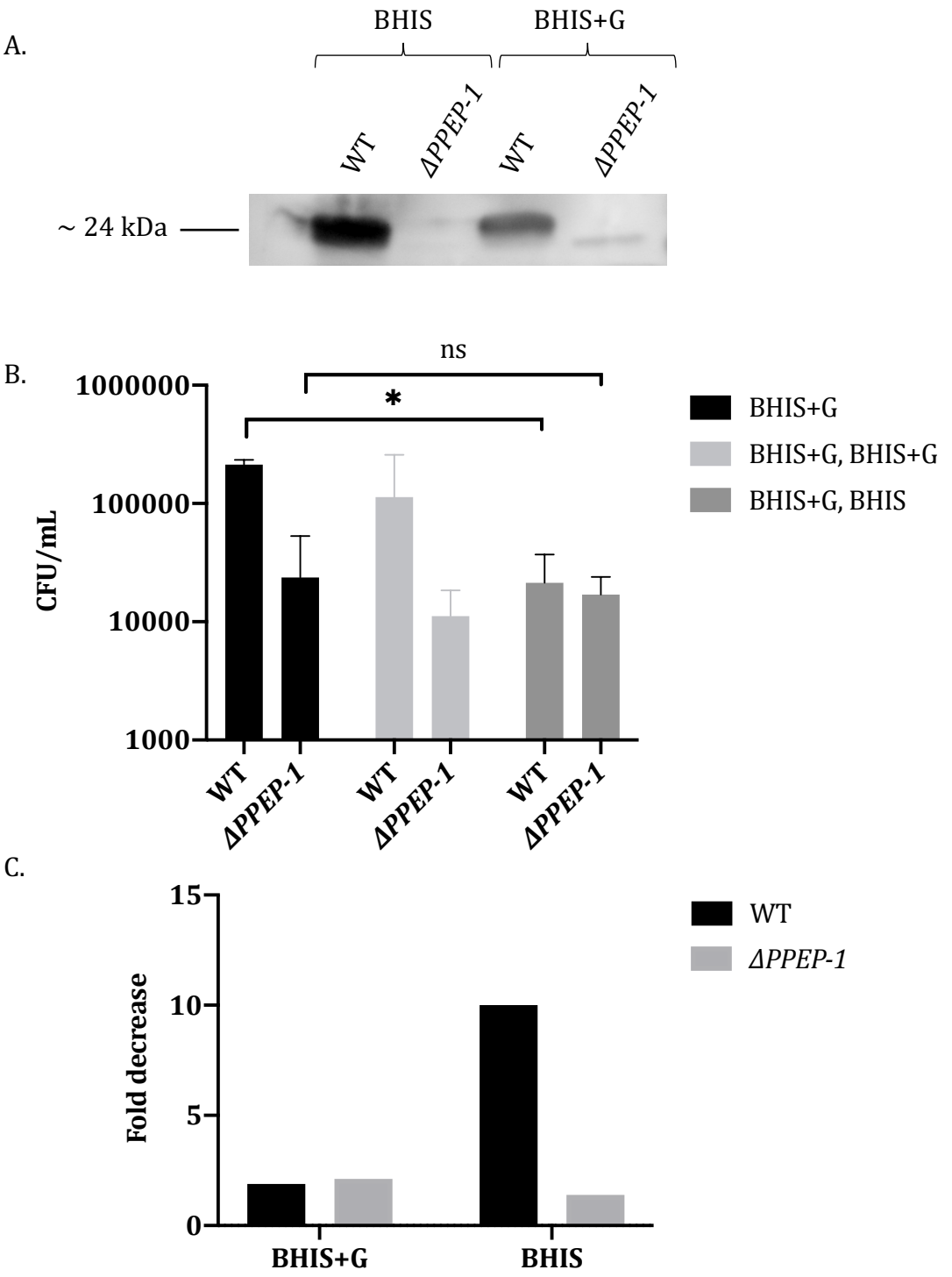


Figure 29. PPEP-1 may have a role in the release of collagen-bound *C. difficile*.

A. Western blot of TCA precipitated WT or *ΔPPEP-1* culture supernatants, grown in BHIS or BHIS+G. B. CFU counts from bacteria incubated in BHIS+G only, BHIS+G then BHIS+G and BHIS+G then BHIS (*p = 0.01 determined with a two-tailed paired t-test) (n=3). C. The fold decrease between the number of cells before and after the subsequent addition of BHIS+G or BHIS media. Bars are representative of the mean +/- standard deviation.

4.3 Discussion

C. difficile colonisation of the gut epithelium is crucial for the establishment of infection. However, the specific alterations in gene expression *C. difficile* employs to establish epithelial cell attachment have been understudied. There have been several previous studies investigating bacterial transcriptomic responses during *C. difficile* infection. Scaria *et al.* (2011) performed a microarray-based transcriptomic analysis to investigate *C. difficile* responses to infection in a pig ligated-loop model. Many genes involved in *C. difficile* virulence were identified as upregulated at the early stages of infection (4 h post infection), including the toxin A gene, *tcdA*, *codY*, CD2830 (*PPEP-1*), CD2592 (fibronectin binding protein), CD2793 (*slpA*), CD1546, CD1208 (hemolysins), and genes involved in sporulation cascade (Scaria *et al.*, 2011). However, *C. difficile* responses to infection of human cells were not investigated. Janvilisri, Scaria and Chang, (2015) investigated the host and bacterial transcriptional responses during *C. difficile* infection of Caco-2 cells for up to 120 minutes under anaerobic conditions using a microarray-based approach. Differential expression of 274 *C. difficile* genes was observed, including many involved in metabolic pathways, such as nucleic acid metabolism, transcription as well as protein synthesis and modification. The gene expression profile resolved in this analysis suggested that a bacterial stress response occurs during infection of Caco-2 cells. However, the timepoints used in this analysis were relatively short and may not be representative of *C. difficile* infection *in vivo* (Janvilisri, Scaria and Chang, 2015). These studies relied on probe-based methods to perform transcriptional profiling which have several flaws, such as reduced sensitivity, risks of cross-hybridisation and the inclusion of only known, coding transcripts. RNA-seq has several advantages over microarray-based

techniques such as providing an unbiased insight into all transcripts, as described in Introduction section 1.5.4. An RNA-seq approach has been previously applied to *C. difficile* VPI 10463 RNA isolated from infected mice ceca at early and late stages of infection to investigate the nutritional needs of *C. difficile* during infection. Fatty acid metabolism, butyrate production and iron uptake pathways were among the top upregulated metabolic processes, highlighting some of the nutritional requirements of *C. difficile* during infection of mice. However, genes required for *C. difficile* colonisation and virulence were not investigated and there are physiological differences in *C. difficile* infection of mice and humans (Best, Freeman and Wilcox, 2012; Fletcher *et al.*, 2018). There remain large gaps in our current knowledge regarding the regulation of genes involved in *C. difficile* colonisation and virulence during infection.

The use of *in vitro* human gut models has contributed towards a better understanding of the mechanisms used by other enteric pathogens to cause disease. For example, *Shigella flexneri* infection of 2D colonoid monolayers plated on permeable membrane scaffolds demonstrated that basolateral infection was significantly more efficient than apical infection, and stimulated an increase in MUC2 production in the host cells (Ranganathan *et al.*, 2019). *In vitro* colonoid monolayer models of EHEC infection have also expanded our knowledge of many features of EHEC pathogenesis, such as the loss of microvilli. In *et al*. demonstrated that the serine protease autotransporter of Enterobacteriaceae (SPATE), EspP, specifically targeted protocadherin 24, an important component of intermicrovillar bridges, as well as induced the redistribution of the tight junction protein occludin, resulting in brush border disruption (In *et al.*, 2016).

These studies demonstrated how the use of *in vitro* models for the human gut can resolve intricate molecular processes which occur during enteric infections, and therefore may also provide insights into the pathogenesis of *C. difficile*.

Single gene expression profiles: in the dual RNA-seq analysis, the gene expression profiles of bacterial RNA from infected samples were compared to that from a bacterial inoculum grown to log-phase in the same media. The use of multiple timepoints after infection made it possible to track the changes in gene expression over the course of infection using single gene expression profiles. There were numerous significantly differentially expressed genes identified which have been previously reported to have a role in *C. difficile* virulence. Interestingly, the virulence-associated genes *codY, PPEP-1* and *fliC* were all downregulated during infection, with statistically significant differences in at least one timepoint (Figure 20).

CodY is a global transcriptional regulator which modulates numerous cellular processes, including the repression of toxin production and sporulation through the suppression of *tcdR* and *spo0A* (Dineen *et al.*, 2007; Dineen, McBride and Sonenshein, 2010; Nawrocki *et al.*, 2016). The activity of CodY is linked to the availability of nutrients, where intracellular branched-chain amino acids and GTP bind to CodY, inhibiting its promotor-binding activity (Dineen, McBride and Sonenshein, 2010). The downregulation of *codY* observed during infection may result in reduced suppression of *tcdR* and *spo0A*, leading to an increase in toxin production and spore formation, two processes which are crucial for *C. difficile* pathogenesis (Dineen *et al.*, 2007; Nawrocki *et al.*, 2016).

FliC is a structural monomer and is a major component of the flagellar filament. *fliC* deletion mutants lack flagella and are non-motile, demonstrating the vital role for *fliC* in flagellar formation and bacterial motility (Dingle, Mulvey and Armstrong, 2011). The role of FliC in *C. difficile* adhesion to intestinal epithelial cells is likely to be strain dependent, where in strain 630 deletion of *fliC* increased adherence to Caco-2 cells, while in R20291 it reduced it (Tasteyre *et al.*, 2001; Dingle, Mulvey and Armstrong, 2011). A *fliC* mutant also exhibited increased toxin production in a hamster model of infection, suggesting it may also play a role in toxin-mediated disease and bacterial motility may not necessarily be required for virulence (Dingle, Mulvey and Armstrong, 2011). *C. difficile* patients have also been reported to exhibit serum antibody responses to FliC, and immunisation with recombinant FliC offered partial protection against *C. difficile* infection and death in hamsters (Péchiné *et al.*, 2005; Ghose *et al.*, 2016). These studies further suggest that FliC may have an important role in *C. difficile* virulence. The downregulation of *fliC* at 12 h post infection was consistent with what was reported by Janoir *et al.* (2013) who also observed a strong downregulation of *fliC* during transcriptomic profiling of *C. difficile* infection in a mouse model (Janoir *et al.*, 2013). Therefore, downregulation of *fliC* may have an important role in *C. difficile* infection, as it exhibits adhesive properties and may play a role in virulence.

The majority of the genes which encode the Opp and App oligopeptide transporters were significantly downregulated in this analysis. This downregulation in oligopeptide transporters may result in a reduction in the

intracellular concentration of oligopeptides. Disruption of the Opp and App transporters has been reported to increase sporulation frequency and virulence in a hamster model, through the activities of the CcpA-dependent transcriptional regulator SinR, a repressor of *spo0A* (Edwards, Nawrocki and McBride, 2014). It has also been established that toxin production in *C. difficile* is upregulated in response to nutrient limitation, including amino acid starvation (Karlsson, Burman and Akerlund, 1999; Dineen *et al.*, 2007; Nawrocki *et al.*, 2016). Therefore, the downregulation in oligopeptide transporters may be linked to an increase in toxin production and spore formation, two processes vital for *C. difficile* pathogenesis.

Interestingly, *slpA* was highly downregulated at all timepoints after infection, and *cwp84* was significantly downregulated at 6 h post infection. These proteins are vital for the formation of the S-layer which has been reported to play a major role in *C. difficile* adhesion to host tissues (Takeoka *et al.*, 1991; Calabi *et al.*, 2002). The transcriptomic analysis by Janoir *et al.* of *C. difficile* 630 in a germfree mouse model reported *slpA* to be upregulated during the early stages of infection, and downregulated in the later stages (Janoir *et al.*, 2013). This suggests that *C. difficile* may modulate its cell surface at different points over the course of infection. Chemical removal or inhibition of SLPs abolished *C. difficile* binding to mouse 929 and human HeLa cells (Takeoka *et al.*, 1991). *C. difficile* SLPs also bind to Hep-2 cells, Vero cells, human gastrointestinal tissues and immobilised ECM components, further demonstrating their important role in colonisation (Calabi *et al.*, 2002). As well as SLPs, Cwp84 has also been reported to cleave components of the ECM and play a role in *C. difficile* biofilm formation *in vitro* (Janoir *et al.*,

2007; Dawson *et al.*, 2012; Dapa *et al.*, 2013; Pantaléon *et al.*, 2015). Therefore, the downregulation of the S-layer genes may result in an alteration in the composition of the S-layer and may regulate *C. difficile* colonisation of intestinal epithelial cells. The coordination of cell-surface adhesins at different timepoints after infection may have a role in the pathogenesis of the disease, and may modulate the exposure of specific adhesins, such as ECM-binding proteins. In this analysis, numerous *C. difficile* cell wall proteins were also modulated during infection. Cwp66 has been reported to play a role in *C. difficile* adhesion to host tissues, where anti-Cwp66 antibodies partially inhibited heat-shocked *C. difficile* adherence to Vero cells. *cwp66* was significantly downregulated in this analysis, as well as *cwp13*, *cwp14*, *cwp17*, *cwp18*, *cwp19* and *cwp20* which do not have defined roles *C. difficile* pathogenesis yet. However, there were also several cell wall proteins which were highly upregulated at all timepoints after infection, including CDR20291_2686, CDR20291_0184 and *cwp10*. There is currently little knowledge regarding the functions of these proteins, but their high levels of upregulation in this analysis suggests they may have unprecedented roles in *C. difficile* pathogenesis. CDR20291_2686 has a putative fibronectin-binding domain, suggesting that it may have a role in *C. difficile* adhesion to host tissues during infection. CDR20291_0184 is a putative cell wall hydrolase and *cwp10* is a paralog of SlpA. Therefore, these genes may encode potential colonisation factors for *C. difficile*.

There were many genes involved in the uptake of iron which were modulated during *C. difficile* infection. Fur is a transcriptional regulator which controls gene expression in response to iron levels by binding to target promotors in the

presence of iron (Ho and Ellermeier, 2015). Fur controls the expression of many molecular processes which are involved in iron acquisition and metabolism, including the repression of *feo1* (ferrous iron transport operon), CDR20291_0516 (a putative cation transporting ATPase) and CDR20291_1639 (putative ferrous iron transport protein A). In this analysis, Fur was highly upregulated at 6 h, 12 h and 24 h post infection. Consistent with this, the *feo1* operon, CDR20291_0516 and CDR20291_1639 were significantly downregulated during infection, likely due to increased suppression by Fur. Fur has also been reported to induce the expression of the CDR20291_3444 ferredoxin gene, which was significantly upregulated at 3 h and 6 h post infection, possibly due to increased levels of Fur. However, Fur has also been reported to repress the expression of *fldX* (flavodoxin), *fhuD* (iron complex transport system substrate-binding protein) and CDR20291_1545 (putative iron compound ABC transporter) which were significantly upregulated during infection in this analysis (Figure 23). The role of Fur in the regulation of iron acquisition systems has been established in *C. difficile* strain 630 but has not been investigated in R20291, and it would be possible that regulation of gene expression by Fur may differ between strains (Ho and Ellermeier, 2015).

The toxin genes were not significantly differentially expressed at any timepoint in this analysis. This is consistent with the microarray-based transcriptomic analysis by Janvilisri *et al.* (2015) who also did not observe any differential expression of the toxin genes during *C. difficile* infection of Caco-2 cells. Small concentrations of TcdA have been previously detected in the supernatants of *C. difficile*-infected cells incubated in the VDC system at 3 h, 6 h and 24 h post

infection (Anonye *et al.*, 2019). In this experiment, the infected intestinal epithelial layers were washed with PBS to remove unadhered bacteria prior to RNA extraction. Therefore, only the RNA from adhered bacteria was sequenced and included in the analysis. Toxin production is induced during the motility associated c-di-GMP ribotype, which is activated during low intracellular concentrations of c-di-GMP, as well as co-expressed with flagella genes following regulation with the *cwpV* genetic switch (McKee *et al.*, 2013; Anjuwon-Foster and Tamayo, 2017). The induction of these motility associated genes during toxin gene expression may result in the bacterial cells becoming unadhered from the intestinal epithelial cell layer, resulting in these bacteria being excluded from the analysis. Furthermore, since there were only small concentrations of TcdA detected in the VDC system previously at these timepoints (~1 ng/mL), it is also possible that the expression of toxin genes was too low to be detected by sequencing.

Pathway enrichment analysis: to investigate *C. difficile* pathways which were modulated during infection, a pathway enrichment analysis was conducted. Many pathways were identified as up- or downregulated during *C. difficile* infection, including the metabolism and biosynthesis of a range of substrates (Figure 24). The purine metabolism pathway was significantly upregulated at 6 h, 12 h and 24 h post infection. This is consistent with what was reported by Fletcher *et al.* who found purine metabolism to be the most significantly upregulated pathway during *C. difficile* infection of a murine model (Fletcher *et al.*, 2018). The upregulation of the glycolysis/gluconeogenesis pathway identified in this study was also consistent with what was found in

transcriptomic analyses by both Scaria *et al.* (2011) and Fletcher *et al.* (2018). The significant upregulation of biosynthesis of secondary metabolites in this analysis, was also consistent with what was found by Scaria *et al.* who performed a transcriptomic analysis of *C. difficile* infection of a pig-loop model. However, there were several differences in the results of these studies, such as pyruvate metabolism, propanoate metabolism, one carbon pool by folate and glyoxylate and dicarboxylate metabolism pathways which were upregulated in the study by Fletcher *et al.* but downregulated in this analysis. There were major differences in the experimental procedures of this analysis and the study by Fletcher *et al.* which may have caused the differences in the regulation of metabolic pathways during infection. Firstly, the study by Fletcher *et al.* was conducted in a mouse model which has several physiological differences to the disease in humans (Best, Freeman and Wilcox, 2012). The Fletcher *et al.* study also compared the expression profiles of early (12 h) and late stages of infection (24 h and 30 h) rather than to a bacterial culture. Therefore, genes and pathways which were modulated at all stages of infection could not be identified. The Fletcher *et al.* study was also conducted in a mouse model of infection, so the bacteria were likely to need to adapt to the inflammatory response evoked by the infected mice, which is not possible in the VDC system. A more recent paper by Fletcher *et al.* (2021) investigated the effect of toxin-mediated inflammation on the gene expression profile of *C. difficile*. The data from this study suggested that carbohydrate metabolism and branched-chain amino acid biosynthesis were enriched in an inflamed gut (Fletcher *et al.*, 2021). These pathways were not identified as differentially expressed in this analysis, potentially because there was no inflammatory response in the VDC system to stimulate their expression.

Interestingly, there was an upregulation in the metabolism of purine and pyrimidine ribonucleotides. As well as being the building blocks for DNA, nucleotides have many other fundamental roles in the molecular processes of all cells. The purinergic signalling molecule, ATP, is vital for cellular processes and acts as a currency for transferring and storing energy within cells. ATP is released from both eukaryotic and prokaryotic cells and has been described as an inter-kingdom signalling molecule (Spari and Beldi, 2020). Follicular T helper cells (Tfh) in the Peyer's patches of the gut have been reported to be modulated by extracellular ATP (eATP) released from the microbiota (Proietti *et al.*, 2019). Bacteria secrete ATP themselves but can also stimulate cells to secrete ATP which can initiate an inflammatory response (Savio *et al.*, 2018; Spari and Beldi, 2020). Bacterial ATP release can act as a virulence factor and can contribute towards the detrimental effects of the inflammatory response in a range of infections, including *Staphylococcus epidermis* (Robertson *et al.*, 2010; Inami, Kiyono and Kurashima, 2018; Liu *et al.*, 2018; Savio *et al.*, 2018). Biofilm formation in *E. coli* was also increased by extracellular pyrimidine nucleotides, through the modulation of the modified pyrimidine ribonucleotide, c-di-GMP (Garavaglia, Rossi and Landini, 2012). This molecule modulates biofilm formation in a range of bacterial species, including *C. difficile* (Purcell *et al.*, 2012; Dawson *et al.*, 2021). As well as having a role in the regulation of biofilm formation, this molecule also regulates the expression of *C. difficile* virulence-associated genes, including toxin production (Garavaglia, Rossi and Landini, 2012; Purcell *et al.*, 2012; McKee *et al.*, 2013; Bordeleau *et al.*, 2015). Foulke-Abel *et al.* also demonstrated that exposure to enterotoxigenic *Escherichia coli* (ETEC) enterotoxins induced the

secretion of cyclic nucleotides in an *in vitro* human gut model. Heat-labile enterotoxin exposure induced significant apical secretion of cyclic AMP, while addition of heat-stable enterotoxin significantly increased basolateral secretion of cyclic GMP (Foulke-Abel *et al.*, 2020). Therefore, the upregulation of these pathways may have a role in *C. difficile* virulence and may contribute towards the detrimental effects of the inflammatory response, as well as regulating *C. difficile* biofilm formation and toxin production through the modulation of intracellular c-di-GMP.

A role for PPEP-1 in *C. difficile* colonisation: PPEP-1 is a zinc metalloprotease which proteolytically cleaves the collagen-binding proteins CD2831 and CD3246, removing them from the *C. difficile* cell-surface (Hensbergen *et al.*, 2014). PPEP-1 has a unique cleavage site and has also been reported to cleave host proteins, including fibronectin and fibrinogen (Cafardi *et al.*, 2013). A PPEP-1 deletion mutant has been previously reported to exhibit an attenuation in virulence in a hamster model, suggesting it may play a role in *C. difficile* virulence (Hensbergen *et al.*, 2015). In this dual RNA-seq analysis, PPEP-1 was significantly downregulated at 3 h, 12 h and 24 h post infection. This downregulation of PPEP-1 during infection is consistent with what was observed by Fletcher *et al.*, (2018), who also found a downregulation in PPEP-1 during RNA-seq profiling of *C. difficile* infection in a mouse model of infection. However, in the analysis by Scaria *et al*. they found an upregulation in PPEP-1 at 4 h post *C. difficile* 630 infection in a pig ligated loop model (Scaria *et al.*, 2011). In this analysis there was no differential expression of CD2831 and CD3246, the substrates of PPEP-1,

suggesting that the regulation of these proteins may be dependent on post-translational cleavage by PPEP-1.

PPEP-1-mediated cleavage of CD2831 and CD3246 may be important for *C. difficile* biofilm formation, where deletion of CD2831 significantly reduced biofilm biomass, while overexpression of CD2831 and CD3246 in *ΔPPEP-1* significantly increased biofilm biomass (Dawson *et al.*, 2021). WT and *ΔPPEP-1* biofilm biomasses were quantified with CV staining, and it was found that there was no significant difference in biofilm formation between WT and *ΔPPEP-1*, consistent with a recent study by Dawson *et al*. (2021). This result may be due to the bacterial cultures being grown into biofilms in BHIS+G, as this media has been previously reported to stimulate biofilm formation, suggesting that BHIS+G media may trigger an adhesion-associated state which is stimulated in the presence of high intracellular concentrations of c-di-GMP (Dapa *et al.*, 2013). However, PPEP-1 is expressed during the motility-associated state, when intracellular concentrations of c-di-GMP are low (Purcell *et al.*, 2012; Bordeleau *et al.*, 2015; Hensbergen *et al.*, 2015; Corver *et al.*, 2017). Therefore, there may not be any observed significant difference in biofilm formation in BHIS+G media because PPEP-1 may not have been highly expressed in the WT. The *ΔPPEP-1*::pRPF185-PPEP-1 complemented strain formed significantly more biofilms than the WT. This is possibly due to the incubation with antibiotics and ATc, which may elicit a stress response in the bacteria, stimulating the formation of biofilms.

Bacterial autoaggregation may have a role in the early stages of biofilm formation and may contribute towards protection against host immune responses and antibiotics during infection (Trunk, S. Khalil and C. Leo, 2018). To determine whether PPEP-1 has a role in *C. difficile* autoaggregation, an aggregation assay was conducted which measured the bacterial autoaggregation over the logarithmic growth phase of cultures grown in BHIS+G. There was no significant difference in the growth rates between any of the tested strains in the growth curve (Figure 26A). However, there was a significantly higher percentage of aggregated cells in *ΔPPEP-1* compared to the WT at the 4 h timepoint (Figure 26B). There was also no significant difference in the percentage of aggregated cells in *ΔPPEP-1*::pRPF185 compared to the WT and complemented strains, which I would expect to behave in a similar way to *ΔPPEP-1*. The expression of the genes on the plasmid (e.g., antibiotic resistance genes) may have altered the cellular aggregation of *C. difficile*. Therefore, further investigations may be necessary to determine the role of PPEP-1 in autoaggregation.

Since PPEP-1 cleaves cell-surface associated collagen-binding proteins, I hypothesised that it may have a role in *C. difficile* colonisation of intestinal epithelial cells. It was found that *ΔPPEP-1* had a significantly higher number of bacterial cells attached to the epithelial cell layers at 6 h post infection (Figure 27). Confocal microscopy analysis of *C. difficile*-infected epithelial cell layers demonstrated that there were significantly more *ΔPPEP-1* cells adhered to cells at 6 h and 24 h post infection compared to the WT (Figure 28). There may be a significant increase in *ΔPPEP-1* adhered to host cells in the microscopy at 24 h, but not in the CFU counts because *C. difficile* may have formed microcolonies or

small biofilms with dead bacterial cells encased in the matrix which could be detected with microscopy, but not CFU counts. This data suggests that PPEP-1 may play a negative role in *C. difficile* colonisation. In the dual RNA-seq analysis, there was a significant downregulation of *PPEP-1* expression at all timepoints except at 6 h post infection. Therefore, the decrease in *PPEP-1* expression at 3 h and 24 h post infection may explain why there was no observed significant difference in bacterial adhesion to host cells between WT and *ΔPPEP-1* at these timepoints. However, at 6 h post infection there was no significant differential expression of *PPEP-1* when comparing a bacterial culture to *C. difficile* from infected cells. The significant increase in the amount of cellular adhesion of *ΔPPEP-1* at 6 h post infection may be the result of an increase in *PPEP-1* expression in comparison to the other timepoints. This may have resulted in increased cleavage of the collagen-binding adhesins in the WT, but not in *ΔPPEP-1*, resulting in differences in cellular adhesion at 6 h post infection. It is also possible that the bacteria expressing PPEP-1 were in a motile state and may not have been included in the RNA-seq analysis since the epithelial cell layers were washed prior to RNA extraction to include only adhered bacteria, resulting in a downregulation in *PPEP-1* in the RNA-seq data.

The bacterial substrates of PPEP-1 are expressed in the adhesion-associated c-di-GMP ribotype, while PPEP-1 is expressed in the motility-associated ribotype (Purcell *et al.*, 2012; Hensbergen *et al.*, 2015; Peltier *et al.*, 2015; Dawson *et al.*, 2021). Therefore, I hypothesised that PPEP-1 may play a role in the shift between the two lifestyles, where PPEP-1 cleaves cell-surface adhesins to release the bacteria from collagen and facilitate bacterial motility. To examine this, a

collagen-binding assay was conducted before and after a phase shift was induced by the addition of different media. To examine the expression levels of PPEP-1 in BHIS or BHIS+G media, a western blot was conducted with supernatants from WT and *ΔPPEP-1* cultures grown in BHIS or BHIS+G. It was found that PPEP-1 is secreted in both media, but more highly in BHIS than BHIS+G (Figure 29A). However, there was no loading control included in this western blot to ensure that equal amounts of protein were loaded into the polyacrylamide gel so the intensities of the bands could be accurately compared. If there was sufficient time remaining to repeat this, the protein concentrations could have been visualised and compared by performing a ponceau stain of the PVDF membrane. Surprisingly, in the collagen binding assay, there were a higher number of bacteria adhered to collagen in the WT than *ΔPPEP-1* in BHIS+G media (Figure 29B). This is inconsistent with what has previously been reported in a collagen-binding assay conducted in PBS, where *ΔPPEP-1* exhibited a higher adhesion to collagen compared to the WT (Hensbergen *et al.*, 2015). This suggests that PPEP-1 may indirectly modulate the levels of other adhesion-associated proteins regulated by the c-di-GMP riboswitch. However, after subsequent incubation in BHIS media there was a significant decrease in the number of WT bacteria remaining adhered to the collagen, while no significant change was observed in the numbers of *ΔPPEP-1* (Figure 29C). This suggests that PPEP-1-mediated cleavage of collagen binding proteins may facilitate the release of *C. difficile* from collagen during the transition between sessile and motile lifestyles. During infection, this may have a role in *C. difficile* virulence, as PPEP-1 may be required to facilitate bacterial motility and dispersal of the infection in the gastrointestinal tract.

To summarise this chapter, the dual RNA-seq analysis successfully identified a range of bacterial genes and pathways modulated during *C. difficile* infection. A potential role for PPEP-1 in the colonisation of intestinal epithelial cells was identified, as well as a potential role for this secreted protease in the transition between sessile and motile lifestyles during infection.

Chapter 5: Host responses to *C. difficile* infection

5.1 Introduction

Host-*C. difficile* interactions are initiated when a bacterial spore or vegetative bacterial cell colonise the gut, a crucial step for disease progression. Infection initiates a cascade of molecular processes which result in alterations of gene expression, leading to adaptations from both host and pathogen to facilitate an infection. *C. difficile* colonisation of the gut epithelium is crucial for the establishment of infection, yet the underlying molecular mechanisms remain poorly understood. Specifically, little is known about the interactions between *C. difficile* and the colonic mucus layer. *C. difficile* carries many ECM-binding proteins on its cell surface which bind to ECM components *in vitro* (Wann, Gurusiddappa and Höök, 2000; Hennequin *et al.*, 2003; Barketi-Klai *et al.*, 2011; Tulli *et al.*, 2013; Arato *et al.*, 2019). However, *C. difficile* interactions with the gut tissue ECM also remain largely unexplored.

Several cellular components have been reported to have a role in the host immune response against *C. difficile* infection. Firstly, *C. difficile* produces PAMPs which are recognised by PRRs on the host cell surface, triggering intracellular signalling cascades and ultimately activates an innate immune response (Jarchum *et al.*, 2011; Yoshino *et al.*, 2013). The innate immune response stimulates the secretion of pro-inflammatory cytokines which attract immune cells to the infected area (Czepiel *et al.*, 2014; Yu *et al.*, 2017). Several colonic epithelial receptors for *C. difficile* toxins have also been identified, where the

binding of a toxin to these receptors results in toxin internalisation and disrupts the usual signalling processes mediated by that receptor (Yuan *et al.*, 2015; Tao *et al.*, 2016; Chen *et al.*, 2018). Toxins also trigger numerous host immune response signalling pathways, including p38 MAPK, TNFα, NF-$\kappa\beta$, and TLR signalling, which stimulate the recruitment of immune cells (He *et al.*, 2002; Ishida *et al.*, 2004; Kim *et al.*, 2005; Arancibia *et al.*, 2007; Sun *et al.*, 2009; Jarchum *et al.*, 2011; Ryan *et al.*, 2011; McDermott *et al.*, 2016; Cowardin *et al.*, 2017). Therefore, these host genes and pathways may be transcriptionally modulated during *C. difficile* infection.

Previous studies have used probe-based methods to investigate host responses to *C. difficile* infection. D'Auria *et al.* (2013) investigated the host gene expression alterations in response to purified *C. difficile* toxins in a mouse model. Frisbee *et al.* (2019) investigated the host response to vegetative *C. difficile* R20291 in a mouse model. Janvilisri *et al.* (2015) investigated transcriptomic profiles for host-pathogen interactions during *C. difficile* infection of Caco-2 cells in anaerobic conditions. Overall, these studies have some shed light on a few of the ways host cells adapt their gene expression in response to *C. difficile* infection, but there are still large gaps in our current knowledge. A better understanding of the host-*C. difficile* interface is vital for the development of novel interventions against the disease. RNA-seq is a powerful tool which overcomes many of the limitations associated with probe-based methods. Dual RNA-seq has emerged as an effective strategy to understand the cellular responses concurrently from both host and pathogen.

In this part of the study, I will present the analysis of host cell transcriptional responses to *C. difficile* infection from the dual RNA-seq analysis. The inclusion of several timepoints over the course of infection allowed the temporal changes in gene expression to be tracked. A pathway enrichment analysis was conducted to investigate pathways which were modulated during *C. difficile* infection. The mucosal immunity protein, MUC13, and ribonucleotide signalling pathways were found to be significantly upregulated during *C. difficile* infection and the importance of these processes in the pathogenesis of disease were investigated.

5.2 Results

5.2.1 Single gene expression profiles

As discussed in chapter 3, many differentially expressed human genes were identified over the course of *C. difficile* infection (complete lists can be found in Appendix tables 11-14). Since total RNA was extracted from *C. difficile*-infected intestinal epithelial cells over several timepoints, single gene expression profiles were generated to visualise the alterations in gene expression over the course of infection. The transcriptomic alterations of selected genes over time were examined by plotting the $\log_2$(fold change) at each timepoint. The single gene expression profiles have been grouped into genes which are involved in the immune response (Figure 30) and other interesting genes which may have a role in the pathogenesis of *C. difficile* infection (Figure 31).

5.2.1.1 Selected immune response genes differentially expressed during infection

5.2.1.1.1 Tumor necrosis factor receptor superfamily members

Tumor necrosis factor receptor superfamily member 12A (TNFRSF12A) was significantly downregulated at 3 h, 12 h and 24 h post infection (Figure 30). TNFRSF12A, also known as Fibroblast growth factor-inducible immediate-early response protein 14 (Fn14) overexpression has been linked to several types of cancer and has been correlated with poor surgical outcome (Li *et al.*, 2013). One of the most significantly differentially expressed genes at 3 h post infection was tumor necrosis factor receptor superfamily member 11B (TNFRSF11B) which was downregulated with a $\log_2$(fold change) of -2.45 and an adjusted *p*-value of 1.12E-14 (Figure 30). This gene encodes osteoprotegerin, a receptor which binds to TNF-related apoptosis-inducing ligand (TRAIL) and prevents it from activating TNF-family death receptors, thereby inhibiting TRAIL-induced apoptosis (Emery *et al.*, 1998; Baud'huin *et al.*, 2013). This secreted protein plays a role in multiple cellular processes, including the regulation of bone density, cell proliferation and apoptosis (Baud'huin *et al.*, 2013).

5.2.1.1.2 NF-κB1

The gene for Nuclear Factor Kappa B Subunit 1 (NF-κB1) was significantly downregulated at 3 h post infection (Figure 30). This gene is involved in the inhibition of the NF-κB signalling pathway, and is also a regulator for MAPK signalling (Beinke and Ley, 2004). The NF-κB and MAPK signalling pathways transcriptionally regulate immune responses and have been previously reported

to be activated during the inflammatory immune response to *C. difficile* infection (Jefferson, Smith and David, 1999).

5.2.1.1.3 Interleukin-8

Surprisingly, the gene for IL-8 (*CXCL8*) was significantly downregulated at 3 h, 6 h and 12 h post infection. IL-8 is highly secreted during *C. difficile* infection and is a chemoattractant for neutrophils (Linevsky *et al.*, 1997a; Hoffmann *et al.*, 2002). IL-8 been strongly associated with *C. difficile* disease prognosis, where fecal and serum levels have been correlated with disease severity and polymorphisms in this gene have been linked to an increased susceptibility to *C. difficile* infection (Linevsky *et al.*, 1997a; He *et al.*, 2002; Jiang *et al.*, 2007; Czepiel *et al.*, 2014) (Figure 30).

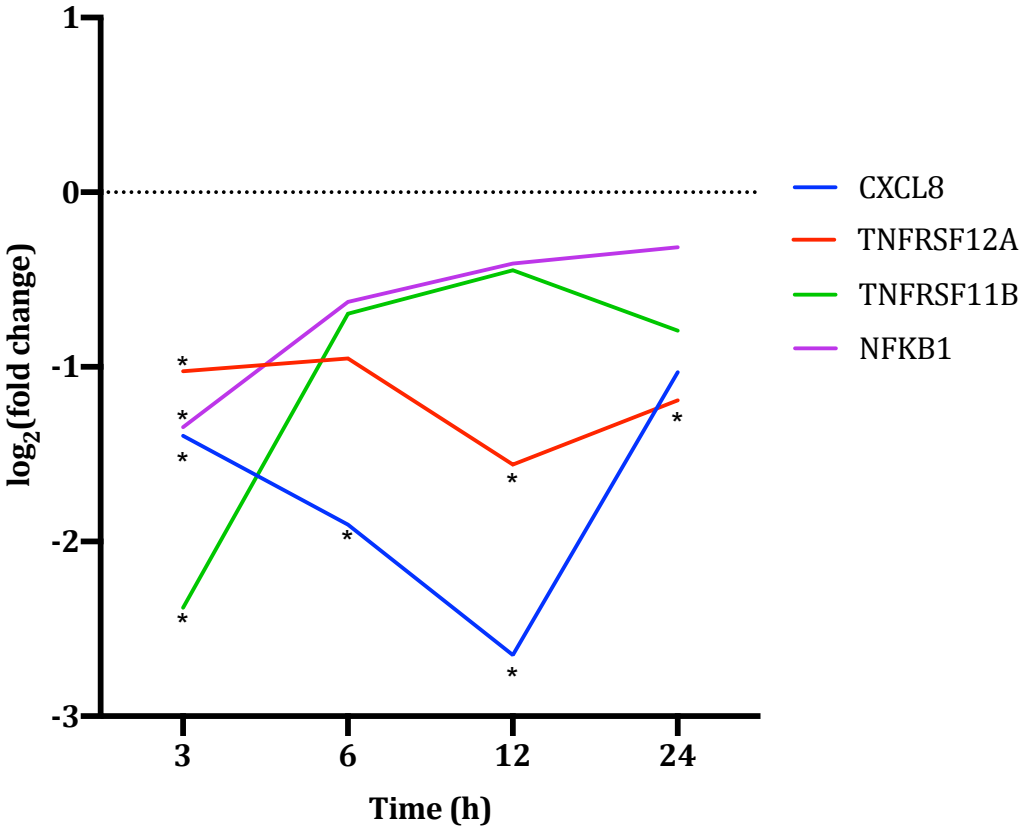


Figure 30. Host single gene expression profiles of selected immune response genes differentially expressed during *in vitro C. difficile* infection.

Log_2(fold change) of selected genes calculated between the expression levels of uninfected controls and infected samples at each timepoint. All selected genes were significantly differentially in at least one timepoint, indicated with an asterisk.

5.2.1.2 Interesting genes differentially expressed in response to *C. difficile* infection

5.2.1.2.1 Fibrinogen

Among the most highly downregulated genes at 3 h, 6 h and 12 h post infection were the fibrinogen alpha, beta and gamma chain genes (*FGA*, *FGB* and *FGG*). However, there was no differential expression observed at 24 h post infection

(Figure 31). The three fibrinogen chain molecules combine to form the fibrinogen glycoprotein complex which plays an important role in blood clotting, wound healing, fibrinolysis, ECM interactions, the inflammatory response, and neoplasia (Mosesson, 2012).

5.2.1.2.2 FZD4 and LRP4

The frizzled-4 protein (*FZD4*) was significantly downregulated at 3 h post infection (log_2(fold change) of 1.57 and adjusted *p*-value of 0.00053) (Figure 31). This protein is a receptor involved in the Wnt/β-catenin canonical signalling pathway, a conserved molecular system which is crucial for cellular processes such as cell fate determination, cell migration and cell polarity (Komiya and Habas, 2008). Low-density lipoprotein receptor-related protein 4 (*LRP4*) was also differentially expressed during *C. difficile* infection, being significantly upregulated at 3 h and 12 h post infection (Figure 31). This protein is involved in the inhibition of Wnt signalling (Tao *et al.*, 2019). LRPs and Frizzled proteins are colonic receptors for *C. difficile* TcdA and TcdB, so their differential expression may have a role in the pathogenesis of *C. difficile* infection (Tao *et al.*, 2016, 2019; Chen *et al.*, 2018).

5.2.1.2.3 APOA4

One of the most upregulated genes at 12 h and 24 h post infection was apolipoprotein A-IV (*APOA4*), an apolipoprotein primarily produced in the small intestine (Figure 31) (Ostos *et al.*, 2001). Apolipoproteins are proteins which bind to lipids to form lipoproteins, and ApoA-IV is involved in multiple

physiological processes, including lipid absorption, metabolism, platelet aggregation, thrombosis and glucose homeostasis (Vowinkel *et al.*, 2004).

5.2.1.2.4 MUC13

The gene for MUC13 was upregulated (log_2(fold change) >1) at 6 h, 12 h and 24 h post infection, but this upregulation was only statistically significantly at 24 h post infection (Figure 31). MUC13 is a transmembrane mucin glycoprotein which is highly expressed on the surface of mucosal epithelial cells in the small and large intestines (McGuckin *et al.*, 2011). This mucin also has a role in the regulation of inflammatory cytokines, including IL-8, a chemoattractant of neutrophils with an important role in the immune response to *C. difficile* infection (Linevsky *et al.*, 1997b; He *et al.*, 2002; Jiang *et al.*, 2007; Sheng *et al.*, 2011).

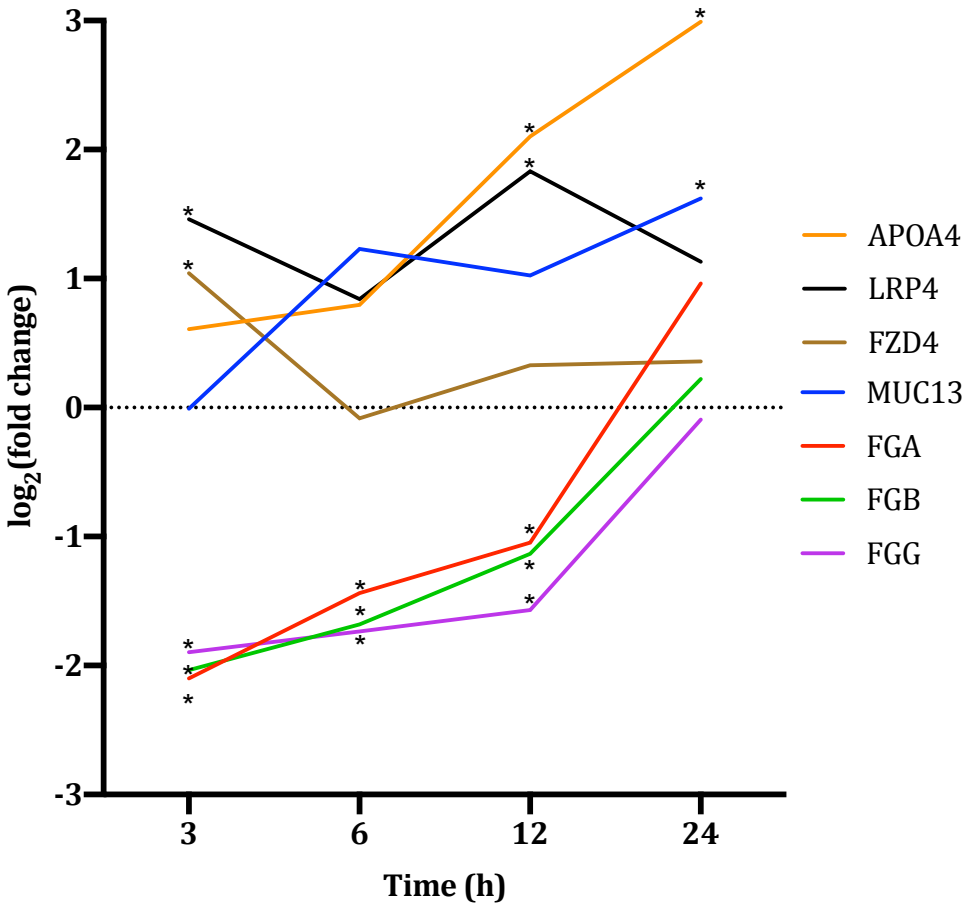


Figure 31. Host single gene expression profiles of interesting genes differentially expressed during *in vitro C. difficile* infection.

Log_2(fold change) of selected genes calculated between the expression levels of uninfected controls and infected samples at each timepoint. All selected genes were significantly differentially in at least one timepoint, indicated with an asterisk.

5.2.2 Transcript quantification using RT-qPCR

RNA-seq analysis resolved the transcriptomic profiles of intestinal epithelial cells during *C. difficile* infection. To confirm that the transcriptomic profiling observed in the RNA-seq analysis reflected the transcript quantities in the RNA samples, RT-qPCR was used to quantitate the expression levels of selected transcripts. The quantities of each target transcript from infected samples and uninfected

controls were normalised to the levels of the housekeeping gene, GAPDH (ΔCt). The differences between the ΔCt values of the uninfected controls and infected samples were calculated for each timepoint, to obtain a value defined as the ΔΔCt. The fold changes during infection were then calculated for each gene of interest using the $2^{-\Delta\Delta Ct}$ (see Methods section 2.9).

The expression levels of *MUC13* transcripts were higher in the infected samples compared to the uninfected controls at 3 h and 24 h post infection. However, the increase in *MUC13* transcript abundances were not statistically significant, suggesting that further repeats may be required to improve the robustness of the data (Figure 32). The RT-qPCR analysis also showed that the FGG transcript was downregulated at 3 h post infection, consistent with what was observed in the dual RNA-seq data. However, the downregulation was also not statistically significant (Figure 32). Finally, *ANXA6* was upregulated at 3 h and 24 h post infection, which was also consistent with the RNA-seq data, but not statistically significant (Figure 32). Overall, the RT-qPCR was mostly consistent with what was observed in the dual RNA-seq analysis, suggesting that the differential gene expression analysis accurately represented the levels of transcripts in the RNA samples. However, more experimental repeats may be required to improve the statistical significance and reliability of the data (Figure 32).

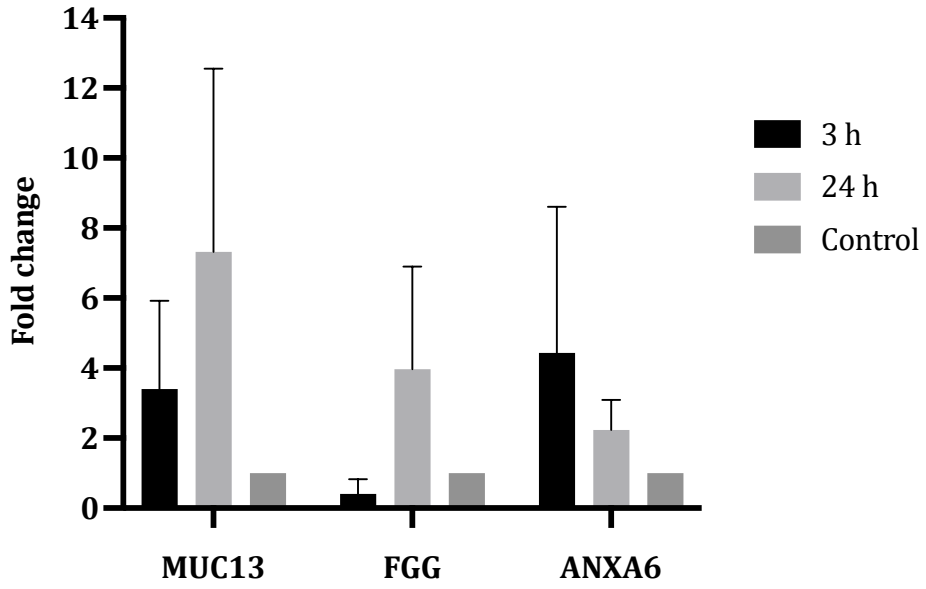


Figure 32. RT-qPCR quantification of selected transcripts from RNA samples extracted at 3 h and 24 h post infection.

Fold changes between the uninfected controls and infected samples were calculated for MUC13 (A), FGG (B) and ANXA6 (C) transcripts, calculated with the $2^{-\Delta\Delta Ct}$ method. Ct values were normalised to the expression levels of the GAPDH housekeeping gene from the same sample. The control bars indicate the $2^{-\Delta\Delta Ct}$ uninfected control value equal to 1. Statistical significance was analysed from the ΔCt values of each gene with multiple unpaired t-tests, and no significant differences were identified, $p > 0.05$ (n=3). Bars are representative of the mean +/- standard deviation.

5.2.3 Investigating the role of MUC13 in *C. difficile* pathogenesis

One of the most interesting, upregulated genes identified in this analysis was *MUC13*, a mucin highly expressed in the human colon (Moehle *et al.*, 2006). Mucus production is crucial for the proper functioning of the colon, and other mucins have been reported to play a role in *C. difficile* pathogenesis (Engevik *et al.*, 2015, 2020; Semenyuk *et al.*, 2015; Soavelomandroso *et al.*, 2017). As well as serving as a protective barrier for the intestinal epithelium, MUC13 is also

involved in the regulation of pro-inflammatory cytokines, such as IL-8 (Sheng *et al.*, 2011). Therefore, I hypothesised that the MUC13 mucin may have an important role in the host mucosal immune response to *C. difficile* infection.

5.2.3.1 Quantification of MUC13 by confocal microscopy

MUC13 undergoes several post-translational modifications to form the mature MUC13 transmembrane glycoprotein (van Putten and Strijbis, 2017). To confirm the upregulation of *MUC13* observed in the dual RNA-seq analysis and investigate whether the levels of mature glycoprotein were also increased during *C. difficile* infection, the levels of the MUC13 glycoprotein were quantified in uninfected and *C. difficile*-infected cells. Intestinal epithelial cell layers were incubated in the VDC system with or without *C. difficile* and stained for imaging with confocal microscopy. An anti-MUC13 primary antibody and an AlexaFluor-488 secondary antibody were applied to fluorescently label the MUC13 glycoprotein. Qualitative analysis of the confocal microscopy images showed that more fluorescently labelled MUC13 was present in the infected samples at 24 h (Figure 33A) and 48 h (Figure 33B) post infection. To quantify the relative amount of MUC13 glycoprotein, the average fluorescence from the 488 nm channel over the z-stack was quantified and compared between the uninfected and infected cell layers at 24 h and 48 h. It was found that there was significantly more 488 nm fluorescence in the infected cell layers compared to the uninfected cell layers at 24 h and 48 h post infection (p <0.0001 quantified by two-way ANOVA) (Figure 33C).

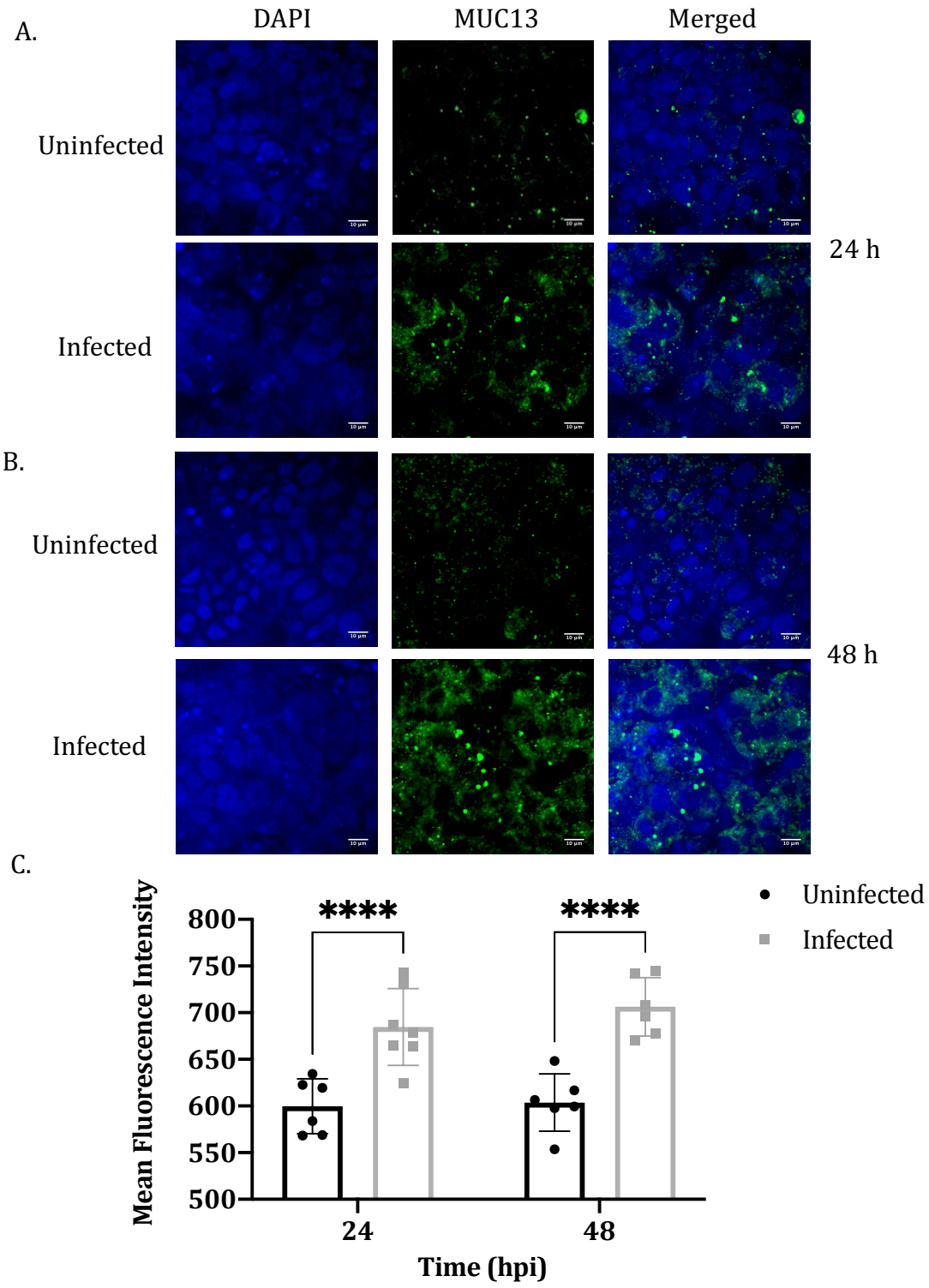


Figure 33. Confocal microscopy of the MUC13 glycoprotein at 24 h and 48 h after *C. difficile* infection.

C. difficile-infected intestinal epithelial cells incubated in the VDC system for 24 h (A) or 48 h (B) were stained with anti-MUC13 antibody (green) and DAPI (nuclear DNA, blue). C. Quantification of mean fluorescence intensity from the average projection of the z-

stack of the 488 nm/green channel. ****p <0.0001 quantified by two-way ANOVA. Microscopy images and fluorescence quantification representative of 3 biological replicates and 5-7 technical replicates. Bars are representative of the mean +/- standard deviation.

5.2.3.2 Quantification of MUC13 by immunoblotting

To further confirm the modulation of the MUC13 glycoprotein during *C. difficile* infection, a western blot analysis was performed with an anti-MUC13 antibody. Proteins were extracted from uninfected, or *C. difficile*-infected intestinal epithelial cells incubated in the VDC system for 3 h or 24 h and analysed by SDS-PAGE and subsequent immunoblotting. Western blot analysis showed that at 3 h post infection there were similar quantities of MUC13 (~55 kDa). However, at 24 h post infection there was more MUC13 present in the *C. difficile*-infected cells than the uninfected cells (Figure 34A). Multiple bands present in the western blot may be due to post-translational modifications of the mucin glycoprotein, which results in different sizes of the mucin following cell lysis. After imaging, the membrane was stripped and re-probed with an anti-actin antibody to quantify the levels of actin, a loading control used to ensure comparable quantities of total protein were loaded to each well. Protein band densities were normalised to the loading control bands to quantify the relative expression levels of MUC13. The results of the densitometry analysis also showed that there was an upregulation of MUC13 during infection at the 24 h timepoint (Figure 34B). These results were consistent with what was shown in the dual RNA-seq results, where there was an upregulation in MUC13 24 h post infection.

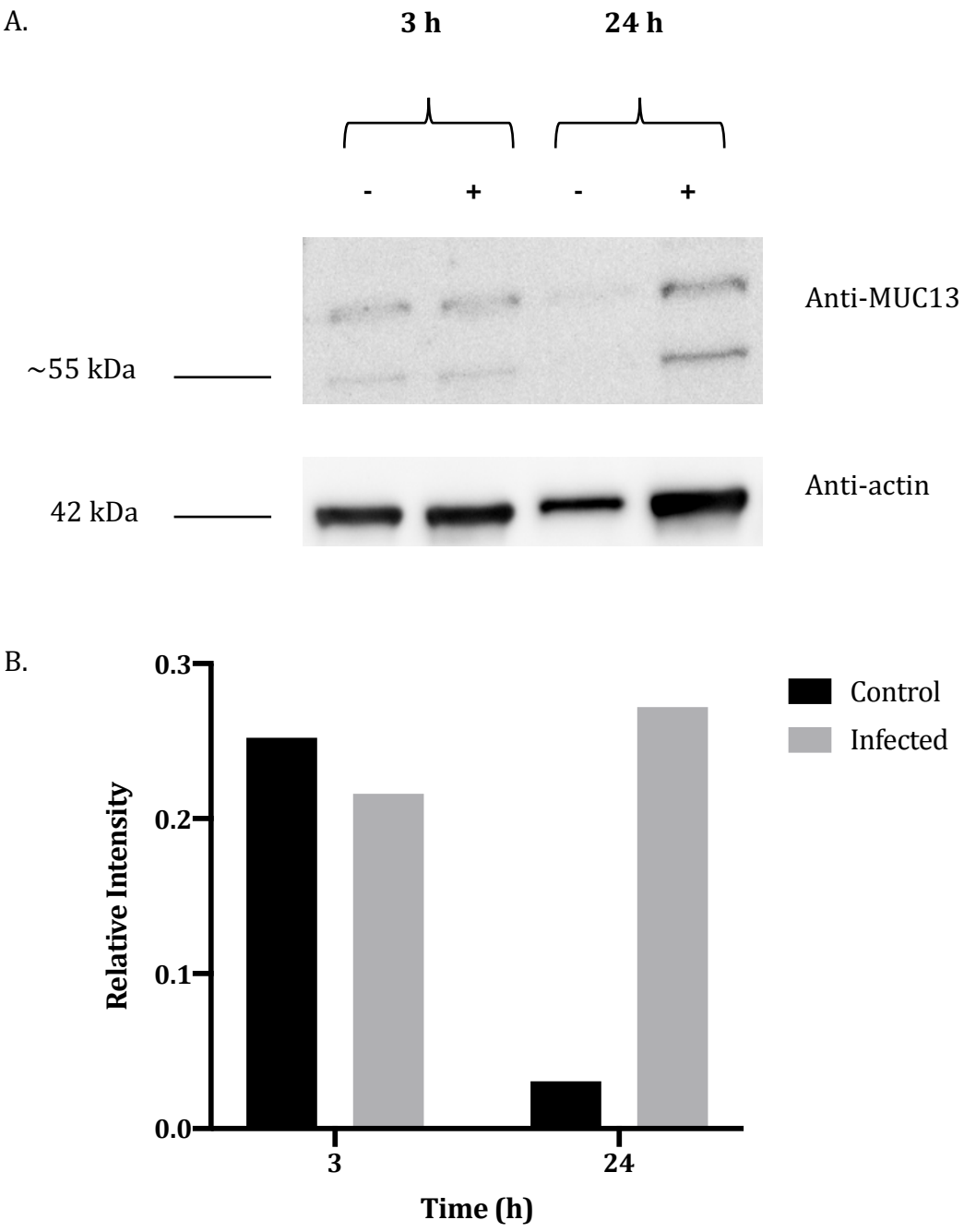


Figure 34. Western blot analysis to quantify the MUC13 glycoprotein.

A. Intestinal epithelial cells were incubated in the VDC system with (+) or without (-) *C. difficile* for 3 h and 24 h. Proteins were extracted and separated by SDS-PAGE on a 10% gel. B. Densitometry analysis performed by normalising the average protein band density to the loading control. Data representative of 2 biological replicates.

5.2.3.3 CRISPR/Cas9 knockout of MUC13

Since *C. difficile* has been reported to bind to mucus and MUC13 regulates the secretion of inflammatory cytokines, I hypothesised that MUC13 may have a role in the pathogenesis of *C. difficile* infection (Tasteyre *et al.*, 2001; Sheng *et al.*, 2011; Engevik *et al.*, 2020). To investigate the role of MUC13 in *C. difficile* infection, the CRISPR/Cas9 system was used to generate MUC13 knockout Caco-2 and HT29-MTX cell lines. Two sgRNAs were designed to target different regions of the MUC13 gene. sgRNA constructs were cloned into the pSpCas9(BB)-2A-GFP vector and transfected into Caco-2 and HT29-MTX cell lines. Transfected cells were isolated with FACS and single transfected clones were isolated and frozen for storage as described in Methods section 2.17. To investigate whether MUC13 knockout cell lines were obtained, target regions were amplified by PCR and sequences were examined with GATC sanger sequencing. Initial screening did not detect knockout clones, although further screening is currently being continued in the laboratory. If knockout mutants are obtained, experiments will be performed to investigate the role of MUC13 in *C. difficile* infection. The VDC system will facilitate infection of WT or *MUC13* knockout cell lines with *C. difficile*. The role of MUC13 in *C. difficile* colonisation will be investigated by comparing the numbers of bacteria adhered to the WT and *MUC13* knockout intestinal epithelial layers with CFU counts. To investigate whether MUC13 has a role in IL-8 secretion during *C. difficile* infection, the release of IL-8 from WT and *MUC13* knockout cell lines will be quantified with an IL-8 ELISA.

5.2.4 Pathway enrichment analysis identified pathways modulated during *C. difficile* infection

5.2.4.1 Identification of differentially expressed pathways

To investigate cellular processes and pathways which were modulated during *C. difficile* infection, a pathway enrichment analysis was conducted. This technique groups individual genes into the pathways they are involved in and can provide insights into the cellular processes which are modulated between two conditions. To obtain advanced mammalian pathway analysis data, we collaborated with Dr Ludmila Rodrigues Pinto Ferreira from The Federal University of Minas Gerais who is a specialist in RNA biology. The dual RNA-seq gene expression data was imported into the QIAGEN IPA pathway analysis software to investigate pathways which were modulated during infection. The outputs of the pathway enrichment analyses were visualised by plotting the percentages of up- or downregulated genes out of the total genes in the pathway at each timepoint (Appendix figure 1). The calculation of the z-score is a method commonly used in RNA-seq analyses to indicate the positive or negative regulation of a gene or pathway on a normalised scale. The z-score was calculated by subtracting the overall average gene abundance from the raw expression value for each gene and dividing by the standard deviation (SD) of all of the measured counts across all samples. The z-scores of the differentially expressed pathways were calculated and illustrated in Figure 35.

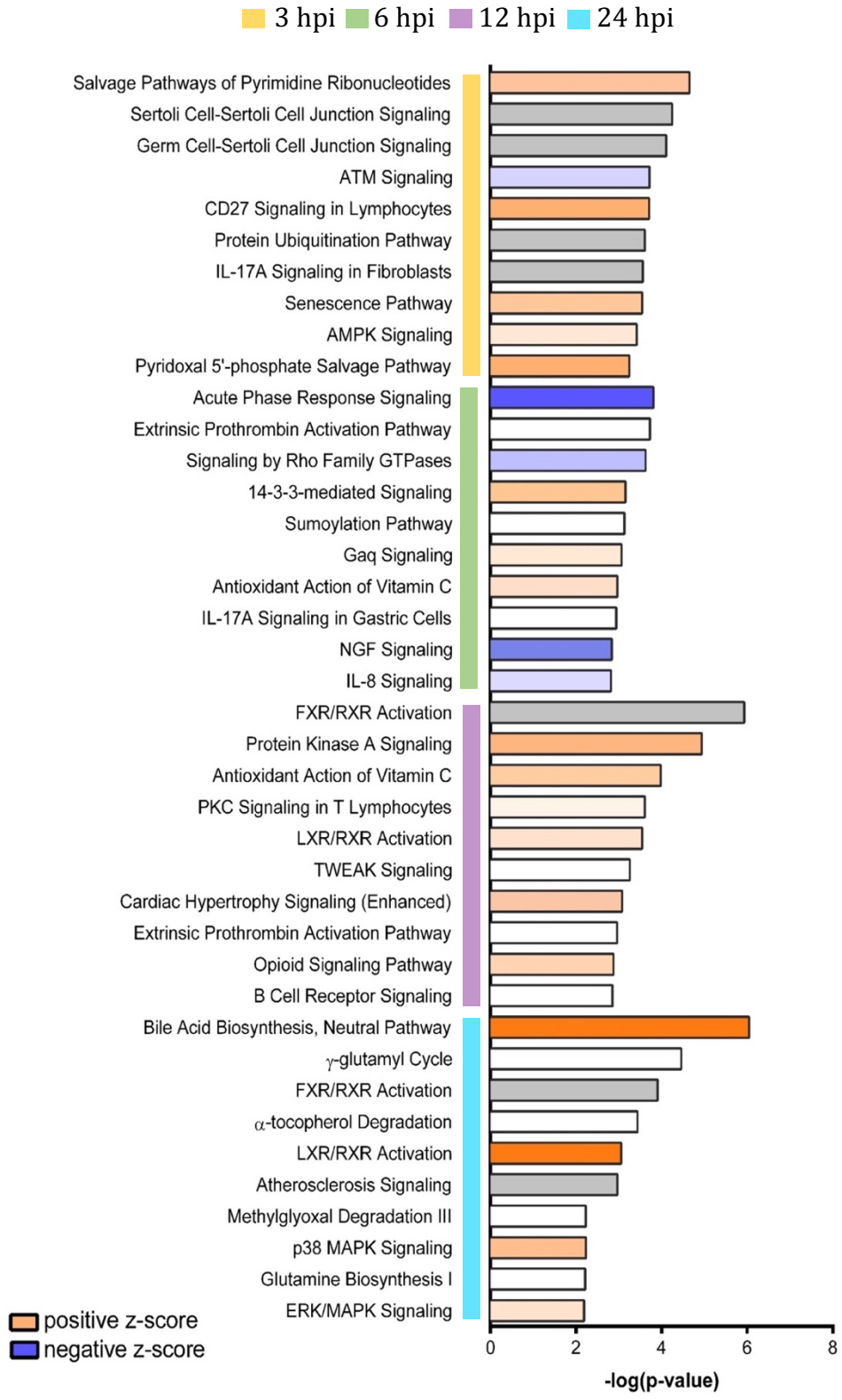


Figure 35. The z-scores of the top 10 most significantly differentially expressed pathways at each timepoint.

Figure produced by Ludmila Rodrigues Pinto Ferreria with the Qiagen IPA software.

At 3 h post infection the most significantly modulated pathway was the salvage pathways of pyrimidine ribonucleotides. This pathway had a positive z-score, suggesting that it was upregulated during infection despite some genes in the pathway being downregulated (Figure 35). This pathway is responsible for the recycling of nucleotides from DNA and RNA breakdown as well as uptake of extracellular nucleotides, enabling them to enter the pathways of pyrimidine biosynthesis (Löffler *et al.*, 2005). Some other interesting differentially expressed pathways at 3 h post infection included Sertoli and germ cell-Sertoli cell junction signalling, CD27 signalling in lymphocytes, IL-17A signalling in fibroblasts and 5′-AMP-activated protein kinase (AMPK) signalling.

At 6 h post infection the most significantly differentially expressed pathway was acute phase response signalling, which had a negative z-score, indicating that it was downregulated during *C. difficile* infection. This pathway is involved in the alterations of gene expression and metabolism in response to inflammatory cytokine signalling (Kushner and Rzewnicki, 1994). Other interesting pathways modulated at 6 h post infection include the extrinsic prothrombin activation pathway, signalling by Rho GTPases, IL-17A signalling in gastric cells, and IL-8 signalling.

At 12 h post infection the most significantly modulated pathway was farnesoid X receptor (FXR)/ retinoid X receptor (RXR) activation, which has a role in the metabolism of bile acids (Stanimirov, Stankov and Mikov, 2012). Other interesting pathways differentially expressed at 12 h post infection include Protein kinase C (PKC) signalling in T lymphocytes, liver X receptor (LXR)/RXR

activation, TNF-like weak inducer of apoptosis (TWEAK) signalling, extrinsic prothrombin activation pathway and B cell receptor signalling.

At 24 h post infection the most significantly modulated pathway was the neutral bile acid biosynthesis pathway which was highly upregulated with a positive z-score. Other interesting differentially expressed pathways included FXR/RXR activation, LXR/RXR activation, p38 MAPK signalling and extracellular signal-regulated kinase (ERK)/MAPK signalling.

To investigate the differential expression of pathways which were modulated during *C. difficile* infection across the course of infection, a heatmap was generated with the z-score to illustrate the up- or downregulation of each pathway at each timepoint. Hierarchical clustering was used to indicate the genetic distances between the timepoints (Figure 36).

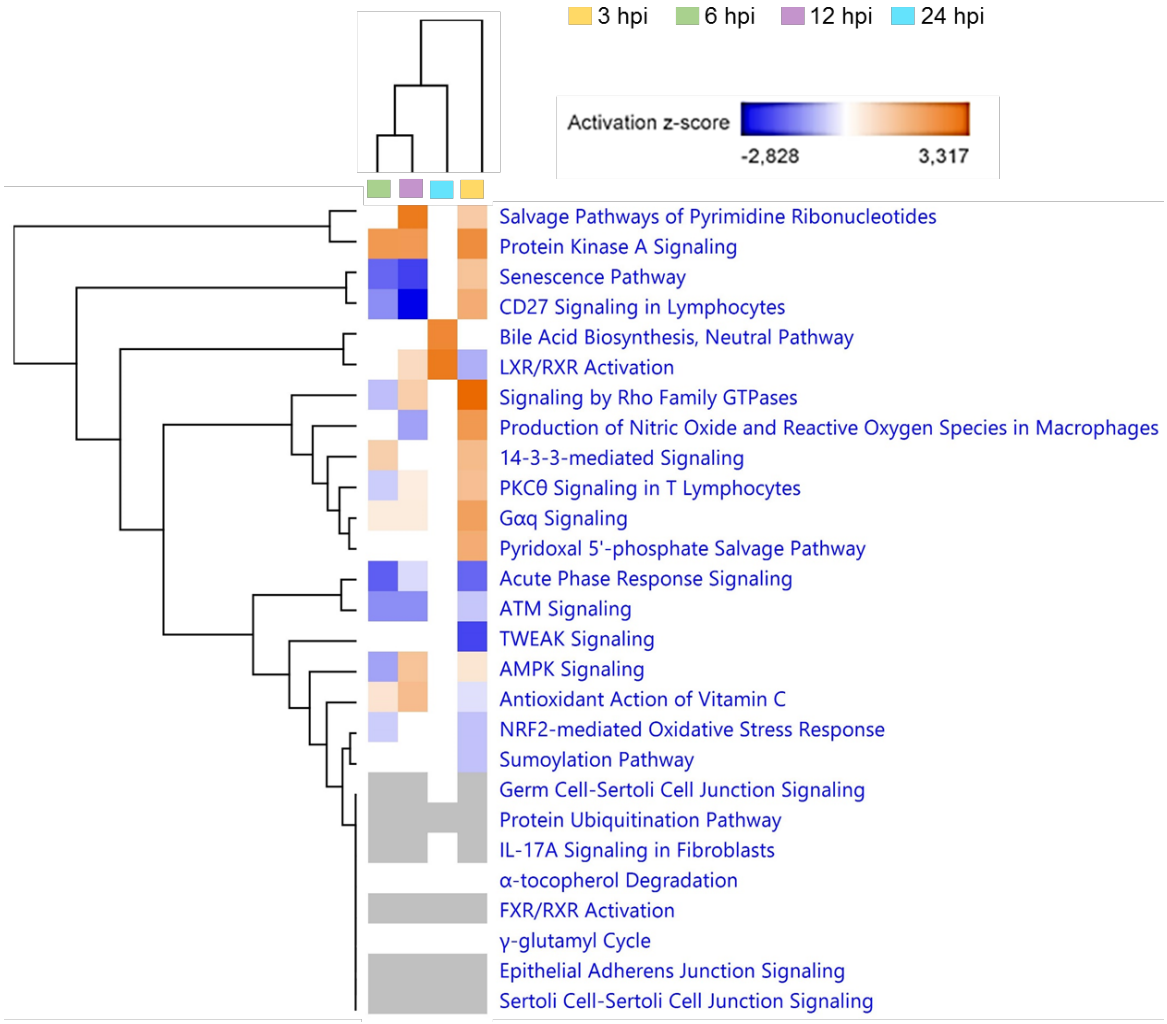


Figure 36. Heat map of top significantly differentially expressed pathways.

The up- (red) or downregulation (blue) of each pathway is indicated by the z-score. Hierarchical clustering was used to indicate the genetic distances in the gene expression profiles between the timepoints. Figure produced by Ludmila Rodrigues Pinto Ferreria with the Qiagen IPA software.

5.2.4.2 Salvage pathways of pyrimidine ribonucleotides were upregulated during infection

The largest transcriptional response to *C. difficile* infection was observed at 3 h post infection, where a high number of genes were differentially expressed. The

most significantly modulated pathway at 3 h post infection was the salvage pathways of pyrimidine ribonucleotides which had a positive z-score, indicating that it is upregulated during infection. The modulation of this pathway during *C. difficile* infection is interesting because extracellular nucleotides have been reported to be regulators of the immune response (Löffler *et al.*, 2005; Garavaglia, Rossi and Landini, 2012; Inami, Kiyono and Kurashima, 2018). To investigate the modulation of this pathway in response to *C. difficile* infection further, the alterations in gene expression of each gene involved the pathway were indicated using red colouring for upregulated genes and green colouring for downregulated genes (Figure 37).

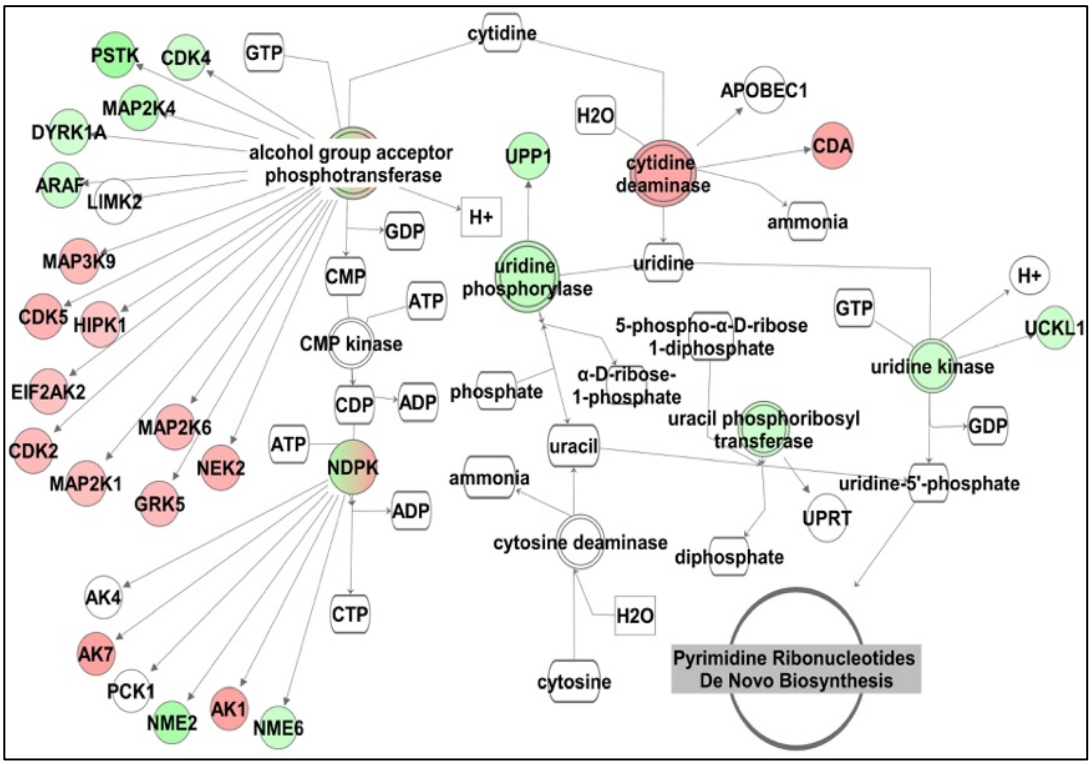


Figure 37. The differential expression of genes involved in the salvage pathways of pyrimidine ribonucleotides at 3 h post infection.

Upregulated genes are indicated by red colouring and downregulated genes are indicated by green colouring. Figure produced by Ludmila Rodrigues Pinto Ferreria with the Qiagen IPA software.

At 3 h post infection, there was an upregulation of 12 genes and downregulation of 13 genes involved in the salvage pathways of pyrimidine ribonucleotides (Figure 37). At the other three timepoints there were also modulation of the genes involved in the salvage pathways for pyrimidine ribonucleotides, including cytidine deaminase which was significantly downregulated at all timepoints (Appendix figure 2). I also analysed the differential expression of genes involved in the salvage pathways of pyrimidine ribonucleotides using a heatmap. The normalised gene expression values of genes involved in the salvage pathways of

pyrimidine ribonucleotides were plotted on a heatmap and hierarchical clustering was applied to group genes and samples with similar gene expression profiles (see Methods section 2.7). In general, the majority of the samples from the four infected groups were clustered together more closely than the uninfected groups. Contrasts in the colour scales can also be observed for most genes, illustrating the differential expression of this pathway between the uninfected controls and infected sample groups. Similar to what was observed during hierarchical clustering with a heatmap containing the majority of the human genome, the 24 h uninfected controls were clustered more closely to the infected sample groups than the other uninfected controls. The differences in gene expression between the infected samples and uninfected control groups demonstrates that *C. difficile* infection is causing a transcriptional shift in the salvage pathways of pyrimidine ribonucleotides (Figure 38).

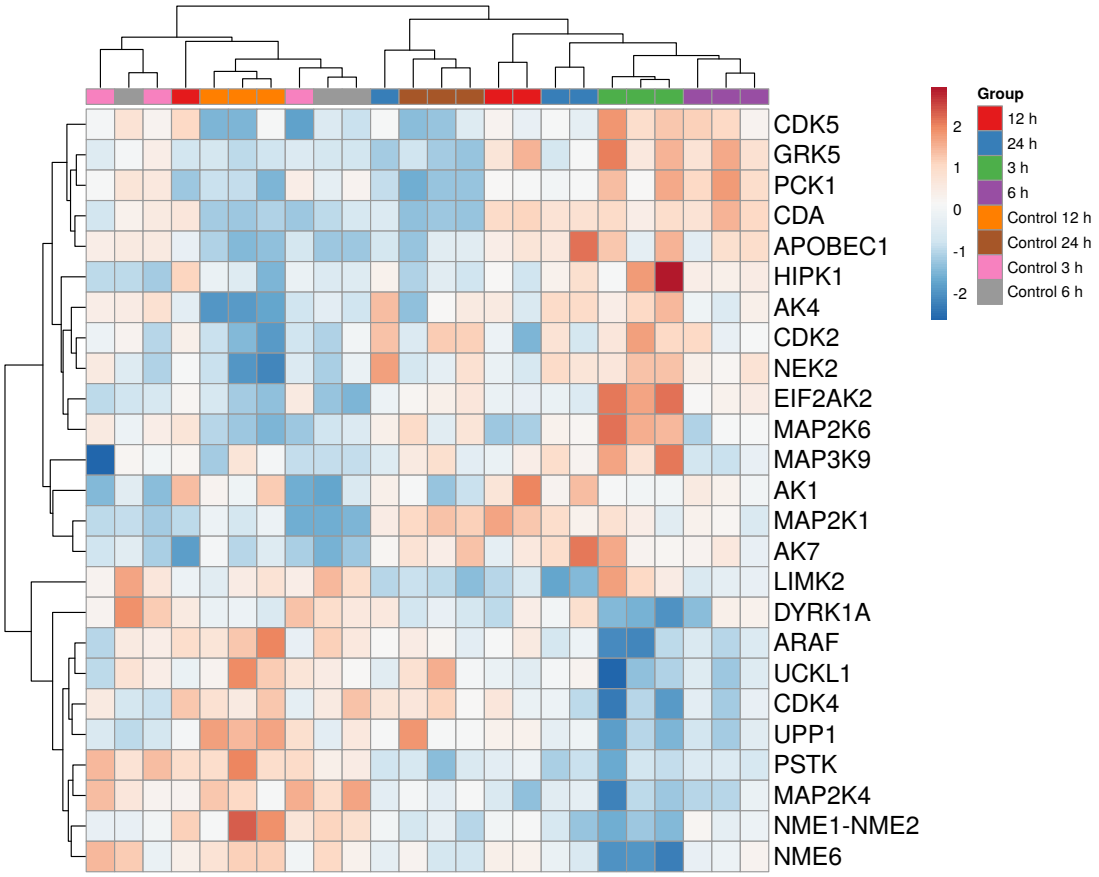


Figure 38. Heatmap of the genes involved in the salvage pathways of pyrimidine ribonucleotides at 3 h post infection.

Regularised logarithmic transformation was applied to the output from DESeq2. The z-score was used to scale the data. The Euclidian distance was used to calculate the genetic distances between samples, and average linkage was used to cluster the rows and columns.

As well as the salvage pathways of pyrimidine ribonucleotides, there were a number of other significantly enriched pathways at 3 h post infection which have a role in the biosynthesis or metabolism of nucleotides. These pathways include the nucleotide excision repair pathway, pyrimidine ribonucleotides interconversion, pyrimidine ribonucleotides de novo biosynthesis, pyrimidine deoxyribonucleotides de novo biosynthesis I, salvage pathways of pyrimidine

deoxyribonucleotides, guanosine nucleotides degradation III, purine nucleotides degradation II (aerobic) and adenosine nucleotides degradation II (Appendix table 16). The analysis of the other pathways indicates that both purine and pyrimidines are modulated during *C. difficile* infection.

5.2.5 Quantification of extracellular ATP

Host pathways involving purine and pyrimidine nucleotides were significantly differentially expressed during *C. difficile* infection. As discussed in chapter 4, the purine and pyrimidine metabolic pathways were also significantly differentially expressed in the bacteria during infection. ATP is a purinergic signalling molecule with a fundamental role as the energy source for cells. Extracellular ATP (eATP) is also involved in the inflammatory response and has been described as an inter-kingdom signalling molecule (Spari and Beldi, 2020). To investigate the role of purinergic signalling during *C. difficile* infection, an ATP release assay was conducted. To quantify the amount of eATP released during *C. difficile* infection, a luminescence-based cell viability assay was used, as described in Methods section 2.16. ATP is required for the conversion of luciferin into oxyluciferin which emits a luminescent signal. The quantity of luminescence produced is directly proportional to the concentration of ATP in the reaction and can therefore be measured to quantify the concentration of eATP when used in conjunction with a standard curve of known ATP concentrations. ATP concentrations in supernatants from uninfected and *C. difficile*-infected intestinal epithelial cells incubated in anaerobic conditions were compared. The quantity of eATP from a bacterial inoculum incubated in prereduced DMEM-10 was also

measured to determine whether a portion of the ATP released during infection may have been secreted from the bacterial cells.

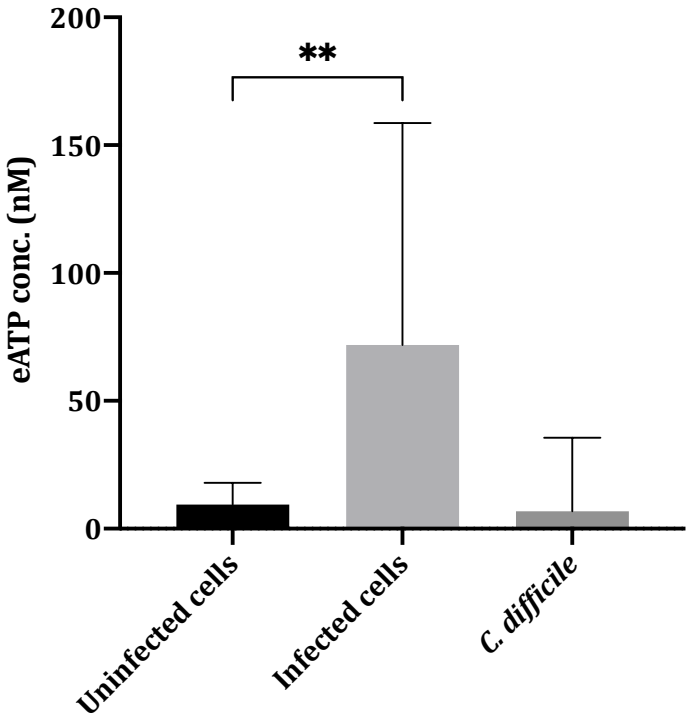


Figure 39. ATP release assay from uninfected cells, *C. difficile*-infected cells or a bacterial inoculum incubated in pre-reduced DMEM-10 for 4 h under anaerobic conditions.

Bars are representative of mean values from 9 technical replicates from 2 biological replicates with standard deviation error bars. $^{**}p = 0.007$, quantified by paired t-test.

At 4 h post infection, there was significantly more ATP released into the supernatant from *C. difficile*-infected intestinal epithelial cells compared to the uninfected cells ($^{**}p = 0.007$, quantified by paired t-test). Interestingly, there was also a substantial amount of ATP released from the bacterial inoculum (Figure 39). This suggests that some of the eATP in the supernatant from *C. difficile*-infected cells may have been released from the bacteria. However, we cannot

determine the amounts of ATP specifically released by the host cells or bacteria in this experiment.

5.3 Discussion

Host responses to *C. difficile* infection play a critical role in *C. difficile* pathogenesis. However, little is known about host molecular mechanisms involved in *C. difficile* infection. A better understanding of host pathways activated or inhibited during infection would provide a clearer understanding of the mechanisms involved in the induction of an immune response to infection. Although there are currently no published reports of dual RNA-seq analysis using gut epithelial cells, previous transcriptional analyses conducted with lung epithelial cells may shed some light on epithelial layer responses to infection. Dual RNA-seq of *Streptococcus pneumoniae*-infected human lung alveolar epithelial cells showed that adherence of *S. pneumoniae* repressed innate immune responses, including the expression of IL-8 and the production of antimicrobial peptides. Several sugar transporters were also activated in response to bacterial adherence to epithelial cells and this activation was dependent on host-derived mucins (Aprianto *et al.*, 2016). Transcriptomic responses to *Haemophilus influenzae* infection in ciliated human bronchial epithelial cells have also been resolved with dual RNA-seq. The host cells elicited a significant dysregulation of the cytoskeleton in response to bacterial infection, with an extensive effect on the filament network, as well as cell junction complexes (Baddal *et al.*, 2015). These studies demonstrated the complex regulatory systems which cells employ in response to epithelial pathogens. In this analysis, host responses to *C. difficile* infection were resolved by comparing

the gene expression profiles of infected samples to uninfected controls at each timepoint from an *in vitro* human gut model. There were numerous differentially expressed genes which are involved in the immune response, as well as other interesting genes which may have a role in *C. difficile* pathogenesis.

TNF Receptor Superfamily Members: some of the most differentially expressed genes were tumor necrosis factor receptor superfamily members TNFRSF12A and TNFRSF11B. These genes are involved in TNF-α signalling, a pathway which has been previously reported to be activated during the immune response against *C. difficile* infection (Czepiel *et al.*, 2014). siRNA knockdown of TNFRSF12A in a hepatocellular carcinoma cell line inhibited cellular proliferation and migration (Wang *et al.*, 2017). Therefore, the downregulation of this gene may contribute towards the arrest of cell proliferation, a process which has been linked to *C. difficile* toxins. TNFRSF11B encodes the gene for osteoprotegerin, a protein which has a role in the immune response towards tumor cells but has not yet been linked to infection-associated apoptosis. The secretion of osteoprotegerin is regulated by a small GTPase, Rab25, and protects cancer cells from the apoptosis through inhibition of TRAIL (Emery *et al.*, 1998; Cheng *et al.*, 2013). Therefore, it is possible that the downregulation of this gene could play a role in the host responses to *C. difficile* infection, as a reduction in TRAIL inhibition may activate TNF-family death receptors (Emery *et al.*, 1998).

Nuclear Factor Kappa B Subunit 1: NF-κB is a major transcriptional regulator of the immune response, apoptosis and cell-growth. NF-κB signalling is upregulated in IBD, ulcerative colitis and Crohn's disease, highlighting its role in

the regulation of inflammation in the colon (Karban *et al.*, 2004). Mice lacking NF-κB subunit genes show defects in clearing bacterial infection along with defects in B-cell and T-cell functions (Gerondakis *et al.*, 1999). Interestingly, the NFKB1 gene was significantly downregulated at 3 h post infection. This gene encodes the NF-κB p105/p50 isoforms which are inhibitors of the NF-κB signalling pathway, also known as IκBs. IkB has been associated with an increased risk of ulcerative colitis, an inflammatory bowel disorder (Karban *et al.*, 2004). Following cellular stimulation with agonists, NF-κB1 is targeted for degradation by the proteasome, reducing the inhibitory effect on NF-κB signalling. Mutations in the NF-κB1 gene have also been linked to defects in the immune response (Beinke and Ley, 2004). NF-κB has been reported to be stimulated during *C. difficile* infection and is essential for chemokine expression in intestinal epithelial cells in response to *C. difficile* TcdA (He *et al.*, 2002; Chae *et al.*, 2006; Kim *et al.*, 2006). Therefore, the downregulation of *NFKB1* during *C. difficile* infection may contribute towards the stimulation of the NF-κB signalling pathway, an important pathway for the inflammatory response in the gut.

Interleukin-8: IL-8 is an inflammatory cytokine and a potent chemoattractant for neutrophils from the blood to the damaged tissue (Hoffmann *et al.*, 2002). In healthy cells, IL-8 is barely detectible. However, IL-8 secretion is rapidly induced upon cell stress, damage or infection (Hoffmann *et al.*, 2002). An increase in IL-8 secretion has been strongly associated with *C. difficile* infection, and IL-8 polymorphisms have been linked to poor disease prognosis (Linevsky *et al.*, 1997a; Steiner *et al.*, 1997; He *et al.*, 2002; Jiang *et al.*, 2007). Surprisingly, in this analysis, *CXCL8*, the gene for IL-8 was significantly downregulated at 3 h, 6 h and

12 h post infection. Rather than *C. difficile* causing a downregulation of *CXCL8*, it is possible that the IL-8 secretion may have been stimulated in uninfected control samples during the incubation in the VDC system. Upon investigation, it was found that *CXCL8* had a high level of expression in the uninfected control samples, particularly individual samples at 6 h and 12 h which were extracted from the same set of cells. Oxygen deprivation has been reported to induce production of IL-8, so perhaps there was not enough oxygen supplied to that particular set of uninfected control samples which stimulated the release of IL-8 during incubation in the VDC system (Desbaillets *et al.*, 1997). This large upregulation observed in the uninfected samples may have skewed the RNA-seq data, resulting in the *CXCL8* being identified as downregulated when comparing infected samples to uninfected controls. It is also possible that IL-8 was not highly induced in the 3 h, 6 h and 12 h infected samples, as IL-8 secretion has only been previously reported after 24 h post infection in the VDC system (Anonye *et al.*, 2019). However, when comparing the gene expression profiles of infected cells between 3 h and 24 h post infection there was also no significant differential expression of *CXCL8* (Appendix table 15). The pathway enrichment analysis demonstrated that although the IL-8 signalling pathway was downregulated overall at 6 h post infection, there were several other genes involved in the pathway which were significantly upregulated which may be involved in the regulation of IL-8 secretion. The p38 MAPK signalling pathway was significantly upregulated at 24 h post infection. p38 MAPK signalling has been reported to be activated by *C. difficile* TcdA and has a role in the secretion of IL-8, monocyte necrosis and enteritis (Warny *et al.*, 2000; Kim *et al.*, 2005). Therefore, the

upregulation of this signalling pathway may contribute towards the secretion of IL-8 in response to *C. difficile* infection.

Fibrinogen: the genes which encode the fibrinogen *α, β* and *γ* chains, which combine to form the fibrinogen glycoprotein, were significantly downregulated at 3 h, 6 h and 12 h post infection. Fibrinogen is present circulating in the bloodstream and has an important role in wound healing, as well as being a component of the ECM (Mosesson, 2012). *C. difficile* has been reported to bind to immobilised fibrinogen *in vitro* (Cerquetti *et al.*, 2002; Hennequin *et al.*, 2003). Specifically the fibronectin-binding protein 68 (Fbp68) has been shown to adhere to both fibronectin and fibrinogen by far-immuno dot-blotting (Hennequin *et al.*, 2003). The *C. difficile* secreted protease, PPEP-1, has also been reported to cleave fibrinogen *in vitro*, suggesting that the degradation of fibrinogen may have a role in *C. difficile* pathogenesis (Cafardi *et al.*, 2013). Therefore, the downregulation in fibrinogen in this analysis may be induced by *C. difficile* colonisation and may result in a reduction of *C. difficile* adhesion to fibrinogen in the ECM of the gut epithelium. Fibrinogen downregulation may also impair the healing of wounds associated with *C. difficile*-induced colitis, and potentially contribute towards *C. difficile* toxin dissemination to the bloodstream to cause systemic disease (Steele *et al.*, 2012).

Wnt receptors: two colonic receptors for Wnt proteins were modulated during *C. difficile* infection. FZD4 was significantly downregulated, and LRP4 was significantly upregulated during *C. difficile* infection. In the intestines, Wnt/*β*-catenin signalling is particularly important for epithelial barrier development

and self-renewal (van der Flier and Clevers, 2009). Wnt proteins are secreted glycoproteins which bind to Frizzled receptor proteins or low-density lipoprotein receptor related proteins (LRP) at the N-terminal cysteine-rich domains (Komiya and Habas, 2008). Upon binding of a Wnt glycoprotein to a Frizzled receptor or LRP5/6, *β*-catenin accumulates in the cell and is then transported to the nucleus following Rac1 GTPase activation, where it binds to several target proteins, including the T-cell factor/lymphocyte enhancer factor (TCF/LEF) family of DNA-binding transcription factors (Komiya and Habas, 2008; Wu *et al.*, 2008). The Wnt/*β*-catenin signalling pathway has been associated with numerous bacterial infections. Bacterial interactions with the intestinal epithelium promote the activation of the pro-inflammatory NF-*κ*B signalling pathway, which is associated with degradation of *β*-catenin and results in the upregulation of pro-inflammatory cytokines such as IL-6, IL-8, and TNF-*a* in *Salmonella*-infected mice (Sun *et al.*, 2005). Moreover, *Bacillus anthracis* edema toxin induces activation of protein kinase A which triggers the reduction of phosphorylated glycogen synthase kinase 3*β* (GSK3*β*) levels in the target cell nucleus. This inactivates *β*-catenin and prevents it from binding to the TCF/LEF transcription factor, and thereby ultimately arrests the cell cycle (Larabee *et al.*, 2008). There are numerous studies which suggest *C. difficile* TcdA and TcdB interfere with Wnt/β-catenin signalling pathways. Frizzled Wnt glycoprotein receptors have been reported to be epithelial receptors for *C. difficile* toxin B, particularly FZD2 and FZD7 (Tao *et al.*, 2016), as well as FZD4, FZD8 and FZD9 which have also been reported to bind to TcdB (Chen *et al.*, 2018). TcdB binding to FZD receptors prevents Wnt protein engagement and blocks the Wnt/*β*-catenin signalling cascade (Chen *et al.*, 2018). The Wnt/*β*-catenin signalling

pathway is also attenuated by TcdA, which has been suggested to contribute towards the cellular proapoptotic and antiproliferative effects induced by this toxin (Lima *et al.*, 2014). The requirement of Rac1 GTPase activation for nuclear accumulation of dephosphorylated *β*-catenin, may suggest a further mechanism of Wnt signalling inhibition during *C. difficile* infection, as TcdB exhibits mono-glycosylation activity to GTPases, including Rac1 (Genth *et al.*, 2006; Wu *et al.*, 2008). However, this has not yet been experimentally confirmed. In summary, the differential expression of FZD4 and LRP4 observed in this analysis may modulate Wnt signalling and the expression of pro-inflammatory cytokines in response to *C. difficile* infection.

APOA4: one of the most significantly upregulated genes at 12 h and 24 h post infection was *APOA4*, an apolipoprotein which has a role in the formation of lipoproteins in the intestines (Ostos *et al.*, 2001). After synthesis, ApoA-IV is packaged into chylomicrons, and secreted into intestinal lymph during fat absorption (Recalde *et al.*, 2004). Studies have also found that ApoA-IV plays an important role in inflammatory responses, including during infections and inflammatory conditions such as allergic reactions. Specifically, when ApoA-IV is not bound to lipids it is predominantly found in circulation where it may play anti-inflammatory roles (Recalde *et al.*, 2004; Roula *et al.*, 2020). A previous study reported that ApoA-IV may suppress the adhesive interactions of leukocytes and platelets and dramatically reduce the upregulation of P-selectin in the colonic endothelium in a DDS-induced colitis mouse model. *APOA4* knockout mice also exhibited a significantly greater inflammatory response to DSS compared to the WT and this increased susceptibility to DSS-induced

inflammation was reversed upon exogenous administration of ApoA-IV to knockout mice (Vowinkel *et al.*, 2004). Furthermore, an increase in ApoA-IV levels in the blood of allergic rhinitis (also known as hay fever) patients after receiving allergen-specific immunotherapy has been reported and ApoA-IV was also found to decrease the secretion of histamine from basophils *in vitro* (Makino *et al.*, 2010). Moreover, ApoA-IV also has an effect on eosinophil cells which can cause tissue damage and have been associated with the severity of many inflammatory diseases such as allergic reactions and asthma (Sanderson, 1992). Systemic administration of ApoA-IV in a murine model of house dust mite extract-induced asthma reduced the quantity of circulating eosinophil cells in the serum and protected against airway hyperresponsiveness and leukocyte infiltration in the airways (Roula *et al.*, 2020). In addition, the expression of human ApoA-IV in $ApoE^{-/-}$ mice reduced oxidative stress, decreased the secretion of pro-inflammatory cytokines and reduced the formation of atherosclerotic lesions after LPS administration (Ostos *et al.*, 2001; Recalde *et al.*, 2004). Together, these findings demonstrate the crucial role of ApoA-IV in protecting against the potentially detrimental inflammatory responses, and therefore may have a role in the inflammatory response to *C. difficile* infection.

Mucin-13: it was demonstrated that the mucosal immunity protein, MUC13, was significantly upregulated during *C. difficile* infection. Mucus has been previously been reported to have a role in *C. difficile* pathogenesis. Patients with CDI exhibited decreased production of the secreted mucin, MUC2, and increased levels of the cell-surface mucin MUC1 in their faeces, yet the underlying molecular mechanisms have not been investigated (Engevik *et al.*, 2015). In a

multispecies biofilm community in a mouse model, communities containing *C. difficile* were found most commonly in the loosely attached mucus layer of the proximal colon, and rarely in the cell-surface mucus layer (Semenyuk *et al.*, 2015; Soavelomandroso *et al.*, 2017). Vegetative bacterial cells, as well as the *C. difficile* flagella components, FliC and FliD, have been reported to bind to mucus *in vitro* (Tasteyre *et al.*, 2001; Engevik *et al.*, 2020). Specifically, it was found that the ability of *C. difficile* to bind to the MUC2 mucin was highly conserved across all tested strains, and *C. difficile* bound to O-linked glycans from MUC2 in similar levels to the full glycoprotein, suggesting that *C. difficile* adheres to the O-linked glycan component of mucus (Engevik *et al.*, 2020).

MUC13 is a surface-associated mucin, which has multiple protective roles in the inflammatory response and response to pathogens. Firstly, as a cell surface-associated mucin, MUC13 provides a physical barrier to prevent pathogens reaching the epithelial cell layer (Williams *et al.*, 2001). MUC13 is also a regulator of the immune response, and is involved in the activation of NF-κB p65 signalling through the induction of phosphorylation and degradation of inhibitor of nuclear factor kappa B (IκBα) in response to TNF-α activation (Sheng *et al.*, 2013). MUC13 has been reported to protect against intestinal epithelial damage and inflammation in DDS-induced colitis and polymorphisms in the gene for MUC13 have been associated with inflammatory bowel disorders (Moehle *et al.*, 2006; Sheng *et al.*, 2011). MUC13 also stimulates the production of pro-inflammatory cytokines, including IL-8, in response to infectious stimuli following activation of NOD1 and TLR3 ligands (Sheng *et al.*, 2013). A recent dual RNA-seq analysis of Plasmodium-infected hepatocyte cells also observed a significant upregulation of

MUC13 in response to infection, and this induction was described as a hallmark of Plasmodium infection. This study demonstrated the role of MUC13 upregulation in response to pathogens. MUC13 is also an inhibitor of apoptosis (Sheng *et al.*, 2011). MUC13 overexpression has been reported to promote cell growth through the activation of Wnt signalling and has been associated with poor disease prognosis in cancer (Dai *et al.*, 2018). Collectively, these studies highlight the important role of MUC13 in the protection against pathogens and intestinal inflammation. Therefore, the upregulation of MUC13 may have a protective role in the response to *C. difficile* infection.

Pathway enrichment analysis: there were many interesting functional pathways identified which were modulated during *C. difficile* infection. Sertoli and germ cell-Sertoli cell junction signalling was significantly enriched at 3 h post infection. The genes which regulate cell junction signalling in Somatic cells are common to all cell-types, including Rho GTPase and actin genes. The glucosyltransferase activity of TcdA and TcdB disrupts cell junctions by inhibiting the activity of Rho GTPases and actin which are essential for the regulation of cell junctions. Ultimately, this results in the opening of tight junctions and loss of epithelial barrier integrity (Nusrat *et al.*, 2001; Walsh *et al.*, 2001; Taek *et al.*, 2012). Therefore, the modulation of these pathways during infection may be in response to *C. difficile* toxins.

The IL-17A signalling in fibroblasts pathway was significantly differentially expressed at 3 h post infection, and the IL-17A signalling in gastric cells pathway at 6 h post infection. IL-17 is a pro-inflammatory cytokine which recruits and

activates neutrophils and monocytes. IL-17 production has been linked to a range of autoimmune and infectious diseases and its pathogenic or protective role has been a controversial subject (Ruiz de Morales *et al.*, 2020). IL-17A neutralisation has been reported to increase tissue damage in a DDS-model of IBD and IL-17A- or IL-17RA-deficient T cells resulted in exacerbated disease in a mouse transfer model of colitis (Ogawa *et al.*, 2004; O'Connor *et al.*, 2009). Interestingly, IL-17 production has also been reported to regulate the permeability of the intestinal epithelial layer during DDS-induced colitis in a mouse model (Lee *et al.*, 2015). IL-17 also has a protective role in the innate immune response to pathogens. In the colon, IL-17A and IL-17F were reported to be critical for the production of the antimicrobial peptides, *β*-defensins, in response to *Citrobacter rodentium* infection (Ishigame *et al.*, 2009). IL-17 has also been reported to be essential in the immune response to cutaneous *Staphylococcus aureus* infection (Cho *et al.*, 2010). Moreover, IL-17 has been identified as a factor for determining the severity of *C. difficile* infection in mice, and IL-17 knockout mice were more resistant to infection (Nakagawa *et al.*, 2016). IL-17A produced by γδ T cells has also been reported to have an essential role in the immune response to *C. difficile* infection (Chen *et al.*, 2020). These studies demonstrate the protective role of IL-17 against T-cell mediated inflammation and bacterial pathogens. Therefore, although the IL-17 gene itself was not significantly differentially expressed, there may be other factors regulating IL-17 signalling during *C. difficile* infection. It is possible that the upregulation in IL-17 signalling may play a role in *C. difficile* infection, as its secretion can attract neutrophils, as well as reduce intestinal inflammation which has been linked to poor disease prognosis (Ng *et al.*, 2010).

AMP-activated protein kinase (AMPK) signalling was one of the top 10 most significantly enriched pathways at 3 h post infection. AMPK is an energy sensor which is activated during metabolic stress conditions and has a range of roles in various physiological and pathological processes (Silwal *et al.*, 2018). AMPK activation is a key bioenergetic regulator which matches mitochondrial ATP production. Mitochondrial ROS production also indirectly activates the AMPK signalling pathway (Hinchy *et al.*, 2018). Intracellular pathogens, such as *Legionella pneumophila* and *Mycobacterium tuberculosis* manipulate the AMPK signalling pathway to promote intracellular survival (Francione *et al.*, 2009; Singhal *et al.*, 2014). Metformin, an AMPK activator and stimulator of mitochondrial ROS production increases murine survival against *Legionella pneumophila* and *Mycobacterium tuberculosis* (Singhal *et al.*, 2014; Kajiwara *et al.*, 2018). AMPKs also have potential roles as regulators of the immune response to infection against pathogens such as *Bacillus thuringiensis*, as demonstrated in a *Caenorhabditis elegans* model (Yang *et al.*, 2015). The small-molecule AMPK activator, compound 13, inhibited *Helicobacter pylori*-induced apoptosis of gastric epithelial cells by modulating ROS levels via the AMPK-heme oxygenase-1 axis (Zhao *et al.*, 2015). Inhibition of AMPK signalling significantly abolished the protective effect of compound 13 against *H. pylori* within gastric epithelial cells (Zhao *et al.*, 2015). *C. difficile* TcdA and TcdB stimulate the production of ROS in the target intestinal epithelial cells which may indirectly activate the AMPK signalling pathway (Frädrich, Beer and Gerhard, 2016). *C. difficile* TcdB also stimulates an autophagy response, a process which is regulated by AMPK (He *et al.*, 2017). Therefore, the activation of the AMPK signalling pathway may

also have a protective role against *C. difficile*-induced apoptosis and may play a role in the immune response to *C. difficile* infection.

At 6 h post infection, there were many interesting significantly enriched pathways. Firstly, the most significantly enriched pathway was acute phase response signalling which had a negative z-score. This pathway is involved in the reprogramming of gene expression and metabolism in response to inflammatory cytokine signalling (Venteclef *et al.*, 2011). *C difficile* LPS, culture extract, and purified toxins A and B have been reported to stimulate the synthesis of acute-phase proteins in murine hepatocytes (Mazuski *et al.*, 2000). This suggests that the host cells may be detecting and responding to inflammatory cytokines which are secreted during *C. difficile* infection.

The extrinsic pro-thrombin pathway was also significantly enriched during *C. difficile* infection at 6 h and 12 h post infection, with approximately 25% of the total genes involved in this pathway being downregulated. This pathway has a vital role in coagulation and wound healing and was suppressed due to the downregulation of the three fibrinogen chain genes, as well as F3. The F3 gene encodes the cell-surface glycoprotein coagulation factor III, an activator of the coagulation cascade. As discussed above, the downregulation of this pathway may inhibit the repair of *C. difficile*-induced intestinal damage and potentially contribute to *C. difficile* dissemination.

The signalling by Rho GTPases pathway was also significantly downregulated 6 h post *C. difficile* infection. Rho GTPases are involved in a variety of molecular

processes within cells and have also been reported to be involved in the pathogenesis of many infectious diseases. For example, Rho GTPases are involved in the remodelling of actin to facilitate the internalisation of bacteria, such as *Salmonella* (Patel and Galán, 2008). Rho GTPases are also cellular targets of *C. difficile* TcdA and TcdB which are inactivated by mono-*O*-glucosylation (Gerhard *et al.*, 2008; Chen *et al.*, 2015). Rho GTPases also regulate cell-cell and cell-ECM adhesions. *C. difficile* toxins disrupt cell junctions and *C. difficile* binds to ECM components, further demonstrating a potential role for these proteins in the pathogenesis of *C. difficile* infection. Therefore, the alterations observed in this pathway in response to *C. difficile* infection may be due to the presence of toxins or *C. difficile* interactions with ECM components.

At 24 h post infection, the most significantly enriched pathway was the neutral bile acid biosynthesis pathway which had a positive z-score and an upregulation of over 30% of the genes involved in the pathway (Appendix figure 1). FXR/RXR activation, a pathway that regulates the transcription of bile acid biosynthesis genes was also significantly enriched at 24 h with a positive z-score (Gadaleta *et al.*, 2011; Stanimirov, Stankov and Mikov, 2012). It is now established that intestinal bile acids play a role in modulating several aspects of the *C. difficile* lifecycle. Primary bile acids, such as taurocholate, stimulate the germination of *C. difficile* spores into vegetative cells, whereas secondary bile acids inhibit spore germination (Sorg and Sonenshein, 2008, 2010; Giel *et al.*, 2010; Thanissery, Winston and Theriot, 2017). Therefore, the upregulation of bile acid biosynthesis 24 h post infection could perhaps increase the concentration of primary bile acids, and therefore increase *C. difficile* spore germination. Interestingly,

intestinal bile acids have also been reported to modulate the structure and function of *C. difficile* TcdB, inhibiting its binding to cell receptors (Tam *et al.*, 2020). Therefore, upregulation of bile acid biosynthesis in response to *C. difficile* infection may provide a mechanism by which the host cells could protect themselves against the action of TcdB. The FXR/RXR activation pathway is also been reported to suppress intestinal inflammation and protect the intestinal barrier in DDS-induced colitis (Gadaleta *et al.*, 2011). The inflammatory response and loss of epithelial barrier integrity have been associated with *C. difficile* pathogenesis, so activation of the FXR/RXR activation pathway may have a protective role during *C. difficile* infection.

Ribonucleotides in *C. difficile* infection: the most significantly enriched pathway at 3 h post infection was the salvage pathways of pyrimidine ribonucleotides. There were also several other differentially expressed pathways involved in the metabolism or biosynthesis of nucleotides at this timepoint, including pyrimidine ribonucleotides interconversion, purine nucleotides degradation II (aerobic) and adenosine nucleotides degradation II. This suggests that there may be a role for purine and pyrimidine nucleotides in *C. difficile* pathogenesis. Nucleotides are the building blocks for nucleic acids and are essential for many cellular processes, including serving as the fundamental currency (ATP) of all cellular energy transactions. It is now established that nucleotides also act as extracellular signalling messengers, with roles in numerous processes, including the stimulation of lymphocyte proliferation, ROS generation, as well as cytokine and chemokine secretion (Coutinho-Silva *et al.*, 1999; Löffler *et al.*, 2005; Spari and Beldi, 2020). In healthy tissues, ATP is almost

exclusively localised intracellularly, and concentrations are low in the extracellular milieu. However, under conditions of stress and apoptosis, human cells secrete ATP through pannexin and connexin hemichannels. eATP has been considered an endogenous adjuvant that acts as a danger signal to initiate an inflammatory response through the activation of purinergic type 2 receptors (P2 receptors) (Coutinho-Silva *et al.*, 1999, 2005). The $P2X_7$ receptor is an ATP-gated ion channel which is involved in both innate and adaptive immune response pathways, including the activation of the inflammasome, as well as TLR, NF-κB, and MAPK signalling. eATP binding to the $P2X_7$ receptor activates a large-scale ATP release via its intrinsic ability to form a cytolytic membrane pore, as well as the stimulation of pannexin hemichannels to secrete ATP, boosting purinergic signalling. This process has been reported to have protective or deleterious effects during infections, inflammatory diseases and cancer. The contrasting effects are dependent on its levels of activation and the type of disease or pathogen (Savio *et al.*, 2018). Liu *et al.* (2018) investigated the role of the ATP-$P2X_7$ signalling axis in the immune response to *C. difficile* infection. In *C. difficile*-infected macrophages, the ATP-$P2X_7$ signalling pathway was reported to have a role in the production of the pro-inflammatory cytokine IL-1β, as well as inflammasome activation and subsequent caspase-1-dependent pyroptotic cell death. These findings demonstrated the role of eATP accumulation and ATP-$P2X_7$ signalling in *C. difficile* infection of macrophages. However, intestinal epithelial cells also express the $P2X_7$ receptor but role of the eATP signalling axis in response to *C. difficile* has not been investigated in these cells (Coutinho-Silva *et al.*, 2005). A study by Proietti *et al.* (2019) found that follicular T helper cells in the Peyer's patches of gut-associated lymphoid tissue were modulated by ATP

released by commensal bacteria. Since eATP binds to the $P2X_7$ receptor of follicular T helper cells which forms a cytolytic pore, eATP may act as a mechanism which suppresses commensal-specific IgA responses. Treatment with apyrase, an ATP/ADP-hydrolysing enzyme, improved the IgA response against live oral vaccines for enteropathogenic *E. coli*. This further demonstrated the important role of eATP in the modulation of the immune response in the gut, as well as a potential mechanism by which bacteria can regulate immune cells (Proietti *et al.*, 2019).

In the bacterial transcriptomic analysis discussed in chapter 4, there was also an upregulation in the metabolism of purine and pyrimidine ribonucleotides. There are several advantages extracellular nucleotides may serve for a bacterial pathogen during infection. Bacteria have been reported to stimulate the secretion of ATP from host cells as well as secreting it themselves. Extracellular ATP may have a role in the suppression of the immune response through the $P2X_7$ receptor, as well as advance the detrimental effects of the inflammatory response during infections (Robertson *et al.*, 2010; Proietti *et al.*, 2019). An ATP release assay showed that significantly more ATP was released from *C. difficile*-infected intestinal epithelial cells than uninfected cells at 4 h post infection (Figure 39). This suggests that *C. difficile* colonisation may be stimulating the release of ATP from the intestinal epithelial cells. Interestingly, there was also a considerable amount of ATP released into the supernatant of the bacterial inoculum. Although we cannot confirm the amount of bacterial ATP released in the presence of cells, this suggests that a fraction of the ATP in the supernatant from infected cells may have been released by the bacteria. This finding is consistent with what was

discussed in chapter 4, where *C. difficile* purine metabolism was upregulated during infection. It is established that the inflammatory response has a role in the pathogenesis of *C. difficile* infection and may be deleterious for the outcome of the disease (Johal *et al.*, 2004; Muñoz-Miralles *et al.*, 2018; Wang *et al.*, 2018). Therefore, it may be a possibility that ATP released by *C. difficile* may contribute towards the detrimental effects of the inflammatory response and exacerbation of the disease.

To summarise this part of the work, there were many interesting human genes and pathways which were modulated during *C. difficile* infection which may have a role in the pathogenesis of the disease. The gene upregulation of MUC13 during *C. difficile* infection was confirmed at a protein level by confocal microscopy and western blot analysis. The role of extracellular nucleotides in *C. difficile* infection was investigated with an ATP release assay. It was found that *C. difficile* stimulated the release of ATP from host cells during infection and may also secrete ATP itself to modulate the inflammatory response.

Chapter 6: Discussion

For *C. difficile* to cause disease, an ingested spore must germinate into a vegetative cell in the anaerobic environment of the large intestines. *C. difficile* adhesion to the gut mucosa is crucial for the establishment of disease, yet we lack a comprehensive understanding of the host and bacterial factors required for colonisation. Host-pathogen interactions of *C. difficile* infection, in general, have been largely understudied. The analysis of transcriptomic responses to infection can greatly improve our understanding of the molecular processes which facilitate infections. Pervious dual RNA-seq analyses have identified new host and bacterial factors which are required to facilitate infections. For example, dual RNA-seq of *Samlonella typhimurium*-infected HeLa cells identified the bacterial small RNA, PinT as a regulator of genes required for intracellular survival and virulence. PinT was also reported to modulate the JAK/STAT signalling pathway in the host cells (Westermann *et al.*, 2016). Dual RNA-seq analysis of uropathogenic *Escherichia coli* (UPEC)-infected mouse macrophages demonstrated that ROS and H_2O_2 responses from the host activated bacterial genes required for protection against oxidative stress to promote intracellular survival (Mavromatis *et al.*, 2015). These studies demonstrated how dual RNA-seq can identify mechanisms by which host cells and pathogens interact during infection.

To perform a dual RNA-seq analysis of *C. difficile*-infected intestinal epithelial cells, a dual environment *in vitro* human gut model was used to facilitate co-

incubation of *C. difficile* with host cells. To our knowledge, this is the first dual RNA-seq study which has been conducted with an anaerobic microbe, in the presence of human gut epithelial cells, probably due to a lack of appropriate infection models. The successful application of the VDC system and dual RNA-seq used in this analysis may be applied to other anaerobic microbes to resolve host-pathogen interactions during infection.

Many interesting differentially expressed host and bacterial genes were identified which may have important roles in *C. difficile* pathogenesis. A downfall of performing dual RNA-seq in a simplified infection model, such as the VDC system, is that there is currently no active immune response element in these systems. The inflammatory response has been recently reported to alter *C. difficile* gene expression during infection, demonstrating that the host response may act as an important environmental cue to regulate *C. difficile* pathogenesis in the gut (Fletcher *et al.*, 2021). Our results also demonstrate that extended incubation in the VDC system caused a stress response in uninfected cells, potentially due to oxygen starvation or the accumulation of metabolic by-products. The development of a fluidic model may overcome these challenges, as well as generate an environment more representative of the human gut (Kulthong *et al.*, 2021). The adaptation of the VDC system to include millifluidic anaerobic and aerobic media channels may improve the oxygen supply to the host cells, as well as remove metabolic by-products during infection.

PPEP-1 was identified as a negative modulator of epithelial cell colonisation and our preliminary data suggest that PPEP-1 may have a role in the release of *C.*

difficile from collagen and intestinal epithelial cell layers. The toxin-mediated opening of cell junctions during *C. difficile* infection may modulate bacterial penetration of the basement membrane, to colonise the ECM (Frost, Cheng and Unnikrishnan, 2021). Therefore, our results suggest that PPEP-1 may be a factor required for *C. difficile* dispersal and penetration of gastrointestinal tissues. *ΔPPEP-1* has been previously reported to exhibit an attenuation in virulence in a hamster model (Hensbergen *et al.*, 2015). This may further suggest that bacteria lacking PPEP-1 may be more readily cleared by the immune response and antibiotic treatments. Therefore, PPEP-1 may be a promising vaccine candidate or a drug target for the prevention of *C. difficile* infection.

Mucus may have a fundamental role in the pathogenesis of *C. difficile* (Tasteyre *et al.*, 2001; Semenyuk *et al.*, 2015; Soavelomandroso *et al.*, 2017; Engevik *et al.*, 2020). Therefore, it was interesting to find that the MUC13 mucosal immunity protein was upregulated during *C. difficile* infection. The differential expression of MUC13 has not been previously identified in a transcriptomic analysis of *C. difficile*-infected cells, probably because probe-based assays only include selected markers. *C. difficile* has been reported to adhere to the O-glycan component of MUC2 (Engevik *et al.*, 2020). MUC13 also contains O-glycans, and therefore may provide an attachment site for *C. difficile* in the gut. *C. difficile* stimulates an inflammatory response, triggering the secretion of pro-inflammatory cytokines, such as IL-8, which have a protective role during infection (Linevsky *et al.*, 1997a; Steiner *et al.*, 1997; Warny *et al.*, 2000). Therefore, the upregulation in MUC13 may contribute towards the secretion of IL-8 and may serve a protective role in the mucosal immune response to *C.*

difficile infection. A dual RNA-seq experiment of *Plasmodium*-infected hepatocytes also identified MUC13 as a hallmark of infection, demonstrating a potential role for this mucin in the mucosal immune response to multiple infections (LaMonte *et al.*, 2019). MUC13 can be detected in peripheral blood and could potentially be used as a valuable clinical biomarker of *C. difficile* infection, which may improve current diagnostic techniques (Maher *et al.*, 2011). It would be particularly interesting to investigate whether MUC13 expression is altered between asymptomatic carriers and patients suffering with CDI, as this could potentially overcome the challenges associated with the current tests which cannot distinguish between them.

Metabolic pathways involving purine and pyrimidine ribonucleotides were identified as highly differentially expressed in both the bacterial and host transcriptomic analyses. This suggests that as well as being the building blocks for nucleic acids, ribonucleotides may also have a role in the pathogenesis of *C. difficile* infection. During conditions of stress, eukaryotic cells release the purinergic signalling molecule, ATP, into the extracellular milieu to act as a danger signal to circulating immune cells and triggering an immune response (Coutinho-Silva *et al.*, 2005; Inami, Kiyono and Kurashima, 2018; Proietti *et al.*, 2019; Spari and Beldi, 2020). *C. difficile* infection evokes an intense inflammatory response in the host, a process which has been linked to poor prognosis of the disease (Czepiel *et al.*, 2014; Rao *et al.*, 2014; Sun and Hirota, 2015; Yu *et al.*, 2017). Therefore, the release of eATP from the host as well as from the bacteria may contribute to the detrimental effects of the inflammatory response, a process which may be advantageous for the bacteria. A recent study also

demonstrated that *C. difficile* toxin-induced inflammation altered *C. difficile* metabolism, suggesting that the inflamed gut may provide an environmental niche to favour *C. difficile* pathogenesis. This study also found that toxin-induced intestinal inflammation upregulated host MMP collagenases, which break down collagen into Stickland metabolism substrates (Fletcher *et al.*, 2021). This may provide another mechanism by which *C. difficile* manipulates the host inflammatory response to favour bacterial growth. Moreover, it has also been suggested that intestinal inflammation may provide an uninhabitable environment for microbiota species, further promoting *C. difficile* growth (Pechine and Collignon, 2016). Therefore, intestinal inflammation induced by eATP may also favour *C. difficile* pathogenesis. eATP released by the microbiota have also been reported to negatively modulate the immune response in the gut through the cytolytic effects of ATP-$P2X_7$ receptor interactions on T follicular helper cells. This limits IgA responses in the gut and may be beneficial for the maintenance of the microbiota. However, during infection with an intestinal pathogen, a reduction in the abundance of T follicular helper cells and limitation of IgA responses by eATP may inhibit the clearance of the infection by the immune response and be detrimental for disease prognosis. Therefore, eATP released by the host and bacteria may act as an inter-kingdom signalling molecule and may have a significant role in the inflammatory response and *C. difficile* pathogenesis.

In conclusion, I demonstrated that transcriptional alterations are elicited from the host and bacteria in response to *C. difficile* infection. This analysis shed some light on the genes and pathways which are modulated during infection and

identified several factors which may be required for *C. difficile* pathogenesis. I identified a novel role for PPEP-1 in *C. difficile* pathogenesis which could potentially be a novel vaccine candidate for the prevention of the disease. I also identified MUC13 as being highly upregulated which may be important for the regulation of pro-inflammatory cytokines during *C. difficile* infection. This mucin could also be potentially used as a novel diagnostic marker for infection, possibly improving the current methods which can't distinguish between asymptomatic carriage or CDI. Finally, I demonstrated that ATP is released during *C. difficile* infection, which may contribute towards the inflammatory response and *C. difficile* pathogenesis.

Future work

This work has given insight into the host and bacterial responses to *C. difficile* infection. There were a number of interesting differentially expressed genes identified in the bacterial transcriptomic analysis which do not currently have functional roles assigned to them. There was a strong upregulation in the *C. difficile* cell wall proteins Cwp10 and CDR20291_2686 at all timepoints after infection, suggesting that they may have important roles during infection. To investigate their role in *C. difficile* pathogenesis, *Δcwp10* and *ΔCDR20291_2686* knockout strains could be generated with allelic exchange mutagenesis. These strains could be compared to the WT to determine whether there are differences in molecular processes, such as epithelial cell colonisation, biofilm formation, and their ability to elicit an IL-8 response in host cells. CDR20291_2686 also has a fibronectin binding domain, so its ability to adhere to immobilised fibronectin and to ECM produced by myofibroblast cells *in vitro* could also be examined. The results of this research may identify new colonisation factors of *C. difficile* which could potentially be therapeutic targets for the treatment and prevention of the disease.

We found evidence that PPEP-1 may be involved in the transition between sessile and motile lifestyles. The induction of a c-di-GMP phase shift by ectopically expressing genes encoding a diguanylate cyclase enzyme, which synthesises c-di-GMP, or a phosphodiesterase enzyme, which degrades c-di-GMP, may be an improved method for studying the role of PPEP-1 in the transition between

sessile and motile lifestyles. WT and *ΔPPEP-1* adhesion to collagen or intestinal epithelial cells before and after an induced c-di-GMP phase shift could be examined to investigate the role of PPEP-1 in *C. difficile* colonisation. Our results suggest that PPEP-1 may be involved in the dispersal of bacteria in the gastrointestinal tract. To investigate this, the numbers of WT and *ΔPPEP-1 C. difficile* adhered to different points along the gastrointestinal tract, as well as penetration into host tissues could be compared in a murine model. PPEP-1 could also be investigated for its effectiveness as a vaccine for the immunisation against *C. difficile* infection.

As discussed in chapter 5, *MUC13* knockout in Caco-2 and HT29 cell lines are currently being generated with the CRISPR/Cas9 system. These cell lines will be infected with *C. difficile* in the VDC system to investigate whether MUC13 has a role in *C. difficile* colonisation, IL-8 production and protection against cellular damage during infection. This research and subsequent analyses may identify MUC13 as a component of the mucosal immune response to *C. difficile*, as well as a novel biomarker for the disease. It would also be interesting to quantify and compare the levels of MUC13 in fecal samples from healthy people, asymptomatic carriers and patients suffering from CDI to determine whether there are alterations in the quantity or composition of excreted MUC13. *C. difficile* patient colon biopsies could also provide insights into the levels of MUC13 in the colon during *C. difficile* infection of the human gastrointestinal tract. Finally, MUC13 is a non-essential gene in mice, so *MUC13* knockout mutants could be generated to investigate whether if *MUC13* knockout mice are more susceptible to colonisation or death from *C. difficile* infection.

It was demonstrated that *C. difficile* infection stimulated the release of eATP from intestinal epithelial cells. Interestingly, there was also an increase in the ATP released into the supernatant of a bacterial inoculum incubated in DMEM, suggesting that *C. difficile* may also release ATP during infection. However, it was not possible to quantify the amount of bacterial ATP produced in response to infection. To investigate the quantities of bacterial ATP released during infection, the host cell ATP release could be obstructed by using pannexin or connexin hemichannel inhibitors, such as carbenoxolone disodium salt which inhibits Panx1. The difference between the eATP concentrations of infected cells with and without the hemichannel inhibitors may indicate the levels of ATP released by the bacteria. Although this is beyond the scope of this project, it would also be interesting to investigate the effect of eATP hydrolysis in the gut on the prognosis of the disease. For example, the large intestines of a *C. difficile*-infected mouse model could be treated with apyrase, an ATP hydrolysing enzyme, to investigate whether an alteration in eATP-induced intestinal inflammation could be linked to protective or damaging functions during infection.

Appendix

Appendix table 1. Bacterial rRNA oligonucleotide probes for RNase H rRNA depletion.

R20291_16S_1	CTT TTA ATT TGT CAT CCA ATT AAA TAT TTT GTT CTT TGT TTG GTA TTT CG
R20291_16S_2	GTC TAT AAA ACT ATT GTT ATC ATA GAC TCG GAC TAT TTG AAA ATA AAC TC
R20291_16S_3	TCA AAC TAG GAC CGA GTC CTA CTT GCG ACC GCC GCA CGG ATT GTG TAC GT
R20291_16S_4	TCA ACT CGC TAA ATG AAG CCA TTT CTC GCC GCC TGC CCA CTC ATT GCG CA
R20291_16S_5	CCC ATT GGA TGG GAC ATG TGT GCC TAT TGT ATG GCT TTC CAT ACG ATT AT
R20291_16S_6	GCC CTA TTA TAT AAA CTC TCC GTA GAG AAC TTA TAG TTT CCA CTC GGT CA
R20291_16S_7	TGT CCT ACC TGG GCG CAG ACT AAT CGA TCA ACC ATT CCA TTG CCG AAT GG
R20291_16S_8	TTC CGC TGC TAG TCA TCG GCT GGA CTC TCC CAC TAG CCG GTG TAA CCT TG
R20291_16S_9	ACT CTG TGC CAG GTT TGA GGA TGC CCT CCG TCG TCA CCC CTT ATA ACG TG
R20291_16S_10	TTA CCC GCT TTC GGA CTA CGT CGT TGC GGC GCA CTC ACT ACT TCC GGA AG
R20291_16S_11	CCC AGC ATT TTG AGA CAG GAG TTC CTT CTA TTA CTG CCA TGA ACT CCT CC
R20291_16S_12	TTC GGG GCC GAT TGA TGC ACG GTC GTC GGC GCC ATT ATG CAT CCC CCG AT
R20291_16S_13	CGC AAT AGG CCT AAA TGA CCC GCA TTT CCC ACG CAT CCG CCA GAA AGT TC
R20291_16S_14	AGT CCT CAC TTT CCG ATG CCG AGT TGG CAT CAT TCG AGA ACT TTG ACC CT
R20291_16S_15	CTG AAC TCA CGT CCT CTC CTC TCA CCT TAA GGA TCA CAT CGC CAC TTT AC
R20291_16S_16	GCA TCT ATA ATC CTC CTT GTG GTC AAC GCT TCC GCC GAG AGA CCT GAC AT
R20291_16S_17	TGA CTG CGA CTC CGT GCT TTC GCA CCC CTC GTT TGT CCT AAT CTA TGG GA
R20291_16S_18	CCA TCA GGT GCG ACA TTT GCT ACT CAT GAT CCA CAG CCC CCA ATG GGG GA
R20291_16S_19	AGC CAC GGC GTC GAT TGC GTA ATT CAT GAG GCG GAC CCT TCA TGC GAG CG
R20291_16S_20	TTC TCA CTT TGA GTT TCC TTA ACT GCC CCT GGG CGT GTT CAT CGC CTC GT
R20291_16S_21	ACA CCA AAT TAA GCT TCG TTG CGC TTC TTG GAA TGG ATT CGA ACT GTA GG
R20291_16S_22	ATT ACT GTA GAG GAA TTA GCC TCT CAA GGG AAG CCC CTG TAA CCA CTG TC
R20291_16S_23	CAC CAC GTA CCA ACA GCA GTC GAG CAC AGC ACT CTA CAA CCC AAT TCA GG
R20291_16S_24	GCG TTG CTC GCG TTG GGA ACA GAA ATC AAC GGT AGT AAT TCA ACC CGT GA
R20291_16S_25	GAT CTC TCT GAC GGT CCC TAT TGG ACC TCC TTC CAC CCC TAC TGC AGT TT
R20291_16S_26	AGT AGT ACG GGG AAT ACG AAT CCC GAT GTG TGC ACG ATG TTA CCC ATC AT

R20291_16S_27	GTC TCC CAA CGG TTC GGC ATT CCA CCT CGA TTA GGG AAT TTC GAT GAG AG
R20291_16S_28	TCA AGC CTA ACA TCC GAC TTT GAG CGG ATG TAC TTC GAC CTC AAT GAT CA
R20291_16S_29	TTA GCG TCT AGT CTT ACG ACG CCA CTT ACG CAA GGG CCC AGA ACA TGT GT
R20291_16S_30	GGC GGG CAG TGT GGT GCC CTC AAC CTC TGC GGG CTT CGG CTA ATA GAT TG
R20291_16S_31	GAA AAC CTT CTT CAG CAG CTT CCA CCT TAG TTA TTG ACC CCA CTT CAG CA
R20291_16S_32	TTG TTC CAT CGG CAT AGC CTT CCA CGC CGA CCT AGT GGA GGA AAG ATT CC
R20291_16S_33	TCT TAA TGG ATG ACA AAT TAA AAC T
R20291_23S_1	ACC AGT TCA ATA ATT CCC ACG TCC CGC CTA CGG AAC CGT GAT CCT CGG CT
R20291_23S_2	ACT TCC TGC ACT ATT CGA CGC TAT TCG AAG CCC CTC AAC GTG CAT TTG AA
R20291_23S_3	ACT AGG CTT CTA AAG GCT TAC TCC TTT GAG TGA ATC TCA TTA CAG ATT CA
R20291_23S_4	TAG TAA TTC ACT TAT GTA TCG AAT TAC TCC CCT TGA GTC CCT TGA CTT TG
R20291_23S_5	TAG ATT CAT GGA CTT CCT TCT CTT TCT TTA AGC TAA GGC ATT CAT CGC CG
R20291_23S_6	CTC GCT TGC GCC TAA TCG GGT TTG GTT ATT TCA AAA GAA ATA ACC CCA AC
R20291_23S_7	GCC TGT ATA GTA TTG CTT CTC CAT AGC ATT AAC TTC TCC AAA CCT TTC TG
R20291_23S_8	GGT GGT GTC TTC CAT TAT CAG GAC ATA CAA TTT GCT CTT CTG AAG TCT AT
R20291_23S_9	ACT AGG TCT CAT GGT GCC CTG TGC ACT TTG GGA CAC CCT TCG TCC TCC CT
R20291_23S_10	GGT GGG AGG TTC CGA TTT ATG ATG GAT CAC TGG CTA TCG CAT ATC ATG GC
R20291_23S_11	ACT CCC TTT CCA CTT TTC TTG GGG CCC TCC CCT CAC TTT ATC TTG GAC TT
R20291_23S_12	TGG GAC GTG AAT GTT CGA CAC CTT CGT GTA AAG AAC ACA CTG GCG CAT GA
R20291_23S_13	AAA ACA TCT TGC CCG GTT GCT CAA TGC AAT TCA TCG TTC CAA TTC GTG AA
R20291_23S_14	TTC CAC GCC TCG GCA TCG CTT TCG CTC AAA ATT GAC TCG CAA GTC AAT GA
R20291_23S_15	ACT GCA TCT GGG CTT TGG CCC GCT GGA TGG GTA CTC GTC CTA CTT CGC TT
R20291_23S_16	TCA TTT TAA AGC ACC TCC AGG CTT GGG TGC TCG CAA CTT TTC GAG CCC CT
R20291_23S_17	ACT GAA CAC CCA TCG CCA CTT TAA GGT TAG CTC GGG CCT CTA TCG ACC AA
R20291_23S_18	GAG GGG CTT TAT CGA AAT CCC GAT CGG AGT TCC ACT CTC TAT GCC TCC AT
R20291_23S_19	CTC GTG ACT TAC AGG ATC CCC CAT AAC GTG GAT GGC TTC TGA TAG TTT GA
R20291_23S_20	GGC TTA CGG CAG TAG AAT ATG AAC CCT CAG TCT GAC ACC CAC TAT TCT AA
R20291_23S_21	GTA TCA GCT TTC CCG TTG TCG GGT CTA GCA GTC GAT TCC AGG GAT TTA CA
R20291_23S_22	TTC AAT TCA CCA TTT CCT ACA CCC TAA CGT GTC TGT TGG TCC TAC AAC CG
R20291_23S_23	AAT CTT CGT CGG TGA GTA AGT TTC TCA CGC ATT ATC GAG TGA CCA GCT CA
R20291_23S_24	CTA GGA CGC GGC TTC TAA AGG CCC CGA TTT TGA ATG ATG GCT TCG ATG CC
R20291_23S_25	GTA GTC ATT ACT ACC CAT CCC CTC GAA GGG TAT GCC CAA CTT CGT ACT GG

R20291_23S_26	CAT TCC TGT ACA CCT GTC ATA CCC TCA CTC TTA CAA CCG TAC TCA TCG CT
R20291_23S_27	CTA CAC CCA CTC TTA GGG TGT CCG GCA TTT GGG TTC CAA AGG TCC CCT TC
R20291_23S_28	CAA GCA GGC GGG ACC CAA TCA GCC CTG GAT TCG ACT CCG GCT TTC CGC AT
R20291_23S_29	CCA CTA CCT GTT GTC CAA CTA TAA GGA CAT GAT GGC TAT TGG CAA ACT CT
R20291_23S_30	CTT CCC TAC TGT GTC ATC CTA TTC GAT TCG TGT GAC AAC CAA TAC ACA CG
R20291_23S_31	GGT TCG TAA CTC CGT CAG TTT CAT CCG TTT AGG CGA AAC TAT TAC GAC CC
R20291_23S_32	TAC ACT ACC CCT CGC TTT AAA TCA TCG CTT CAT CGA CTA AAG TGT GAC AG
R20291_23S_33	TTC TTT TCA GAG ATA GCT CCA ATT TCC ATG GGC ATG GCG TTT GGC TGT GT
R20291_23S_34	CCA CCC ACT CCT CTC ATA GGA TTC CGG TCG CTC TCT TGA CAA CAA TTC CT
R20291_23S_35	TGA GCC GTT TTA CTG GGG CAT TGA ATC CCT ATT CCC CAC GGT GGT AGT CC
R20291_23S_36	ACC GGC GTC TCT TAT CCG GGT TCG CTG ACA AAT GGT TTT TGT ATC CAA AG
R20291_23S_37	ACG ATT CAG CGT TCT GCT ACA TAT CCT CGA CTG CGG ACG GGC CAC GAC CT
R20291_23S_38	TCC AAT TCC CCT AGA CAA TCT CGT TAG CTT CGT CAC TTG AAT TCG GGG TC
R20291_23S_39	ATT TGC CGC CGG CAT TGA TAT TGC CAG GAT TCC ATC GCT TTA AGG AAC AG
R20291_23S_40	CCC ATT CAA GGC TGG GCG TGC TTT CCG CAT TGC TAA ACC CGT GAC AGA GT
R20291_23S_41	TGT TGT CTG AGC CAC TTT AAC ATT AAG GCC ACT TCT ACG GCC TAT GGA CG
R20291_23S_42	CTG TCC TGC CTT TCT GGG GTA CCT CGA AAT GAC ATC GAA CTG TAA CCC AG
R20291_23S_43	AAC CAT GAT GTA CAT GTC CTA TCC ACC CTC CGA AAC TTT GGT CCT GCG GT
R20291_23S_44	CAA AAC CGC CTC GGT AGG AAC CCT ATG GTG GGA ACA TCA TGA CCC TGA GA
R20291_23S_45	TTG GTA TCC GGT ACT TAG ACC AGA ACC CTG TGA CAG TCC ACC CGT CAA AC
R20291_23S_46	TGA CCC CGC CAG CGG AGG GTT TTC CAT TGC CTC CGC GAG TTT CCA AGA GA
R20291_23S_47	GTC ATG CCA GCC TTT AGC ATG CAT CTC ACA TTT CCG TTT TCT CTC GAA CT
R20291_23S_48	AAC GTT CTG TAT GTC CAG CTC GTT CCT ACT TTT AGC CTG AAT CAC TAG GC
R20291_23S_49	CAC CAA GAC GCA CCT TCC CGG TAG CGA GTT GCC TAT TTT CGA TGG GAC CC
R20291_23S_50	CTA TTG TCC GAA TAG AGG GGG TTC TCA GGT GTA GCT GCC CCT CCA AAC CG
R20291_23S_51	TGG AGC TAC AGC CGA GTA GTG TAG GAC CCC GAC ATC ATC CAG GGT TCC CA
R20291_23S_52	ACC CGA CAA GCG GGT AAT TTC ACC ATG CGC TCG ACC CAA GTC TTG CAG CA
R20291_23S_53	CTC TGT CAA GCC AGG GAT AGG CAG CGT CCG CAT CCT TTA AAC TCT TCT GG
R20291_23S_54	ACA GGA ATC ATG CTC TCC TGG CCC TAC CTG CAT GGA GAC CAC ATG GTC AA
R20291_23S_55	CAG GAC GGT TCC CGT ACC GAC CCA TCG ATA CAT GCC TTA CCT ATT CGC GA
R20291_23S_56	CTT TCG TAG ATT CGC GCT TCG GTT GAA GTT CTA TTC TAA AGG GTG GCG TT
R20291_23S_57	CCC ATT CTG GGG TCT TTC TGA TAG ACC AAC TAT CCA GCT TCC ACA TTC AC

R20291_23S_58	GTC GTT ACA TAA ATC GAA TAG CTA TGA TTA TCC AGC TCC TGA ACT GGT T

Appendix table 2. FastQC reports for the FASTQ files from sequenced samples in the small-scale pilot study.

✓ indicates a pass, ×, a failure and ⚠, a warning.

Sample	Per base seq qual.	Per base qual. scores	Per base seq cont.	Per base GC cont.	Per seq GC cont.	Per base N cont.	Seq length distrib.	Seq duplic. levels	Over-represented seq	Kmer cont.
3 h 1	✓	✓	×	×	×	✓	✓	×	✓	×
3 h 2	✓	✓	×	×	×	✓	✓	×	⚠	⚠
3 h 3	✓	✓	×	×	×	✓	✓	×	✓	×
24 h 1	✓	✓	×	×	×	✓	✓	×	✓	×
24 h 2	✓	✓	×	×	×	✓	✓	×	✓	×
24 h 3	✓	✓	×	×	×	✓	✓	×	✓	×

Appendix table 3. FastQC reports for the FASTQ files from the large-scale dual RNA-seq sequenced samples.

✓ indicates a pass, ×, a failure and ⚠, a warning.

Sample	Per base seq qual.	Per base qual. scores	Per base seq cont.	Per base GC cont.	Per seq GC cont.	Per base N cont.	Seq length distrib.	Seq duplic. levels	Over-repre-sented seq	Kmer cont.
Bacterial control 1	✓	✓	×	×	×	✓	⚠	×	×	×

Bacterial control 2	✓	✓	×	×	×	✓	⚠	×	×	×
Bacterial control 3	✓	✓	×	×	×	✓	⚠	×	×	×
Human control 3 h 1	✓	✓	✓	×	×	✓	⚠	×	⚠	×
Human control 3 h 2	✓	✓	✓	×	×	✓	⚠	×	⚠	⚠
Human control 3 h 3	✓	✓	✓	×	×	✓	⚠	×	⚠	⚠
Human control 6 h 1	✓	✓	✓	×	×	✓	⚠	×	⚠	×
Human control 6 h 2	✓	✓	✓	×	×	✓	⚠	×	⚠	⚠
Human control 6 h 3	✓	✓	✓	×	×	✓	⚠	×	⚠	×
Human control 12 h 1	✓	✓	✓	×	×	✓	⚠	×	⚠	×
Human control 12 h 2	✓	✓	✓	×	×	✓	⚠	×	⚠	⚠

Human control 12 h 3	✓	✓	✓	×	×	✓	⚠	×	⚠	×
Human control 24 h 1	✓	✓	✓	×	×	✓	⚠	×	⚠	⚠
Human control 24 h 2	✓	✓	✓	×	⚠	✓	⚠	×	⚠	⚠
Human control 24 h 3	✓	✓	✓	×	×	✓	⚠	×	⚠	⚠
3 h infected 1	✓	✓	×	×	×	✓	⚠	×	✓	×
3 h infected 2	✓	✓	×	×	×	✓	⚠	×	✓	×
3 h infected 3	✓	✓	×	×	×	✓	⚠	×	✓	×
6 h infected 1	✓	✓	✓	×	×	✓	⚠	×	⚠	×
6 h infected 2	✓	✓	×	×	×	✓	⚠	×	✓	×
6 h infected 3	✓	✓	×	×	×	✓	⚠	×	✓	×

12 h infected 1	✓	✓	✓	×	×	✓	⚠	×	⚠	×
12 h infected 2	✓	✓	×	×	×	✓	⚠	×	✓	×
12 h infected 3	✓	✓	×	×	×	✓	⚠	×	✓	×
24 h infected 1	✓	✓	×	×	×	✓	⚠	×	✓	×
24 h infected 2	✓	✓	×	×	×	✓	⚠	×	✓	×
24 h infected 3	✓	✓	×	×	×	✓	⚠	×	✓	×

Appendix table 4. Bacterial differentially expressed genes between uninfected controls and infected samples at 3 h post infection.

Gene ID	Gene name	log2(fold change)	padj	annotation
CDR20291_1556	gcvPB	-5.359123293	8.03E-31	glycine cleavage system P protein
CDR20291_2167	-	4.128957729	3.97E-30	PTS system, IIbc component
CDR20291_1591	-	-4.409902515	8.70E-28	putative dinitrogenase iron-molybdenum cofactor
CDR20291_2427	nox	3.643442669	7.01E-26	NADH oxidase
CDR20291_2024	-	3.847977463	3.33E-22	thioredoxin reductase
CDR20291_0118	-	3.311560371	1.29E-19	phosphoglucomutase/ phosphomannomutase mutase
CDR20291_1395	-	-4.119988092	6.97E-19	putative hemolysin-like membrane protein
CDR20291_0783	oppB	-3.010419133	3.07E-17	oligopeptide ABC transporter, permease protein
CDR20291_1555	-	-3.413469564	9.96E-17	putative bi-functional glycine dehydrogenase/ aminomethyl transferase protein

CDR20291_1329	-	-3.286505721	3.42E-15	hypothetical protein
CDR20291_0366	hadA	-3.068045974	4.44E-15	isocaprenoyl-CoA:2-hydroxyisocaproate CoA-transferase
CDR20291_0177	-	2.508809579	4.68E-15	putative oxidoreductase, NAD/FAD binding subunit
CDR20291_0689	-	3.156886318	2.87E-14	putative ATP-dependent RNA helicase
CDR20291_2340	-	-3.023716342	3.21E-14	histidine triad nucleotide-binding protein
CDR20291_2266	-	-2.731762298	5.97E-14	butyrate kinase
CDR20291_0787	oppF	-3.35353094	9.45E-14	oligopeptide ABC transporter, ATP-binding protein
CDR20291_2554	crr	-3.279884682	1.58E-13	PTS system, glucose-specific IIa component
CDR20291_2087	-	2.823754048	1.80E-13	putative aromatic compounds hydrolase
CDR20291_1643	-	2.633463606	3.44E-13	putative 2-hydroxyacyl-CoA dehydratase
CDR20291_0892	-	-2.966770658	5.66E-13	cell surface protein (putative N-acetylmuramoyl-L-alanine amidase) / cwp17
CDR20291_1180	-	-2.408195853	6.15E-13	aspartate aminotransferase
CDR20291_2686	-	2.751670081	1.39E-12	hypothetical protein
CDR20291_0785	oppA	-3.94720158	1.80E-12	oligopeptide ABC transporter, substrate-binding lipoprotein
CDR20291_0187	pyrD	3.028804931	1.95E-12	dihydroorotate dehydrogenase, catalytic subunit
CDR20291_1778	-	2.639700994	5.43E-12	hypothetical protein
CDR20291_0119	glmS	2.451142313	1.17E-11	glucosamine--fructose-6-phosphate aminotransferase [isomerizing]
CDR20291_0492	-	2.829061462	2.95E-11	hypothetical protein
CDR20291_0913	crt2	-2.948442064	4.04E-11	3-hydroxybutyryl-CoA dehydratase
CDR20291_2022	-	-3.003046209	4.40E-11	putative copper-transporting P-type ATPase
CDR20291_0175	-	2.851505729	9.51E-11	putative oxidoreductase, acetyl-CoA synthase subunit
CDR20291_0786	oppD	-3.041147449	1.19E-10	oligopeptide ABC transporter, ATP-binding protein
CDR20291_3217	pfkA	2.516041491	1.30E-10	6-phosphofructokinase
CDR20291_0470	-	-2.217286135	3.46E-10	putative lipoprotein
CDR20291_2624	-	-2.160239678	3.88E-10	cell surface protein / cwp14
CDR20291_1328	feoB1	-3.684294469	4.43E-10	ferrous iron transport protein B
CDR20291_0060	rpoB	2.158829931	5.44E-10	DNA-directed RNA polymerase beta chain
CDR20291_2555		-3.250426404	5.65E-10	
CDR20291_0915	thlA1	-3.220896835	6.52E-10	acetyl-CoA acetyltransferase
CDR20291_0912	etfA2	-2.885685471	8.43E-10	electron transfer flavoprotein alpha-subunit
CDR20291_0180	gluD	-2.522943712	1.08E-09	NAD-specific glutamate dehydrogenase
CDR20291_2685	-	2.254384983	1.09E-09	cell surface protein / cwp10
CDR20291_0516	-	-2.857960841	1.14E-09	putative cation transporting ATPase
CDR20291_0655	-	-1.883213105	1.14E-09	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, beta subunit

CDR20291_0122	murA	2.314779483	1.74E-09	UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1
CDR20291_1182	exoA	3.078645029	2.00E-09	putative exodeoxyribonuclease
CDR20291_3334	-	2.350079154	2.50E-09	putative bifunctional protein [include tetrapyrrole (Corrin/Porphyrin) methylase and nucleoside triphosphate pyrophosphohydrolase]
CDR20291_1639	-	-2.902948947	3.98E-09	putative ferrous iron transport protein A
CDR20291_1718	-	2.594833709	3.98E-09	putative cations-transporting ATPase
CDR20291_1876	-	2.623140914	3.98E-09	hypothetical protein
CDR20291_3527	-	2.459713215	4.17E-09	hypothetical protein
CDR20291_2676	cwp84	-2.2510842	4.30E-09	cell surface protein (putative cell surface-associated cysteine protease)
CDR20291_3135	feoB3	-2.336502284	6.06E-09	putative ferrous iron transport protein B
CDR20291_0368	hadB	-2.887839998	7.49E-09	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase
CDR20291_0176	-	2.894093184	9.14E-09	putative oxidoreductase, electron transfer subunit
CDR20291_1799	-	-1.774385548	1.74E-08	hypothetical protein
CDR20291_0507	gapN	2.002531855	1.90E-08	NADP-dependent glyceraldehyde-3-phosphate dehydrogenase
CDR20291_1719	cysD	-2.902847841	1.97E-08	putative O-acetylhomoserine sulfhydrylase
CDR20291_0184	-	2.055757695	3.17E-08	putative cell wall hydrolase
CDR20291_0372	etfA1	-2.832082634	3.97E-08	electron transfer flavoprotein alpha-subunit
CDR20291_0371	etfB1	-2.466751936	5.69E-08	electron transfer flavoprotein beta-subunit
CDR20291_1373	-	2.159253755	7.18E-08	putative rubrerythrin
CDR20291_0318	-	-3.069530828	8.57E-08	putative heavy-metal-transporting ATPase
CDR20291_0490	nth	1.991950558	1.28E-07	endonuclease iii
CDR20291_3147	-	2.09875593	1.39E-07	putative sodium-dependent phosphate transporter
CDR20291_0354	-	2.333177247	1.55E-07	hypothetical protein
CDR20291_1318	-	-2.134113226	1.94E-07	cell surface protein (putative penicillin-binding protein) / cwp20
CDR20291_0367	hadI	-2.362720281	2.46E-07	activator of 2-hydroxyisocaproyl-CoA dehydratase
CDR20291_2558	appF	2.189572956	2.63E-07	oligopeptide ABC transporter, ATP-binding protein
CDR20291_3031	cggR	2.683708836	2.63E-07	central glycolytic genes regulator
CDR20291_0822	cspA	2.68247838	3.31E-07	cold shock protein
CDR20291_1975	-	1.940764248	3.49E-07	ABC transporter, ATP-binding protein
CDR20291_2678	cwp66	-1.971089192	3.64E-07	cell surface protein
CDR20291_0815	glgP	-2.039677154	4.63E-07	glycogen phosphorylase
CDR20291_1038	-	2.543420341	5.29E-07	hypothetical protein
CDR20291_0914	hbd	-2.477116322	5.59E-07	3-hydroxybutyryl-CoA dehydrogenase
CDR20291_0005	dnaH	2.235041415	6.57E-07	DNA polymerase III subunit gamma/tau
CDR20291_2655	-	-1.771779015	6.67E-07	cell surface protein / cwp19

CDR20291_0015	clpC	-1.998600307	7.00E-07	ATP-dependent Clp protease
CDR20291_1104	-	-2.455828313	7.64E-07	hypothetical protein
CDR20291_3306	atpA	1.540508968	8.28E-07	ATP synthase alpha chain
CDR20291_3308	atpF	1.843152121	8.28E-07	ATP synthase B chain
CDR20291_0014	-	-2.044918396	8.55E-07	putative ATP:guanido phosphotransferase
CDR20291_3168	-	1.638601997	8.55E-07	putative phosphoesterase
CDR20291_0065	-	-2.010073239	9.09E-07	elongation factor TU
CDR20291_0911	etfB2	-2.208393896	1.13E-06	electron transfer flavoprotein beta-subunit
CDR20291_2369	-	2.338293838	1.17E-06	hypothetical protein
CDR20291_3234	uvrB	1.934558333	1.23E-06	excinuclease ABC subunit B
CDR20291_3393	-	1.877527021	1.23E-06	hypothetical protein
CDR20291_1644	hgdC	1.958747762	1.35E-06	putative (R)-2-hydroxyglutaryl-CoA dehydratase alpha-subunit
CDR20291_0329	cbiM	-2.339526559	1.36E-06	putative cobalt transport protein
CDR20291_0958	-	2.21877211	1.42E-06	AraC-family transcriptional regulator
CDR20291_2744	dltA	2.195886798	1.42E-06	D-alanine--poly(phosphoribitol) ligase subunit 1 (D-alanine-activating enzyme)
CDR20291_3390	greA	2.336992018	1.42E-06	transcription elongation factor grea
CDR20291_0595		2.081960744	1.47E-06	
CDR20291_0106	nrdD	2.077734783	1.56E-06	anaerobic ribonucleoside-triphosphate reductase
CDR20291_0053	secE	2.221628512	1.86E-06	preprotein translocase SecE subunit
CDR20291_0217	purF	1.983692085	2.09E-06	amidophosphoribosyltransferase
CDR20291_0445	-	-2.090408058	3.13E-06	hypothetical protein
CDR20291_1963	-	-1.656718551	4.38E-06	hypothetical protein
CDR20291_1386	aspB	-1.754033861	5.12E-06	putative glutamate synthase [NADPH] small chain
CDR20291_3352	gcaD	1.816074786	9.02E-06	bifunctional protein
CDR20291_3409	-	2.310564155	9.02E-06	putative preprotein translocase
CDR20291_1493	cysA	-1.595892014	9.34E-06	serine acetyltransferase
CDR20291_3444	-	4.694570844	9.67E-06	ferredoxin
CDR20291_3392	-	1.423027701	1.13E-05	hypothetical protein
CDR20291_0711	-	-1.658351745	1.15E-05	cytidine/deoxycytidylate deaminase family protein
CDR20291_0944	-	-2.038713266	1.38E-05	hypothetical protein
CDR20291_3167	tig	1.287777081	1.56E-05	trigger factor
CDR20291_1693	-	1.812072722	1.61E-05	LysR-family transcriptional regulator
CDR20291_0355	-	1.994497573	1.89E-05	hypothetical protein
CDR20291_0097	rplQ	-1.337115544	1.97E-05	50S ribosomal protein L17
CDR20291_1157	rpsO	2.51246877	2.13E-05	30S ribosomal protein S15
CDR20291_0084	rpsE	-1.611075968	2.94E-05	30S ribosomal protein S5
CDR20291_3524	-	-1.528221335	3.49E-05	putative amino acid aminotransferase
CDR20291_0370	acdB	-2.123269291	3.88E-05	acyl-CoA dehydrogenase, short-chain specific

CDR20291_1100	-	-1.781289023	3.97E-05	branched chain amino acid transport system carrier protein
CDR20291_0982	mreB2	1.222086593	4.17E-05	rod shape-determining protein
CDR20291_0003	serS1	-2.040058776	4.49E-05	seryl-tRNA synthetase
CDR20291_0081	rpsH	-1.55257646	4.49E-05	30S ribosomal protein S8
CDR20291_0083	rplR	-1.571648278	4.49E-05	50S ribosomal protein L18
CDR20291_0369	hadC	-2.2697263	4.98E-05	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase
CDR20291_2733	-	-1.40616266	4.98E-05	nitrilase (carbon-nitrogen hydrolase)
CDR20291_0685	-	1.61689079	5.50E-05	putative signaling protein
CDR20291_3028	tpi	1.569727676	5.52E-05	triosephosphate isomerase
CDR20291_2708	-	-1.95893583	5.66E-05	putative amino acid racemase
CDR20291_0277	htpG	1.566559666	6.13E-05	chaperone protein (heat shock protein)
CDR20291_3358	-	1.344434074	6.16E-05	putative peptidase
CDR20291_0086	rplO	-1.855836668	6.64E-05	50S ribosomal protein L15
CDR20291_1013	-	1.832110364	6.65E-05	hypothetical protein
CDR20291_0102	-	1.30508751	6.76E-05	50S ribosomal protein L13 (pseudogene)
CDR20291_1645	-	-1.622930137	7.16E-05	cell surface protein (putative cell surface-associated cysteine protease) / cwp13
CDR20291_2731	serS2	-1.864347344	7.16E-05	seryl-tRNA synthetase
CDR20291_3528	-	1.930465277	7.16E-05	putative transcriptional regulator
CDR20291_2116	-	1.626575624	7.67E-05	putative GTP-binding protein
CDR20291_3355	murC	1.498130073	7.67E-05	UDP-N-acetylmuramate--L-alanine ligase
CDR20291_0051	-	-1.735414223	8.43E-05	elongation factor TU
CDR20291_2252	-	-1.609710203	8.74E-05	putative sulfonate ABC transporter, solute-binding lipoprotein
CDR20291_2055	-	1.70038796	9.33E-05	hypothetical protein
CDR20291_1327	feoA1	-1.988268098	9.35E-05	putative ferrous iron transport protein A
CDR20291_0643	cooS	-1.491130356	9.72E-05	putative bifunctional carbon monoxide dehydrogenase/ acetyl-CoA synthase
CDR20291_3530	-	-1.428660699	0.000102314	putative selenocysteine lyase
CDR20291_2353	dnaJ	-1.794403607	0.000129941	chaperone protein
CDR20291_0687	plfB	1.584160292	0.000136126	formate acetyltransferase
CDR20291_3299	alr	-1.399982415	0.000141282	alanine racemase
CDR20291_0721	-	-1.64523749	0.000143473	putative NUDIX-family hydrolase
CDR20291_0082	rplF	-1.518358153	0.000143931	50S ribosomal protein L6
CDR20291_1022	fabF	-1.104835237	0.000144949	3-oxoacyl-[acyl-carrier-protein] synthase II
CDR20291_1772	-	-1.961159968	0.000172274	hypothetical protein
CDR20291_3169	rph	1.485032012	0.000177067	ribonuclease Ph
CDR20291_2579	hpt	2.05612762	0.000177723	putative phosphoribosyltransferase
CDR20291_2520	-	1.424416163	0.000182149	putative GTP-binding protein
CDR20291_0611	infC	1.73452823	0.000186085	translation initiation factor IF-3

CDR20291_1264	-	-1.829670866	0.000205072	putative ATP-binding protein
CDR20291_0079	rplE	-1.211890243	0.000233674	50S ribosomal protein L5
CDR20291_1250	-	-1.221354277	0.000358459	putative oligopeptide transporter
CDR20291_2571	-	-1.341201267	0.000376798	putative propanediol utilization protein
CDR20291_3397	-	1.638801645	0.000380828	putative ATPase
CDR20291_0983	mreC	1.453210508	0.000382057	putative rod shape-determining protein precursor
CDR20291_3320	prfA	1.228570276	0.000400463	peptide chain release factor 1
CDR20291_2854	-	-1.14743799	0.000434187	hypothetical protein
CDR20291_3402	-	1.203227146	0.000463843	GntR-family transcriptional regulator
CDR20291_0515	-	-1.111376311	0.00049697	hypothetical protein
CDR20291_1053	-	1.557737854	0.00049697	putative pyrophosphokinase
CDR20291_0945	-	-1.798206378	0.000501301	putative peptidase
CDR20291_2354	dnaK	-1.47075616	0.000573258	chaperone protein
CDR20291_2030	-	1.475332173	0.000589356	putative multiprotein-complex assembly protein
CDR20291_0054	nusG	1.432384614	0.000629743	transcription antitermination protein
CDR20291_0218	purG	1.799152411	0.000749925	phosphoribosylformylglycinamidine cyclo-ligase
CDR20291_2127	-	1.63582388	0.000749925	hypothetical protein
CDR20291_0039	-	-2.238030746	0.000801093	putative dual-specificity prolyl/cysteinyl-tRNA synthetase
CDR20291_0088	adk	-1.39950887	0.000846663	adenylate kinase
CDR20291_3307	atpH	1.77560131	0.000915658	ATP synthase subunit delta
CDR20291_0910	bcd2	-1.433577015	0.000955591	butyryl-CoA dehydrogenase
CDR20291_2721	-	-1.798128049	0.000955591	hypothetical protein / zmp1
CDR20291_1617	-	-1.482416367	0.001098286	putative hydantoinase
CDR20291_3048	-	-1.218114933	0.001216569	cell surface protein / cwp21
CDR20291_0096	rpoA	-1.330344407	0.001265931	DNA-directed RNA polymerase alpha chain
CDR20291_0330	cbiN	-1.744455905	0.001375516	cobalt transport protein
CDR20291_0663	-	-1.18646333	0.001390342	hypothetical protein
CDR20291_0725	nadE	1.154964374	0.001430637	NH3-dependent NAD(+) synthetase
CDR20291_0089	map1	-1.32846976	0.001575005	methionine aminopeptidase
CDR20291_2649	-	1.355678081	0.001609268	putative N-acetylmuramoyl-L-alanine amidase
CDR20291_2464	stk	-1.159805136	0.001624484	serine/threonine-protein kinase and phosphatase
CDR20291_3541	rpmH	3.220009523	0.001640966	50S ribosomal protein L34
CDR20291_2410	leuS	-1.577691594	0.001663848	leucyl-tRNA synthetase
CDR20291_1854	dinR	-1.245451206	0.001789698	SOS regulatory protein
CDR20291_1534	-	1.52110168	0.001859338	tellurium resistance protein
CDR20291_1066	-	1.160908416	0.001872737	putative protease
CDR20291_0087	prlA	-1.733729249	0.001880232	preprotein translocase SecY subunit
CDR20291_2681	secA2	1.264685018	0.002106482	preprotein translocase SecA subunit
CDR20291_0221	purD	1.334365462	0.002469488	phosphoribosylamine--glycine ligase

CDR20291_2072	msrAB	1.389641272	0.002563891	putative peptide methionine sulfoxide reductase
CDR20291_2774	fhuD	1.380154843	0.002563891	putative ferrichrome ABC transporter, substrate-binding protein
CDR20291_0662	-	-1.163113119	0.002734608	hypothetical protein
CDR20291_1101	-	1.171952405	0.002741746	putative ribonucleotide-diphosphate reductase
CDR20291_2874	-	1.422773654	0.003004711	hypothetical protein
CDR20291_0085	rpmD	-1.151328946	0.00301435	50S ribosomal protein L30
CDR20291_0847	-	-1.582798355	0.00301435	putative aminotransferase
CDR20291_1161	-	1.38117573	0.00301435	putative peptidase
CDR20291_1859	accA	-1.347080486	0.00301435	acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha
CDR20291_1955	-	-1.137884502	0.00301435	RpiR-family transcriptional regulator
CDR20291_2211	tkt	1.210269286	0.00301435	transketolase
CDR20291_2609	-	-1.231731722	0.003519387	hypothetical protein
CDR20291_0903	-	-1.277616524	0.003842915	cell surface protein / cwp18
CDR20291_3029	pgK	1.305604878	0.003869586	phosphoglycerate kinase
CDR20291_0818	speA	1.501156967	0.004043695	arginine decarboxylase
CDR20291_3231	hprK	1.196240991	0.004177307	HPr(Ser) kinase/phosphorylase
CDR20291_2561	appB	-1.425820752	0.00444158	oligopeptide ABC transporter, permease protein
CDR20291_0057	rplJ	1.099018719	0.004509592	50S ribosomal protein L10
CDR20291_2437	-	1.462306246	0.004644596	putative sugar transporter, substrate-binding lipoprotein
CDR20291_2336	-	-1.770968035	0.004725388	putative sigma 54 modulation protein
CDR20291_0612	rpmI	1.200855693	0.004796788	50S ribosomal protein L35
CDR20291_2478	ltaE	-1.166477143	0.004877573	low-specificity L-threonine aldolase
CDR20291_2668	pgm2	1.169162836	0.005179375	putative phosphomannomutase/ phosphoglycerate mutase
CDR20291_1017	fabH	1.158437292	0.005612197	3-oxoacyl-[acyl-carrier-protein] synthase III
CDR20291_2034	-	-1.028577117	0.005647202	hypothetical protein
CDR20291_1933	clpB	-3.012089616	0.005831755	chaperone
CDR20291_1014	rpmF	1.255086455	0.005888284	50S ribosomal protein L32
CDR20291_2415	gbeA	1.150808232	0.006314955	glycogen branching enzyme
CDR20291_3225	-	-1.371934948	0.006361039	putative formate/nitrite transporter
CDR20291_0774	-	-1.021814397	0.006536241	cell surface protein / cwp25
CDR20291_0923	ccpA	1.094269084	0.006776629	LacI-family transcriptional regulator (catabolite control protein)
CDR20291_1119	iscS2	1.183762512	0.006776629	cysteine desulfurase
CDR20291_2199	cspD	3.023528729	0.006809195	cold shock protein
CDR20291_2126	-	1.098026151	0.007159017	putative radical SAM superfamily lipoprotein
CDR20291_1779	-	1.97702469	0.007244683	hypothetical protein
CDR20291_0219	purN	1.631944685	0.007304131	phosphoribosylglycinamide formyltransferase
CDR20291_3405	veg	-1.29763657	0.007341969	hypothetical protein

CDR20291_0107	nrdG	1.231621313	0.007667327	anaerobic ribonucleoside-triphosphate reductase activating protein
CDR20291_3011	smpB	1.439761062	0.007977575	SsrA-binding protein
CDR20291_1862	accB	-1.217379152	0.00851087	biotin carboxyl carrier protein of acetyl-CoA carboxylase
CDR20291_0760	-	-1.234836928	0.008598528	hypothetical protein
CDR20291_3529	-	-1.209550119	0.008792974	hypothetical protein
CDR20291_0811	-	-1.142218489	0.009030253	hypothetical protein
CDR20291_1197	cspB	-1.228622862	0.009153277	putative cold shock protein
CDR20291_2428	-	1.382449473	0.00919006	sodium:dicarboxylate symporter family protein
CDR20291_2302	-	-1.148944087	0.009376187	hypothetical protein
CDR20291_1323	-	1.438689612	0.009838324	putative ruberythrin
CDR20291_0656	gcvH	-1.422087449	0.009962902	putative glycine cleavage system H protein
CDR20291_0883	potD	-1.077479827	0.010930376	spermidine/putrescine ABC transporter, substrate-binding lipoprotein
CDR20291_1346	-	1.130670202	0.011039744	hypothetical protein
CDR20291_1675	-	1.160008135	0.011039744	putative methylase
CDR20291_2083	-	1.233099104	0.011899512	putative amino-acid ABC transporter, substrate-binding protein
CDR20291_3064	-	1.155244711	0.011911809	hypothetical protein
CDR20291_0078	rplX	-1.210279576	0.012126173	50S ribosomal protein L24
CDR20291_0345	-	-1.410355008	0.012387735	hypothetical protein
CDR20291_2866	-	1.294475174	0.012508341	transcription antiterminator
CDR20291_2373	-	-1.270297646	0.012614186	hypothetical protein
EBG00000018546		-6.383285544	0.012614186	
CDR20291_0332	cbiO	-1.229761873	0.012983482	cobalt ABC transporter, ATP-binding protein
CDR20291_2045	pyrH	1.154868605	0.01343985	uridylate kinase
CDR20291_2117	-	-1.006127148	0.013515	ABC transporter, ATP-binding/permease protein
EBG00000018535		-6.006096446	0.013823988	
EBG00000018538		-6.229824786	0.014191811	
CDR20291_3428	pyrAB2	1.100880879	0.015837182	carbamoyl-phosphate synthase, pyrimidine-specific, large chain
CDR20291_3461	-	1.281599206	0.015837182	chloramphenicol o-acetyltransferase
CDR20291_0978	rnfB	-1.169932128	0.015928154	electron transport complex protein
CDR20291_0141	-	-1.388954141	0.016281313	putative RNA-binding protein
CDR20291_3337	prsA	-1.133623738	0.016281313	putative foldase lipoprotein (late stage protein export lipoprotein)
CDR20291_2562	appC	-1.125205891	0.01643852	oligopeptide ABC transporter, permease protein
CDR20291_2329	era	1.125249953	0.016537727	GTP-binding protein
CDR20291_0633	-	-1.207974522	0.017026409	putative signaling protein
CDR20291_0851	-	1.139574795	0.017026409	putative radical SAM protein
CDR20291_1657	-	1.11609464	0.017391731	hypothetical protein
CDR20291_1120	-	1.167663454	0.017522383	NifU-like protein

CDR20291_2341	-	1.237742075	0.017522383	putative radical SAM superfamily protein
CDR20291_2301	ppdK	-1.228308884	0.018131111	pyruvate, phosphate dikinase
CDR20291_1861	accC	-1.17746547	0.018986945	biotin carboxylase (acetyl-CoA carboxylase subunit A)
CDR20291_2628	aspS	-1.079056797	0.019356597	putative aspartyl-tRNA synthetase
CDR20291_1860	accD	-1.055418676	0.020235331	acetyl-coenzyme A carboxylase carboxyl transferase subunit beta
CDR20291_2063	-	-1.226177577	0.020779315	hypothetical protein
CDR20291_0105	aspC	-1.388261653	0.021107353	aspartate aminotransferase
CDR20291_0849	-	-2.447280713	0.023771955	hypothetical protein
CDR20291_1537	-	-1.004688709	0.023771955	putative tellurite resistance protein
CDR20291_2556	licT	1.06578121	0.023771955	putative transcription antiterminator
CDR20291_2044	rrf	-1.038433248	0.024249347	ribosome recycling factor
CDR20291_1375	pyrC	1.08578776	0.024801356	dihydroorotase
CDR20291_3132	-	1.294823907	0.025698143	hypothetical protein
CDR20291_0802	-	-1.144541367	0.025703738	ABC transporter, substrate-binding lipoprotein
CDR20291_2023	bipA	1.199308598	0.025843113	GTP-binding protein
CDR20291_3233	uvrA	1.006954316	0.025843113	excinuclease ABC subunit A
CDR20291_2763	tldD	1.206529336	0.02717582	putative regulatory protease
CDR20291_0918	acpP	2.23555518	0.033076765	acyl carrier protein
EBG00000018541		-5.48381199	0.033076765	
EBG00000018522		-5.531209033	0.035662092	
EBG00000018531		-5.281288666	0.04636169	

Appendix table 5. Bacterial differentially expressed genes between uninfected controls and infected samples at 6 h post infection.

Gene ID	Gene name	log_2(fold change)	padj	annotation
CDR20291_2167	-	5.07117665	1.04E-47	PTS system, IIbc component
CDR20291_2024	-	4.93538419	1.87E-39	thioredoxin reductase
CDR20291_0507	gapN	3.46620557	4.40E-27	NADP-dependent glyceraldehyde-3-phosphate dehydrogenase
CDR20291_2410	leuS	-5.2396857	4.40E-27	leucyl-tRNA synthetase
CDR20291_2554	crr	-4.8160845	2.19E-26	PTS system, glucose-specific IIa component
CDR20291_2685	-	3.65656163	2.19E-26	cell surface protein / cwp10
CDR20291_2686	-	3.69100638	8.47E-26	hypothetical protein
CDR20291_0944	-	-4.570234	3.38E-23	hypothetical protein
CDR20291_2555		-5.1777539	4.65E-23	
CDR20291_0787	oppF	-4.4810509	1.81E-21	oligopeptide ABC transporter, ATP-binding protein
CDR20291_1719	cysD	-5.2654983	3.24E-21	putative O-acetylhomoserine sulfhydrylase

CDR20291_3217	pfkA	3.4607712	5.43E-21	6-phosphofructokinase
CDR20291_3028	tpi	3.33615124	1.38E-20	triosephosphate isomerase
CDR20291_0783	oppB	-3.7712385	4.84E-20	oligopeptide ABC transporter, permease protein
CDR20291_0445	-	-3.9494009	6.43E-20	hypothetical protein
CDR20291_2340	-	-4.0205561	1.97E-19	histidine triad nucleotide-binding protein
CDR20291_0802	-	-4.5458138	2.58E-19	ABC transporter, substrate-binding lipoprotein
CDR20291_3031	cggR	4.17488483	2.58E-19	central glycolytic genes regulator
CDR20291_0912	etfA2	-4.1082951	4.72E-19	electron transfer flavoprotein alpha-subunit
CDR20291_0119	glmS	3.07077797	5.92E-19	glucosamine--fructose-6-phosphate aminotransferase [isomerizing]
CDR20291_0913	crt2	-3.8320931	5.92E-19	3-hydroxybutyryl-CoA dehydratase
CDR20291_0785	oppA	-4.7747284	1.25E-18	oligopeptide ABC transporter, substrate-binding lipoprotein
CDR20291_0915	thlA1	-4.2957474	1.61E-18	acetyl-CoA acetyltransferase
CDR20291_0786	oppD	-4.1905417	2.87E-18	oligopeptide ABC transporter, ATP-binding protein
CDR20291_1556	gcvPB	-4.2501763	3.53E-18	glycine cleavage system P protein
CDR20291_2427	nox	3.06474734	5.37E-18	NADH oxidase
CDR20291_0118	-	3.13925391	5.91E-18	phosphoglucomutase/ phosphomannomutase mutase
CDR20291_0945	-	-4.1059689	6.34E-18	putative peptidase
CDR20291_2369	-	3.80593512	9.04E-18	hypothetical protein
CDR20291_0643	cooS	-3.457777	1.68E-17	putative bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase
CDR20291_2774	fhuD	3.55434953	1.79E-17	putative ferrichrome ABC transporter, substrate-binding protein
CDR20291_3168	-	2.71437086	7.87E-17	putative phosphoesterase
CDR20291_0329	cbiM	-3.9286995	3.22E-16	putative cobalt transport protein
CDR20291_0847	-	-4.4559431	3.22E-16	putative aminotransferase
CDR20291_0122	murA	3.0610199	5.95E-16	UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1
CDR20291_1100	-	-3.3524481	2.44E-15	branched chain amino acid transport system carrier protein
CDR20291_0914	hbd	-3.9812252	3.75E-15	3-hydroxybutyryl-CoA dehydrogenase
CDR20291_1182	exoA	3.7651705	7.32E-15	putative exodeoxyribonuclease
CDR20291_0760	-	-3.8418292	1.39E-14	hypothetical protein
CDR20291_3167	tig	2.07138462	3.55E-14	trigger factor
CDR20291_1318	-	-3.0116338	5.92E-14	cell surface protein (putative penicillin-binding protein) / cwp20
CDR20291_0217	purF	3.03254352	5.96E-14	amidophosphoribosyltransferase
CDR20291_0180	gluD	-3.1769047	6.58E-14	NAD-specific glutamate dehydrogenase
CDR20291_0106	nrdD	3.06271218	7.72E-14	anaerobic ribonucleoside-triphosphate reductase
CDR20291_0470	-	-2.8754271	1.03E-13	putative lipoprotein
CDR20291_3029	pgK	2.93122439	1.55E-13	phosphoglycerate kinase
CDR20291_0911	etfB2	-3.3676613	2.02E-13	electron transfer flavoprotein beta-subunit

CDR20291_2655	-	-2.5804372	2.47E-13	cell surface protein / cwp19
CDR20291_2676	cwp84	-2.8116807	2.56E-13	cell surface protein (putative cell surface-associated cysteine protease)
CDR20291_0195	groEL	2.02780351	2.64E-13	60 kDa chaperonin
CDR20291_0815	glgP	-3.0589506	3.97E-13	glycogen phosphorylase
CDR20291_0655	-	-2.4496541	5.49E-13	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, beta subunit
CDR20291_1493	cysA	-2.5079883	5.49E-13	serine acetyltransferase
CDR20291_1118	-	3.04667317	6.80E-13	RRF2-family transcriptional regulator
CDR20291_1119	iscS2	2.66042599	1.42E-12	cysteine desulfurase
CDR20291_0177	-	2.1855052	4.46E-12	putative oxidoreductase, NAD/FAD binding subunit
CDR20291_1180	-	-2.3093159	6.79E-12	aspartate aminotransferase
CDR20291_1250	-	-2.3264729	8.37E-12	putative oligopeptide transporter
CDR20291_0910	bcd2	-2.8699168	8.41E-12	butyryl-CoA dehydrogenase
CDR20291_2252	-	-2.7693603	1.43E-11	putative sulfonate ABC transporter, solute-binding lipoprotein
CDR20291_0354	-	3.0108489	1.57E-11	hypothetical protein
CDR20291_1395	-	-3.4279559	1.57E-11	putative hemolysin-like membrane protein
CDR20291_0492	-	2.88602198	1.96E-11	hypothetical protein
CDR20291_3306	atpA	1.97447644	2.29E-11	ATP synthase alpha chain
CDR20291_3169	rph	2.39797971	6.78E-11	ribonuclease Ph
CDR20291_2266	-	-2.7578631	7.16E-11	butyrate kinase
CDR20291_3218	dnaE	1.78615756	8.88E-11	DNA polymerase III alpha subunit
CDR20291_3334	-	2.53118495	9.82E-11	putative bifunctional protein [include tetrapyrrole (Corrin/Porphyrin) methylase and nucleoside triphosphate pyrophosphohydrolase]
CDR20291_0595		2.82853741	1.36E-10	
CDR20291_1120	-	2.6969413	1.37E-10	NifU-like protein
CDR20291_1963	-	-2.5471411	2.00E-10	hypothetical protein
CDR20291_1555	-	-2.7908637	2.77E-10	putative bi-functional glycine dehydrogenase/aminomethyl transferase protein
CDR20291_2510	-	-2.2774778	5.59E-10	cinserved hypothetical protein
CDR20291_0721	-	-2.8275319	7.60E-10	putative NUDIX-family hydrolase
CDR20291_3027	gpmI	2.64067162	1.21E-09	2,3-bisphosphoglycerate-independent phosphoglycerate mutase
CDR20291_1718	-	2.48631883	1.42E-09	putative cations-transporting ATPase
CDR20291_1876	-	2.76532561	1.67E-09	hypothetical protein
CDR20291_1017	fabH	2.34439815	1.85E-09	3-oxoacyl-[acyl-carrier-protein] synthase III
CDR20291_0187	pyrD	2.75189574	1.88E-09	dihydroorotate dehydrogenase, catalytic subunit
CDR20291_2874	-	2.7601258	3.10E-09	hypothetical protein
CDR20291_0515	-	-1.9609461	3.37E-09	hypothetical protein
CDR20291_0755	rbr	1.83980191	3.58E-09	rubrerythrin
CDR20291_3434	-	-2.8525376	3.92E-09	putative homocysteine S-methyltransferase

CDR20291_1104	-	-3.1812916	3.94E-09	hypothetical protein
CDR20291_1645	-	-2.50562	4.51E-09	cell surface protein (putative cell surface-associated cysteine protease) / cwp13
CDR20291_2624	-	-2.3989517	6.45E-09	cell surface protein / cwp14
CDR20291_0490	nth	2.29101185	6.76E-09	endonuclease iii
CDR20291_1308	-	-2.1400296	7.32E-09	putative 5-nitroimidazole reductase
CDR20291_2055	-	2.42416842	7.50E-09	hypothetical protein
CDR20291_1328	feoB1	-3.0092363	1.20E-08	ferrous iron transport protein B
CDR20291_3299	alr	-2.2221191	1.46E-08	alanine racemase
CDR20291_1778	-	2.02056849	1.55E-08	hypothetical protein
CDR20291_1639	-	-2.668911	2.21E-08	putative ferrous iron transport protein A
CDR20291_3405	veg	-2.5226379	2.41E-08	hypothetical protein
CDR20291_2087	-	2.34411807	2.45E-08	putative aromatic compounds hydrolase
CDR20291_1492	cysM	-1.8807664	2.83E-08	putative O-acetylserine sulfhydrylase
CDR20291_3234	uvrB	2.2884045	2.83E-08	excinuclease ABC subunit B
CDR20291_1386	aspB	-2.2039807	4.18E-08	putative glutamate synthase [NADPH] small chain
CDR20291_0175	-	2.22969421	7.68E-08	putative oxidoreductase, acetyl-CoA synthase subunit
CDR20291_0218	purG	2.69131549	1.07E-07	phosphoribosylformylglycinamidine cyclo-ligase
CDR20291_2152	kamA	-2.6368737	1.53E-07	L-lysine 2,3-aminomutase
CDR20291_0219	purN	2.99259224	2.01E-07	phosphoribosylglycinamide formyltransferase
CDR20291_2845	-	2.06873961	2.25E-07	putative pyridine-nucleotide-disulfide oxidoreductase
CDR20291_2571	-	-2.022286	2.36E-07	putative propanediol utilization protein
CDR20291_2826	-	2.1026211	2.93E-07	putative ABC transporter, permease protein
CDR20291_1018	fabK	1.69211015	3.62E-07	trans-2-enoyl-ACP reductase
CDR20291_1329	-	-2.0472153	3.80E-07	hypothetical protein
CDR20291_1799	-	-1.8078151	3.81E-07	hypothetical protein
CDR20291_0355	-	2.40223102	4.00E-07	hypothetical protein
CDR20291_3216	pykF	1.70396044	4.55E-07	pyruvate kinase
CDR20291_3030	gapB	2.51292148	4.59E-07	glyceraldehyde-3-phosphate dehydrogenase 2
CDR20291_2324	glyS	-1.6844621	4.76E-07	glycyl-tRNA synthetase beta chain
CDR20291_0107	nrdG	2.27525434	4.83E-07	anaerobic ribonucleoside-triphosphate reductase activating protein
CDR20291_1115	codY	-1.9091144	4.97E-07	GTP-sensing transcriptional pleiotropic repressor
CDR20291_0756	-	2.1457845	7.38E-07	putative oxidative stress regulatory protein
CDR20291_1373	-	1.81935975	7.38E-07	putative rubrerythrin
CDR20291_0221	purD	2.19629803	7.55E-07	phosphoribosylamine--glycine ligase
CDR20291_0689	-	2.1296313	8.68E-07	putative ATP-dependent RNA helicase
CDR20291_1127	fur	1.99562644	9.09E-07	ferric uptake regulation protein

CDR20291_0079	rplE	-1.6767546	9.62E-07	50S ribosomal protein L5
CDR20291_2708	-	-2.4962172	1.08E-06	putative amino acid racemase
CDR20291_0051	-	-2.2771019	1.15E-06	elongation factor TU
CDR20291_0757	rbo	2.05694057	1.16E-06	rubredoxin oxidoreductase (desulfoferrodoxin)
CDR20291_2678	cwp66	-1.9124937	1.23E-06	cell surface protein
CDR20291_0039	-	-2.8581205	1.66E-06	putative dual-specificity prolyl/cysteinyl-tRNA synthetase
CDR20291_2615	glyA	-1.6351585	1.91E-06	putative serine hydroxymethyltransferase
CDR20291_0645	fhs	-2.1237146	1.92E-06	formate--tetrahydrofolate ligase
CDR20291_0771	-	-1.984253	1.96E-06	putative formyltransferase
CDR20291_1545	-	2.11269405	2.13E-06	putative iron compound ABC transporter, permease protein
CDR20291_1197	cspB	-2.0125131	2.14E-06	putative cold shock protein
CDR20291_0685	-	1.9335544	2.27E-06	putative signaling protein
CDR20291_3527	-	2.05187674	2.46E-06	hypothetical protein
CDR20291_0082	rplF	-1.8928633	2.48E-06	50S ribosomal protein L6
CDR20291_1589	trxB1	1.5135328	2.83E-06	thioredoxin reductase
CDR20291_0059	-	1.97606456	2.83E-06	NADP-dependent 7-alpha-hydroxysteroid dehydrogenase
CDR20291_1535	-	-1.3923691	3.04E-06	putative calcium-transporting ATPase
CDR20291_1588	trxA1	1.58040468	3.70E-06	thioredoxin
CDR20291_0892	-	-2.0325107	3.93E-06	cell surface protein (putative N-acetylmuramoyl-L-alanine amidase) / cwp17
CDR20291_3308	atpF	1.76656687	4.17E-06	ATP synthase B chain
CDR20291_0330	cbiN	-2.5453944	4.81E-06	cobalt transport protein
CDR20291_0958	-	2.22246832	6.21E-06	AraC-family transcriptional regulator
CDR20291_2558	appF	1.96664863	6.21E-06	oligopeptide ABC transporter, ATP-binding protein
CDR20291_0081	rpsH	-1.7436962	6.35E-06	30S ribosomal protein S8
CDR20291_1617	-	-2.0578174	7.35E-06	putative hydantoinase
CDR20291_0633	-	-2.4224657	7.54E-06	putative signaling protein
CDR20291_1101	-	1.7357711	8.07E-06	putative ribonucleotide-diphosphate reductase
CDR20291_3428	pyrAB2	1.96994675	8.51E-06	carbamoyl-phosphate synthase, pyrimidine-specific, large chain
CDR20291_0982	mreB2	1.319138	9.96E-06	rod shape-determining protein
CDR20291_0516	-	-1.9636502	1.03E-05	putative cation transporting ATPase
CDR20291_3524	-	-1.5465391	1.18E-05	putative amino acid aminotransferase
CDR20291_1693	-	1.875741	1.31E-05	LysR-family transcriptional regulator
CDR20291_2744	dltA	2.08679545	1.36E-05	D-alanine--poly(phosphoribitol) ligase subunit 1 (D-alanine-activating enzyme)
CDR20291_2127	-	2.08279376	1.70E-05	hypothetical protein
CDR20291_0649	-	-2.0524904	1.93E-05	putative methylenetetrahydrofolate reductase
CDR20291_0003	serS1	-2.1039776	2.52E-05	seryl-tRNA synthetase
CDR20291_0656	gcvH	-2.2321198	2.52E-05	putative glycine cleavage system H protein

CDR20291_0923	ccpA	1.62783071	2.54E-05	LacI-family transcriptional regulator (catabolite control protein)
CDR20291_3352	gcaD	1.78755684	2.82E-05	bifunctional protein
CDR20291_1019	fabD	1.53619136	3.00E-05	malonyl coa-acyl carrier protein transacylase
CDR20291_2560	appA	-1.6591077	3.26E-05	oligopeptide ABC transporter, substrate-binding protein
CDR20291_0332	cbiO	-1.9920841	3.39E-05	cobalt ABC transporter, ATP-binding protein
CDR20291_1955	-	-1.7694058	3.44E-05	RpiR-family transcriptional regulator
CDR20291_1264	-	-2.0970397	4.07E-05	putative ATP-binding protein
CDR20291_0725	nadE	1.48442875	4.14E-05	NH3-dependent NAD(+) synthetase
CDR20291_2373	-	-2.0572643	4.35E-05	hypothetical protein
CDR20291_0074	rplP	-1.4345121	5.04E-05	50S ribosomal protein L16
CDR20291_1346	-	1.78500267	5.17E-05	hypothetical protein
CDR20291_2199	cspD	3.96977455	7.19E-05	cold shock protein
CDR20291_3231	hprK	1.70579223	8.16E-05	HPr(Ser) kinase/phosphorylase
CDR20291_0176	-	2.04916312	9.00E-05	putative oxidoreductase, electron transfer subunit
CDR20291_0073	rpsC	-1.4356532	9.15E-05	30S ribosomal protein S3
CDR20291_0095	rpsD	-1.2423235	9.90E-05	30S ribosomal protein S4
CDR20291_2601	-	-1.0821267	0.00010308	cell surface protein / cwp22
CDR20291_1657	-	1.84760246	0.00010453	hypothetical protein
CDR20291_3132	-	2.17677911	0.00011779	hypothetical protein
CDR20291_3026	eno	1.84215396	0.00011838	enolase
CDR20291_3305	atpG	1.2758258	0.00011868	ATP synthase subunit gamma
CDR20291_1975	-	1.52732979	0.00012988	ABC transporter, ATP-binding protein
CDR20291_0903	-	-1.6904928	0.00014386	cell surface protein / cwp18
CDR20291_0005	dnaH	1.84831072	0.00014643	DNA polymerase III subunit gamma/tau
CDR20291_1038	-	1.95280797	0.00014946	hypothetical protein
CDR20291_0086	rplO	-1.7148344	0.00015639	50S ribosomal protein L15
CDR20291_0077	rplN	-1.5391432	0.00017413	50S ribosomal protein L14
CDR20291_0096	rpoA	-1.5029822	0.00017672	DNA-directed RNA polymerase alpha chain
CDR20291_3135	feoB3	-1.4814623	0.0001783	putative ferrous iron transport protein B
CDR20291_0078	rplX	-1.772655	0.00019146	50S ribosomal protein L24
CDR20291_2248	-	-1.5422201	0.00020964	putative aliphatic sulfonate ABC transporter, ATP-binding protein
CDR20291_0070	rplB	-1.4734479	0.00021837	50S ribosomal protein L2
CDR20291_0978	rnfB	-1.8141073	0.00022115	electron transport complex protein
CDR20291_2671	-	1.24688811	0.00022115	putative glycosyltransferase
CDR20291_3521	rpsR	-1.5146856	0.00022277	30S ribosomal protein S18
CDR20291_2825	-	1.79627671	0.00024113	ABC transporter, ATP-binding protein
CDR20291_1157	rpsO	2.0430733	0.00024355	30S ribosomal protein S15
CDR20291_0085	rpmD	-1.5474514	0.00024566	50S ribosomal protein L30
CDR20291_3297	-	-1.5007445	0.00026731	putative regulator of cell growth

CDR20291_2907	-	-1.8458035	0.00028766	hypothetical protein
CDR20291_3233	uvrA	1.55978286	0.00031975	excinuclease ABC subunit A
CDR20291_2126	-	1.47820029	0.00032779	putative radical SAM superfamily lipoprotein
CDR20291_2030	-	1.61950611	0.00033308	putative multiprotein-complex assembly protein
CDR20291_3397	-	1.66743712	0.00034818	putative ATPase
CDR20291_1794	-	-1.6006363	0.00037014	hypothetical protein
CDR20291_2520	-	1.44768696	0.00037014	putative GTP-binding protein
CDR20291_3461	-	1.8676761	0.00037559	chloramphenicol o-acetyltransferase
CDR20291_0065	-	-1.5461543	0.00038314	elongation factor TU
CDR20291_2668	pgm2	1.6053753	0.00039239	putative phosphomannomutase/ phosphoglycerate mutase
CDR20291_1129	-	1.10016799	0.00039808	metallo beta-lactamase superfamily protein
CDR20291_2763	tldD	1.86995415	0.00040422	putative regulatory protease
CDR20291_1170	-	-1.1028285	0.00043463	putative nucleic acid-binding protein
CDR20291_1299	folP	1.40055632	0.00043757	dihydropteroate synthase
CDR20291_0141	-	-1.836354	0.00049389	putative RNA-binding protein
CDR20291_2325	glyQ	-1.3391281	0.00049992	glycyl-tRNA synthetase alpha chain
CDR20291_2682	slpA	-1.1828344	0.00051138	cell surface protein (S-layer precursor protein)
CDR20291_0087	prlA	-1.8442563	0.00051529	preprotein translocase SecY subunit
CDR20291_1854	dinR	-1.3928453	0.000582	SOS regulatory protein
CDR20291_1066	-	1.35988331	0.00062702	putative protease
CDR20291_3518	-	1.32507628	0.00062702	putative RNA/single-stranded DNA exonuclease
CDR20291_2081	-	-1.6074941	0.00064767	putative amino-acid ABC transporter, permease protein
CDR20291_0088	adk	-1.4211985	0.00072166	adenylate kinase
CDR20291_2570	nifJ	-1.6706062	0.00072166	pyruvate-flavodoxin oxidoreductase
CDR20291_0097	rplQ	-1.0855423	0.00072975	50S ribosomal protein L17
CDR20291_1759	-	-1.4538172	0.00079809	addiction module toxin, rele/stbe family
CDR20291_3087	dapD	1.05803474	0.00079809	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase
CDR20291_1013	-	1.5689422	0.00080199	hypothetical protein
CDR20291_0083	rplR	-1.2683067	0.00093053	50S ribosomal protein L18
CDR20291_1644	hgdC	1.31789298	0.00098989	putative (R)-2-hydroxyglutaryl-CoA dehydratase alpha-subunit
CDR20291_1334	-	1.94309672	0.0009908	hypothetical protein
CDR20291_0849	-	-3.2306233	0.00100318	hypothetical protein
CDR20291_1760	-	-1.6209336	0.00101742	addiction module antitoxin, relb/dinj family
CDR20291_0084	rpsE	-1.2125713	0.00118302	30S ribosomal protein S5
CDR20291_3541	rpmH	2.99600178	0.00118989	50S ribosomal protein L34
CDR20291_2253	-	-1.5039125	0.00127765	hypothetical protein
CDR20291_0076	rpsQ	-1.3343648	0.00146699	30S ribosomal protein S17
CDR20291_0818	speA	1.55709598	0.00160942	arginine decarboxylase

CDR20291_0973	rnfC	-1.1566543	0.00164261	electron transport complex protein
CDR20291_1643	-	1.19973757	0.00164261	putative 2-hydroxyacyl-CoA dehydratase
CDR20291_2824	-	1.11623619	0.00172097	ABC transporter, substrate-binding protein
CDR20291_0758	-	1.39703431	0.00178201	putative oxidative stress protein
CDR20291_2561	appB	-1.6920155	0.00179092	oligopeptide ABC transporter, permease protein
CDR20291_3523	rpsF	-1.160151	0.00197042	30S ribosomal protein S6
CDR20291_3307	atpH	1.76933914	0.00198653	ATP synthase subunit delta
CDR20291_3230	-	1.41262458	0.00205243	putative DNA mismatch repair protein
CDR20291_3402	-	1.21670615	0.00205858	GntR-family transcriptional regulator
CDR20291_0060	rpoB	1.06725547	0.00206112	DNA-directed RNA polymerase beta chain
CDR20291_0094	rpsK	-1.0550523	0.00223311	30S ribosomal protein S11
CDR20291_0184	-	1.17151178	0.00230051	putative cell wall hydrolase
CDR20291_0983	mreC	1.31531552	0.00237984	putative rod shape-determining protein precursor
CDR20291_1537	-	-1.4312171	0.00260116	putative tellurite resistance protein
CDR20291_0071	rpsS	-1.7887099	0.00285112	30S ribosomal protein S19
CDR20291_1664	-	1.01194202	0.00301455	hypothetical protein
CDR20291_2336	-	-1.6917374	0.00310164	putative sigma 54 modulation protein
CDR20291_1298	folE	1.26374622	0.00324117	putative GTP cyclohydrolase I
CDR20291_2250	-	-1.3078221	0.00333558	hypothetical protein
CDR20291_0711	-	-1.1820521	0.00343159	cytidine/deoxycytidylate deaminase family protein
CDR20291_0173	fdxA	-1.9706247	0.00350214	ferredoxin
CDR20291_2733	-	-1.1237361	0.00360803	nitrilase (carbon-nitrogen hydrolase)
CDR20291_1020	fabG	1.23135957	0.00364892	3-oxoacyl-[acyl-carrier protein] reductase
CDR20291_2099	-	-1.1838678	0.00381546	cell surface protein
CDR20291_3320	prfA	1.05759802	0.00396135	peptide chain release factor 1
CDR20291_2800	adhE	1.6876753	0.00397607	aldehyde-alcohol dehydrogenase [includes: alcohol dehydrogenase and pyruvate-formate-lyase deactivase
CDR20291_3409	-	1.61017553	0.00423413	putative preprotein translocase
CDR20291_3225	-	-1.4641828	0.00489984	putative formate/nitrite transporter
CDR20291_1532	-	1.07601528	0.00517407	tellurium resistance protein
CDR20291_1344	-	-1.4138832	0.00546786	putative transcriptional regulator
CDR20291_1862	accB	-1.4285059	0.00587629	biotin carboxyl carrier protein of acetyl-CoA carboxylase
CDR20291_3291	-	1.31140418	0.00587629	putative exported carboxy-terminal processing protease
CDR20291_0075	rpmC	-1.2419799	0.00618084	50S ribosomal protein L29
CDR20291_2044	rrf	-1.3445672	0.0064983	ribosome recycling factor
CDR20291_2254	-	-1.1765446	0.00666602	putative permease
CDR20291_1385	-	-1.3548953	0.00668763	putative dehydrogenase, electron transfer subunit
EBG00000018527		-4.4164926	0.00668763	
CDR20291_2644	ptsH	1.05957778	0.00670464	PTS system, phosphocarrier protein

CDR20291_1327	feoA1	-1.323862	0.00678323	putative ferrous iron transport protein A
CDR20291_2072	msrAB	1.26315822	0.00702578	putative peptide methionine sulfoxide reductase
CDR20291_1323	-	1.33901621	0.00704539	putative ruberythrin
CDR20291_2338	rpsU	-1.0165067	0.007632	putative tRNA binding protein
CDR20291_3337	prsA	-1.3696952	0.0081275	putative foldase lipoprotein (late stage protein export lipoprotein)
CDR20291_3530	-	-1.0958516	0.00842799	putative selenocysteine lyase
CDR20291_0687	plfB	1.02619623	0.00885547	formate acetyltransferase
CDR20291_2659	rkpK	1.2742871	0.00934616	putative UDP-glucose 6-dehydrogenase
CDR20291_3086	dapB1	1.04749243	0.00935339	dihydrodipicolinate reductase
CDR20291_0072	rplV	-1.0027807	0.0098289	50S ribosomal protein L22
CDR20291_2256	-	-1.2313101	0.00991778	hypothetical protein
CDR20291_1053	-	1.24548078	0.01037582	putative pyrophosphokinase
CDR20291_1301	folK	1.05082156	0.01172782	2-amino-4-hydroxy-6-hydroxymethyldihydropteridi ne pyrophosphokinase
CDR20291_3143	pflD	-1.4303476	0.01193383	putative formate acetyltransferase / down in dupuy BF paper
CDR20291_0277	htpG	1.03958355	0.01201347	chaperone protein (heat shock protein)
CDR20291_1900	mutS	1.062734	0.01255006	DNA mismatch repair protein
CDR20291_2762	-	1.11841509	0.01256357	putative regulatory protease
CDR20291_1772	-	-1.212013	0.01320495	hypothetical protein
CDR20291_3528	-	1.28881474	0.01359973	putative transcriptional regulator
CDR20291_2355	grpE	1.12437996	0.01447643	heat shock protein
CDR20291_1954	-	2.5954282	0.0158501	hypothetical protein
CDR20291_3379	-	-1.169779	0.0163713	AsnC-family transcriptional regulator
CDR20291_1375	pyrC	1.17098277	0.01661204	dihydroorotase
CDR20291_1014	rpmF	1.0825952	0.01749591	50S ribosomal protein L32
CDR20291_2502	-	-1.0693062	0.01767492	isoleucyl-tRNA synthetase
CDR20291_2562	appC	-1.2354836	0.01767492	oligopeptide ABC transporter, permease protein
CDR20291_2649	-	1.10985355	0.01767492	putative N-acetylmuramoyl-L-alanine amidase
CDR20291_2866	-	1.21722234	0.0178612	transcription antiterminator
CDR20291_2292	-	-1.1307312	0.02148455	putative cell wall hydrolase
CDR20291_1548	-	1.13759217	0.0221947	putative iron compound ABC transporter, substrate-binding protein
CDR20291_2045	pyrH	1.08404922	0.02328628	uridylate kinase
CDR20291_1672	-	2.00209076	0.02717869	putative arsenate reductase
CDR20291_3147	-	1.03293539	0.02774833	putative sodium-dependent phosphate transporter
CDR20291_1866	-	1.05489788	0.02920065	hypothetical protein
CDR20291_1861	accC	-1.1296023	0.02934772	biotin carboxylase (acetyl-CoA carboxylase subunit A)
CDR20291_2579	hpt	1.2785186	0.03121939	putative phosphoribosyltransferase
CDR20291_2118	-	1.08307188	0.03195353	ABC transporter, ATP-binding/permease protein

CDR20291_0370	acdB	1.06206267	0.03214753	acyl-CoA dehydrogenase, short-chain specific
CDR20291_0918	acpP	2.01054875	0.03458976	acyl carrier protein
CDR20291_1859	accA	-1.0002914	0.03458976	acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha
CDR20291_2341	-	1.21847257	0.03458976	putative radical SAM superfamily protein
CDR20291_1675	-	1.06852396	0.03540544	putative methylase
CDR20291_2445	recG	1.10151308	0.0361845	ATP-dependent DNA helicase
CDR20291_3229	hymC	-1.0355191	0.0368305	putative iron-only hydrogenase, catalytic subunit
CDR20291_1925	fldX	1.34673959	0.0373868	flavodoxin
CDR20291_1860	accD	-1.0211901	0.0385727	acetyl-coenzyme A carboxylase carboxyl transferase subunit beta
CDR20291_2063	-	-1.1927297	0.03955513	hypothetical protein
CDR20291_1481	-	1.09165683	0.04073452	hypothetical protein
CDR20291_2082	-	-1.1264193	0.04514509	putative amino-acid ABC transporter, permease protein
CDR20291_3390	greA	1.11980161	0.0496842	transcription elongation factor grea

Appendix table 6. Bacterial differentially expressed genes between uninfected controls and infected samples at 12 h post infection.

Gene ID	**Gene name**	**log_2(fold change)**	**padj**	**annotation**
CDR20291_2167	-	5.17069612	1.98E-52	PTS system, IIbc component
CDR20291_2845	-	3.967875	4.47E-29	putative pyridine-nucleotide-disulfide oxidoreductase
CDR20291_3031	cggR	4.8610279	2.08E-26	central glycolytic genes regulator
CDR20291_0119	glmS	3.35671925	1.60E-23	glucosamine--fructose-6-phosphate aminotransferase [isomerizing]
CDR20291_0445	-	-4.2496721	1.61E-23	hypothetical protein
CDR20291_2340	-	-4.2616308	2.25E-23	histidine triad nucleotide-binding protein
CDR20291_2410	leuS	-4.4342508	2.41E-22	leucyl-tRNA synthetase
CDR20291_2554	crr	-4.0576151	1.37E-21	PTS system, glucose-specific IIa component
CDR20291_3028	tpi	3.31762883	3.00E-21	triosephosphate isomerase
CDR20291_2024	-	3.62769105	5.69E-21	thioredoxin reductase
CDR20291_1719	cysD	-4.8045739	1.07E-20	putative O-acetylhomoserine sulfhydrylase
CDR20291_1100	-	-3.8656527	2.51E-20	branched chain amino acid transport system carrier protein
CDR20291_0944	-	-3.9605694	1.04E-19	hypothetical protein
CDR20291_0760	-	-4.4256703	3.13E-19	hypothetical protein
CDR20291_0122	murA	3.21067621	1.56E-18	UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1
CDR20291_0118	-	3.03031702	1.95E-17	phosphoglucomutase/ phosphomannomutase mutase
CDR20291_2555		-4.0687479	1.03E-16	
CDR20291_3029	pgK	3.21936605	1.20E-16	phosphoglycerate kinase

CDR20291_1118	-	3.34008158	3.38E-16	RRF2-family transcriptional regulator
CDR20291_0470	-	-2.9480732	1.22E-15	putative lipoprotein
CDR20291_2427	nox	2.76254908	3.39E-15	NADH oxidase
CDR20291_1119	iscS2	2.90003115	6.00E-15	cysteine desulfurase
CDR20291_1180	-	-2.5717954	6.00E-15	aspartate aminotransferase
CDR20291_3217	pfkA	2.86906573	6.89E-15	6-phosphofructokinase
CDR20291_1250	-	-2.5747771	8.67E-15	putative oligopeptide transporter
CDR20291_1718	-	3.0332384	2.55E-14	putative cations-transporting ATPase
CDR20291_1556	gcvPB	-3.4053891	2.55E-14	glycine cleavage system P protein
CDR20291_0815	glgP	-3.0839132	2.58E-14	glycogen phosphorylase
CDR20291_1963	-	-2.8953823	7.77E-14	hypothetical protein
CDR20291_2686	-	2.6642188	8.55E-14	hypothetical protein
CDR20291_1318	-	-2.8467603	2.35E-13	cell surface protein (putative penicillin-binding protein) / cwp20
CDR20291_0329	cbiM	-3.3507595	2.84E-13	putative cobalt transport protein
CDR20291_2685	-	2.55843215	3.27E-13	cell surface protein / cwp10
CDR20291_2655	-	-2.4699496	4.48E-13	cell surface protein / cwp19
CDR20291_0141	-	-3.7560373	1.47E-12	putative RNA-binding protein
CDR20291_1308	-	-2.5797999	1.47E-12	putative 5-nitroimidazole reductase
CDR20291_2774	fhuD	2.89825521	7.13E-12	putative ferrichrome ABC transporter, substrate-binding protein
CDR20291_3027	gpmI	2.91010329	1.17E-11	2,3-bisphosphoglycerate-independent phosphoglycerate mutase
CDR20291_1120	-	2.82367796	1.46E-11	NifU-like protein
CDR20291_0655	-	-2.1256663	1.93E-11	putative carbon monoxide dehydrogenase/acetyl-CoA synthase complex, beta subunit
CDR20291_2676	cwp84	-2.4799584	2.07E-11	cell surface protein (putative cell surface-associated cysteine protease)
CDR20291_0039	-	-3.941826	3.32E-11	putative dual-specificity prolyl/cysteinyl-tRNA synthetase
CDR20291_0847	-	-3.3058589	3.32E-11	putative aminotransferase
CDR20291_0180	gluD	-2.663526	3.48E-11	NAD-specific glutamate dehydrogenase
CDR20291_2826	-	2.53007978	3.53E-11	putative ABC transporter, permease protein
CDR20291_1104	-	-3.4383006	4.15E-11	hypothetical protein
CDR20291_0187	pyrD	2.86274219	5.26E-11	dihydroorotate dehydrogenase, catalytic subunit
CDR20291_2266	-	-2.5301626	5.44E-11	butyrate kinase
CDR20291_0643	cooS	-2.4190007	5.70E-11	putative bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase
CDR20291_0945	-	-3.0019148	5.95E-11	putative peptidase
CDR20291_0689	-	2.61745985	1.10E-10	putative ATP-dependent RNA helicase
CDR20291_0755	rbr	1.95498071	1.34E-10	rubrerythrin
CDR20291_2624	-	-2.441257	1.69E-10	cell surface protein / cwp14
CDR20291_3334	-	2.41446146	3.01E-10	putative bifunctional protein [include tetrapyrrole (Corrin/Porphyrin) methylase and nucleoside triphosphate pyrophosphohydrolase]

CDR20291_0802	-	-2.8897828	3.63E-10	ABC transporter, substrate-binding lipoprotein
CDR20291_3434	-	-2.8657089	4.35E-10	putative homocysteine S-methyltransferase
CDR20291_3428	pyrAB2	2.53805195	7.82E-10	carbamoyl-phosphate synthase, pyrimidine-specific, large chain
CDR20291_0507	gapN	2.03818778	8.00E-10	NADP-dependent glyceraldehyde-3-phosphate dehydrogenase
CDR20291_3392	-	1.86677542	1.10E-09	hypothetical protein
CDR20291_1395	-	-2.7985187	1.45E-09	putative hemolysin-like membrane protein
CDR20291_0217	purF	2.45280756	1.55E-09	amidophosphoribosyltransferase
CDR20291_2324	glyS	-1.906225	3.85E-09	glycyl-tRNA synthetase beta chain
CDR20291_2055	-	2.41012504	4.05E-09	hypothetical protein
CDR20291_1017	fabH	2.24536345	4.25E-09	3-oxoacyl-[acyl-carrier-protein] synthase III
CDR20291_3030	gapB	2.87635642	5.50E-09	glyceraldehyde-3-phosphate dehydrogenase 2
CDR20291_1778	-	2.05483309	5.64E-09	hypothetical protein
CDR20291_1876	-	2.56367544	1.18E-08	hypothetical protein
CDR20291_0060	rpoB	1.83392624	1.32E-08	DNA-directed RNA polymerase beta chain
CDR20291_1182	exoA	2.81685452	1.32E-08	putative exodeoxyribonuclease
CDR20291_2369	-	2.59089633	1.76E-08	hypothetical protein
CDR20291_0003	serS1	-2.7549953	2.24E-08	seryl-tRNA synthetase
CDR20291_2678	cwp66	-2.145417	2.24E-08	cell surface protein
CDR20291_0595		2.43834931	2.46E-08	
CDR20291_0354	-	2.48275494	3.14E-08	hypothetical protein
CDR20291_3168	-	1.85181304	3.49E-08	putative phosphoesterase
CDR20291_0005	dnaH	2.42941904	4.43E-08	DNA polymerase III subunit gamma/tau
CDR20291_2336	-	-3.0410212	6.88E-08	putative sigma 54 modulation protein
CDR20291_3527	-	2.24240449	6.94E-08	hypothetical protein
CDR20291_2252	-	-2.1035659	7.23E-08	putative sulfonate ABC transporter, solute-binding lipoprotein
CDR20291_3299	alr	-1.9891386	7.36E-08	alanine racemase
CDR20291_1555	-	-2.1988492	7.66E-08	putative bi-functional glycine dehydrogenase/aminomethyl transferase protein
CDR20291_3167	tig	1.47901911	9.44E-08	trigger factor
CDR20291_0221	purD	2.24395821	1.15E-07	phosphoribosylamine--glycine ligase
CDR20291_0818	speA	2.42797978	1.41E-07	arginine decarboxylase
CDR20291_0218	purG	2.59738975	1.71E-07	phosphoribosylformylglycinamidine cyclo-ligase
CDR20291_2510	-	-1.867872	1.90E-07	cinserved hypothetical protein
CDR20291_0785	oppA	-2.6810445	1.93E-07	oligopeptide ABC transporter, substrate-binding lipoprotein
CDR20291_2030	-	2.14291358	2.15E-07	putative multiprotein-complex assembly protein
CDR20291_1617	-	-2.2789403	3.20E-07	putative hydantoinase
CDR20291_0721	-	-2.1649467	3.91E-07	putative NUDIX-family hydrolase
CDR20291_2325	glyQ	-1.9034207	3.91E-07	glycyl-tRNA synthetase alpha chain

CDR20291_0516	-	-2.2175491	4.34E-07	putative cation transporting ATPase
CDR20291_1545	-	2.18474145	4.94E-07	putative iron compound ABC transporter, permease protein
CDR20291_0783	oppB	-1.7520987	5.08E-07	oligopeptide ABC transporter, permease protein
CDR20291_0106	nrdD	2.11047819	5.71E-07	anaerobic ribonucleoside-triphosphate reductase
CDR20291_2558	appF	2.0908381	5.71E-07	oligopeptide ABC transporter, ATP-binding protein
CDR20291_1794	-	-2.1717084	7.60E-07	hypothetical protein
CDR20291_1018	fabK	1.60576823	7.85E-07	trans-2-enoyl-ACP reductase
CDR20291_2825	-	2.24875767	7.92E-07	ABC transporter, ATP-binding protein
CDR20291_1197	cspB	-2.0521167	9.38E-07	putative cold shock protein
CDR20291_1975	-	1.83665567	9.79E-07	ABC transporter, ATP-binding protein
CDR20291_3352	gcaD	1.96083405	1.29E-06	bifunctional protein [includes: UDP-N-acetylglucosamine pyrophosphorylase and glucosamine-1-phosphate N-acetyltransferase]
CDR20291_0492	-	2.11906463	1.54E-06	hypothetical protein
CDR20291_3393	-	1.77758708	2.31E-06	hypothetical protein
CDR20291_1493	cysA	-1.5890677	2.37E-06	serine acetyltransferase
CDR20291_3234	uvrB	1.9259002	2.55E-06	excinuclease ABC subunit B
CDR20291_0102	-	1.44165892	3.11E-06	50S ribosomal protein L13 (pseudogene)
CDR20291_1328	feoB1	-2.4765195	3.52E-06	ferrous iron transport protein B
CDR20291_2373	-	-2.2699897	3.60E-06	hypothetical protein
CDR20291_3405	veg	-2.0343924	3.79E-06	hypothetical protein
CDR20291_0685	-	1.8318899	4.22E-06	putative signaling protein
CDR20291_1643	-	1.63324805	4.89E-06	putative 2-hydroxyacyl-CoA dehydratase
CDR20291_0880	potA	1.97432509	5.45E-06	spermidine/putrescine ABC transporter, ATP-binding protein
CDR20291_1115	codY	-1.6501443	5.45E-06	GTP-sensing transcriptional pleiotropic repressor
CDR20291_0645	fhs	-1.9115162	7.11E-06	formate--tetrahydrofolate ligase
CDR20291_2127	-	2.10369539	7.19E-06	hypothetical protein
CDR20291_0355	-	2.09670592	7.63E-06	hypothetical protein
CDR20291_0982	mreB2	1.29268436	8.85E-06	rod shape-determining protein
CDR20291_2671	-	1.43897089	1.06E-05	putative glycosyltransferase
CDR20291_0914	hbd	-2.0564734	1.07E-05	3-hydroxybutyryl-CoA dehydrogenase
CDR20291_1639	-	-2.0494389	1.07E-05	putative ferrous iron transport protein A
CDR20291_2087	-	1.85556695	1.07E-05	putative aromatic compounds hydrolase
CDR20291_0912	etfA2	-1.9143978	1.14E-05	electron transfer flavoprotein alpha-subunit
CDR20291_0915	thlA1	-2.0770678	1.36E-05	acetyl-CoA acetyltransferase
CDR20291_1644	hgdC	1.66521038	1.38E-05	putative (R)-2-hydroxyglutaryl-CoA dehydratase alpha-subunit
CDR20291_2733	-	-1.6179454	1.40E-05	nitrilase (carbon-nitrogen hydrolase)
CDR20291_3320	prfA	1.46796929	1.45E-05	peptide chain release factor 1
CDR20291_2744	dltA	2.01078802	1.52E-05	D-alanine--poly(phosphoribitol) ligase subunit 1 (D-alanine-activating enzyme)

CDR20291_1657	-	1.94574855	1.63E-05	hypothetical protein
CDR20291_0317	-	2.4129644	1.66E-05	ArsR-family transcriptional regulator
CDR20291_0330	cbiN	-2.2867027	1.66E-05	cobalt transport protein
CDR20291_2800	adhE	2.4680355	1.66E-05	aldehyde-alcohol dehydrogenase [includes: alcohol dehydrogenase and pyruvate-formate-lyase deactivase
CDR20291_3026	eno	2.02523534	1.94E-05	enolase
CDR20291_1854	dinR	-1.6621783	2.29E-05	SOS regulatory protein
CDR20291_1013	-	1.87085181	2.56E-05	hypothetical protein
CDR20291_1645	-	-1.6797515	2.67E-05	cell surface protein (putative cell surface-associated cysteine protease) / cwp13
CDR20291_2571	-	-1.5548538	2.67E-05	putative propanediol utilization protein
CDR20291_3402	-	1.48380639	3.02E-05	GntR-family transcriptional regulator
CDR20291_3114	valS	-1.7120431	3.02E-05	valyl-tRNA synthetase
CDR20291_0107	nrdG	1.86436199	3.52E-05	anaerobic ribonucleoside-triphosphate reductase activating protein
CDR20291_0366	hadA	-1.599074	4.37E-05	isocaprenoyl-CoA:2-hydroxyisocaproate CoA-transferase
CDR20291_3169	rph	1.54596638	4.61E-05	ribonuclease Ph
CDR20291_0184	-	1.4756409	5.17E-05	putative cell wall hydrolase
CDR20291_0656	gcvH	-2.0826123	5.32E-05	putative glycine cleavage system H protein
CDR20291_0219	purN	2.35555355	5.35E-05	phosphoribosylglycinamide formyltransferase
CDR20291_2116	-	1.622182	5.71E-05	putative GTP-binding protein
CDR20291_1264	-	-1.9708449	5.72E-05	putative ATP-binding protein
CDR20291_2824	-	1.35462678	5.90E-05	ABC transporter, substrate-binding protein
CDR20291_0849	-	-3.9083266	6.03E-05	hypothetical protein
CDR20291_3216	pykF	1.35300043	6.77E-05	pyruvate kinase
CDR20291_0771	-	-1.548534	6.92E-05	putative formyltransferase
CDR20291_2152	kamA	-1.8410441	6.92E-05	L-lysine 2,3-aminomutase
CDR20291_0978	rnfB	-1.8757631	7.29E-05	electron transport complex protein
CDR20291_0756	-	1.72508475	7.50E-05	putative oxidative stress regulatory protein
CDR20291_1386	aspB	-1.4994581	7.76E-05	putative glutamate synthase [NADPH] small chain
CDR20291_2983	-	1.27168962	8.19E-05	abc-type fe3+ transport system periplasmic component-like protein precursor
CDR20291_0173	fdxA	-2.6078148	8.71E-05	ferredoxin
CDR20291_1038	-	1.96017079	8.71E-05	hypothetical protein
CDR20291_3530	-	-1.5638047	8.74E-05	putative selenocysteine lyase
CDR20291_2708	-	-1.8753778	9.34E-05	putative amino acid racemase
CDR20291_1373	-	1.44139061	0.00010182	putative rubrerythrin
CDR20291_0081	rpsH	-1.4418842	0.00010438	30S ribosomal protein S8
CDR20291_2126	-	1.52361183	0.00010915	putative radical SAM superfamily lipoprotein
CDR20291_0086	rplO	-1.7007077	0.00011864	50S ribosomal protein L15
CDR20291_3306	atpA	1.17063112	0.00012622	ATP synthase alpha chain

CDR20291_0881	potB	1.53638598	0.00012648	spermidine/putrescine ABC transporter, permease protein
CDR20291_1127	fur	1.56442856	0.00012983	ferric uptake regulation protein
CDR20291_0903	-	-1.6231477	0.00014555	cell surface protein / cwp18
CDR20291_0911	etfB2	-1.6235223	0.00015938	electron transfer flavoprotein beta-subunit
CDR20291_1693	-	1.59385171	0.00018056	LysR-family transcriptional regulator
CDR20291_0787	oppF	-1.5823056	0.00019694	oligopeptide ABC transporter, ATP-binding protein
CDR20291_2063	-	-2.1287951	0.00020066	hypothetical protein
CDR20291_1170	-	-1.1299103	0.00022316	putative nucleic acid-binding protein
CDR20291_1019	fabD	1.34131208	0.00022652	malonyl coa-acyl carrier protein transacylase
CDR20291_0757	rbo	1.57507705	0.00023881	rubredoxin oxidoreductase (desulfoferrodoxin)
CDR20291_0633	-	-1.8339032	0.00024582	putative signaling protein
CDR20291_3397	-	1.65523337	0.00024582	putative ATPase
CDR20291_1548	-	1.72857464	0.00025239	putative iron compound ABC transporter, substrate-binding protein
CDR20291_0786	oppD	-1.6162457	0.00027608	oligopeptide ABC transporter, ATP-binding protein
CDR20291_1664	-	1.17658282	0.00029216	hypothetical protein
CDR20291_1329	-	-1.4253785	0.00031304	hypothetical protein
CDR20291_3524	-	-1.2427115	0.00033967	putative amino acid aminotransferase
CDR20291_1535	-	-1.0359635	0.00036909	putative calcium-transporting ATPase
CDR20291_3409	-	1.85129636	0.00046597	putative preprotein translocase
CDR20291_1799	-	-1.1361101	0.00048446	hypothetical protein
CDR20291_0014	-	-1.5735972	0.00052998	putative ATP:guanido phosphotransferase
CDR20291_0892	-	-1.4284127	0.00054227	cell surface protein (putative N-acetylmuramoyl-L-alanine amidase) / cwp17
CDR20291_1129	-	1.0454481	0.00056278	metallo beta-lactamase superfamily protein
CDR20291_0240	fliC	-1.8093292	0.00058187	flagellin subunit
CDR20291_0649	-	-1.4983857	0.00058187	putative methylenetetrahydrofolate reductase
CDR20291_0053	secE	1.69063026	0.00058545	preprotein translocase SecE subunit
CDR20291_3147	-	1.43687318	0.00061723	putative sodium-dependent phosphate transporter
CDR20291_1345	proC1	-1.3958351	0.00062424	pyrroline-5-carboxylate reductase
CDR20291_3308	atpF	1.32114994	0.00063026	ATP synthase B chain
CDR20291_3461	-	1.74922465	0.00064284	chloramphenicol o-acetyltransferase
CDR20291_0958	-	1.68269288	0.00070312	AraC-family transcriptional regulator
CDR20291_0626	pheT	-1.3338814	0.00070692	phenylalanyl-tRNA synthetase beta chain
CDR20291_2253	-	-1.4818893	0.00070692	hypothetical protein
CDR20291_2668	pgm2	1.46535945	0.0007189	putative phosphomannomutase/phosphoglycerate mutase
CDR20291_0087	prlA	-1.7561796	0.00072802	preprotein translocase SecY subunit
CDR20291_0083	rplR	-1.2571731	0.00073959	50S ribosomal protein L18

CDR20291_0096	rpoA	-1.3155522	0.00073959	DNA-directed RNA polymerase alpha chain
CDR20291_3518	-	1.27126525	0.00075621	putative RNA/single-stranded DNA exonuclease
CDR20291_2615	glyA	-1.1225885	0.00076906	putative serine hydroxymethyltransferase
CDR20291_2649	-	1.43593229	0.00077867	putative N-acetylmuramoyl-L-alanine amidase
CDR20291_2081	-	-1.4703379	0.00081034	putative amino-acid ABC transporter, permease protein
CDR20291_0097	rplQ	-1.0354421	0.00085214	50S ribosomal protein L17
CDR20291_0983	mreC	1.37260585	0.00085796	putative rod shape-determining protein precursor
CDR20291_0088	adk	-1.3540148	0.00086499	adenylate kinase
CDR20291_2631	-	-1.3223214	0.00087222	putative hydrolase
CDR20291_0913	crt2	-1.3711101	0.00098642	3-hydroxybutyryl-CoA dehydratase
CDR20291_2248	-	-1.2829665	0.00098642	putative aliphatic sulfonate ABC transporter, ATP-binding protein
CDR20291_2353	dnaJ	-1.5356537	0.00099205	chaperone protein
CDR20291_0175	-	1.38818937	0.00118179	putative oxidoreductase, acetyl-CoA synthase subunit
CDR20291_2520	-	1.27794081	0.00121331	putative GTP-binding protein
CDR20291_1933	clpB	-3.2525167	0.00122482	chaperone
CDR20291_2874	-	1.60346003	0.00132271	hypothetical protein
CDR20291_0725	nadE	1.15595402	0.00133652	NH3-dependent NAD(+) synthetase
CDR20291_1346	-	1.4075864	0.00134895	hypothetical protein
CDR20291_3521	rpsR	-1.2648738	0.00136507	30S ribosomal protein S18
CDR20291_1862	accB	-1.5527168	0.00141724	biotin carboxyl carrier protein of acetyl-CoA carboxylase
CDR20291_0084	rpsE	-1.164546	0.00144289	30S ribosomal protein S5
CDR20291_2628	aspS	-1.3938099	0.00146706	putative aspartyl-tRNA synthetase
CDR20291_3337	prsA	-1.5535041	0.00150858	putative foldase lipoprotein (late stage protein export lipoprotein)
CDR20291_0687	plfB	1.20334276	0.00157727	formate acetyltransferase
CDR20291_0711	-	-1.2138946	0.00161967	cytidine/deoxycytidylate deaminase family protein
CDR20291_2659	rkpK	1.42005921	0.00162352	putative UDP-glucose 6-dehydrogenase
CDR20291_0051	-	-1.363282	0.00178763	elongation factor TU
CDR20291_0094	rpsK	-1.0349507	0.00194581	30S ribosomal protein S11
CDR20291_0819	speD	1.28819699	0.00194581	S-adenosylmethionine decarboxylase proenzyme
CDR20291_3231	hprK	1.32731178	0.00194581	HPr(Ser) kinase/phosphorylase
CDR20291_3528	-	1.50166989	0.00216442	putative transcriptional regulator
CDR20291_3390	greA	1.55666861	0.00221168	transcription elongation factor grea
CDR20291_1053	-	1.38996596	0.00221268	putative pyrophosphokinase
CDR20291_0910	bcd2	-1.2291263	0.00225873	butyryl-CoA dehydrogenase
CDR20291_2561	appB	-1.5511816	0.00226964	oligopeptide ABC transporter, permease protein
CDR20291_2907	-	-1.5243504	0.00227689	hypothetical protein
CDR20291_1157	rpsO	1.69764066	0.00228922	30S ribosomal protein S15
CDR20291_2360	lepA	1.38732099	0.00245967	GTP-binding elongation factor

CDR20291_0054	nusG	1.20976467	0.0027495	transcription antitermination protein
CDR20291_1344	-	-1.439957	0.00284514	putative transcriptional regulator
CDR20291_2478	ltaE	-1.3338663	0.00287764	low-specificity L-threonine aldolase
CDR20291_3541	rpmH	2.7421437	0.00297225	50S ribosomal protein L34
CDR20291_0082	rplF	-1.1411565	0.00316683	50S ribosomal protein L6
CDR20291_0758	-	1.29475138	0.00316683	putative oxidative stress protein
CDR20291_1334	-	1.72548633	0.0032068	hypothetical protein
CDR20291_2502	-	-1.2683624	0.00337919	isoleucyl-tRNA synthetase
CDR20291_0089	map1	-1.1836802	0.00342819	methionine aminopeptidase
CDR20291_0073	rpsC	-1.0523797	0.0034587	30S ribosomal protein S3
CDR20291_3143	pflD	-1.5816349	0.00351929	putative formate acetyltransferase / down in dupuy BF paper
CDR20291_0105	aspC	-1.727822	0.00365474	aspartate aminotransferase
CDR20291_3355	murC	1.14041375	0.00365474	UDP-N-acetylmuramate--L-alanine ligase
CDR20291_2118	-	1.33800343	0.00393126	ABC transporter, ATP-binding/permease protein
CDR20291_1547	-	1.45199415	0.00476156	putative iron compound ABC transporter, ATP-binding protein
CDR20291_0637	argS	1.11260608	0.00502667	arginyl-tRNA synthetase
CDR20291_2866	-	1.37091217	0.00510982	transcription antiterminator
CDR20291_0093	rpsM	-1.5043226	0.00548583	30S ribosomal protein S13
CDR20291_3132	-	1.55784063	0.0066855	hypothetical protein
CDR20291_0332	cbiO	-1.2538977	0.00713747	cobalt ABC transporter, ATP-binding protein
CDR20291_1066	-	1.05289342	0.00731601	putative protease
CDR20291_3048	-	-1.0949278	0.00765183	cell surface protein / cwp21
CDR20291_0957	-	1.0413471	0.00792704	nitroreductase-family protein
CDR20291_1663	-	-1.3121759	0.00824315	hypothetical protein
CDR20291_2630	-	-1.377909	0.00887991	putative coproporphyrinogen III oxidase
CDR20291_2297	-	1.03170818	0.0090109	putative multidrug efflux pump, membrane protein
CDR20291_0345	-	-1.4814064	0.00925525	hypothetical protein
CDR20291_0085	rpmD	-1.0227169	0.00929412	50S ribosomal protein L30
CDR20291_0078	rplX	-1.1825366	0.010184	50S ribosomal protein L24
CDR20291_1385	-	-1.2132977	0.01055383	putative dehydrogenase, electron transfer subunit
CDR20291_2540	ftsW	1.10637605	0.01255721	cell division/stage V sporulation protein
CDR20291_2579	hpt	1.39704698	0.01282755	putative phosphoribosyltransferase
CDR20291_2473	coaBC	-1.0671795	0.01324836	coenzyme A biosynthesis bifunctional protein
CDR20291_1287	-	1.00216677	0.01364572	putative multiprotein complex assembly protein
CDR20291_2199	cspD	2.55492423	0.0147326	cold shock protein
CDR20291_2763	tldD	1.29265278	0.01684861	putative regulatory protease
CDR20291_0490	nth	1.03966649	0.01797253	endonuclease iii
CDR20291_0368	hadB	-1.2039998	0.01815327	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase
CDR20291_1925	fldX	1.50908459	0.0185813	flavodoxin

CDR20291_3224	murB	1.10816173	0.01871931	UDP-N-acetylenolpyruvoylglucosamine reductase
CDR20291_2499	-	1.32498007	0.02018961	hypothetical protein
CDR20291_1029	fbp	1.15120718	0.02027081	putative fructose-1,6-bisphosphatase
CDR20291_2437	-	1.0750479	0.02484277	putative sugar transporter, substrate-binding lipoprotein
CDR20291_3379	-	-1.0316086	0.02567462	AsnC-family transcriptional regulator
CDR20291_2044	rrf	-1.0427867	0.02591061	ribosome recycling factor
CDR20291_3064	-	1.0692454	0.02605354	hypothetical protein
CDR20291_2835	-	-1.0856298	0.0306347	putative iron ABC transporter, substrate-binding protein
CDR20291_2721	-	-1.1229921	0.03153164	hypothetical protein / zmp1
CDR20291_2062	-	-1.1456844	0.03174819	radical SAM-superfamily protein
CDR20291_2445	recG	1.06994121	0.0325563	ATP-dependent DNA helicase
CDR20291_3307	atpH	1.2354051	0.03377624	ATP synthase subunit delta
CDR20291_1327	feoA1	-1.0299921	0.03398578	putative ferrous iron transport protein A
CDR20291_3225	-	-1.0418965	0.0371333	putative formate/nitrite transporter
CDR20291_2023	bipA	1.06316242	0.04050796	GTP-binding protein
CDR20291_0822	cspA	1.11425774	0.04228817	cold shock protein
CDR20291_1324	-	-1.0922885	0.04713185	hypothetical protein
CDR20291_2731	serS2	-1.0061413	0.04713185	seryl-tRNA synthetase

Appendix table 7. Bacterial differentially expressed genes between uninfected controls and infected samples at 24 h post infection.

Gene ID	**Gene name**	**log_2(fold change)**	**padj**	**annotation**
CDR20291_2410	leuS	-5.709829758	4.76E-29	leucyl-tRNA synthetase
CDR20291_2167	-	4.088378751	2.06E-27	PTS system, IIbc component
CDR20291_3031	cggR	5.435392934	4.02E-27	central glycolytic genes regulator
CDR20291_0445	-	-4.811914101	1.33E-25	hypothetical protein
CDR20291_3028	tpi	3.861673085	2.56E-25	triosephosphate isomerase
CDR20291_2845	-	3.918672851	7.17E-25	putative pyridine-nucleotide-disulfide oxidoreductase
CDR20291_1118	-	4.235017874	3.04E-23	RRF2-family transcriptional regulator
CDR20291_0119	glmS	3.38386132	4.87E-21	glucosamine--fructose-6-phosphate aminotransferase [isomerizing]
CDR20291_1719	cysD	-5.006063371	1.67E-19	putative O-acetylhomoserine sulfhydrylase
CDR20291_0944	-	-4.217679363	2.96E-19	hypothetical protein
CDR20291_2340	-	-4.012102946	3.01E-19	histidine triad nucleotide-binding protein
CDR20291_0039	-	-5.919484015	4.33E-19	putative dual-specificity prolyl/cysteinyl-tRNA synthetase
CDR20291_3029	pgK	3.444080825	6.78E-16	phosphoglycerate kinase
CDR20291_1119	iscS2	3.279415837	1.54E-15	cysteine desulfurase

CDR20291_0122	murA	3.012593363	2.16E-14	UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1
CDR20291_0118	-	2.89886765	3.95E-14	phosphoglucomutase/ phosphomannomutase mutase
CDR20291_1250	-	-2.668277738	7.06E-14	putative oligopeptide transporter
CDR20291_2510	-	-2.903571244	1.18E-13	cinserved hypothetical protein
CDR20291_0945	-	-3.747683743	1.27E-13	putative peptidase
CDR20291_2554	crr	-3.263177548	5.66E-13	PTS system, glucose-specific IIa component
CDR20291_3392	-	2.319534962	6.47E-13	hypothetical protein
CDR20291_1718	-	3.145731135	8.61E-13	putative cations-transporting ATPase
CDR20291_1100	-	-3.097239314	1.63E-12	branched chain amino acid transport system carrier protein
CDR20291_3217	pfkA	2.652116592	4.37E-11	6-phosphofructokinase
CDR20291_2555		-3.473760774	6.71E-11	
CDR20291_3027	gpmI	3.125934542	6.71E-11	2,3-bisphosphoglycerate-independent phosphoglycerate mutase
CDR20291_2676	cwp84	-2.592295595	8.70E-11	cell surface protein (putative cell surface-associated cysteine protease)
CDR20291_2655	-	-2.361049512	1.30E-10	cell surface protein / cwp19
CDR20291_0141	-	-3.627081519	2.23E-10	putative RNA-binding protein
CDR20291_3434	-	-3.176573309	3.87E-10	putative homocysteine S-methyltransferase
CDR20291_0470	-	-2.375533309	6.02E-10	putative lipoprotein
CDR20291_0760	-	-2.983549305	7.55E-10	hypothetical protein
CDR20291_1963	-	-2.478313602	1.03E-09	hypothetical protein
CDR20291_3030	gapB	3.37816638	1.03E-09	glyceraldehyde-3-phosphate dehydrogenase 2
CDR20291_2427	nox	2.345810575	1.09E-09	NADH oxidase
CDR20291_1308	-	-2.399825576	1.11E-09	putative 5-nitroimidazole reductase
CDR20291_0847	-	-3.256601544	1.19E-09	putative aminotransferase
CDR20291_2686	-	2.378623301	2.08E-09	hypothetical protein
CDR20291_1120	-	2.807221003	2.19E-09	NifU-like protein
CDR20291_0106	nrdD	2.592072598	4.85E-09	anaerobic ribonucleoside-triphosphate reductase
CDR20291_2685	-	2.240739399	5.81E-09	cell surface protein / cwp10
CDR20291_0755	rbr	1.949120348	6.16E-09	rubrerythrin
CDR20291_0815	glgP	-2.445200488	8.94E-09	glycogen phosphorylase
CDR20291_2336	-	-3.537453443	9.94E-09	putative sigma 54 modulation protein
CDR20291_1318	-	-2.369812976	1.57E-08	cell surface protein (putative penicillin-binding protein) / cwp20
CDR20291_1104	-	-3.103199381	1.66E-08	hypothetical protein
CDR20291_1617	-	-2.719848789	2.55E-08	putative hydantoinase
CDR20291_2774	fhuD	2.568969257	2.63E-08	putative ferrichrome ABC transporter, substrate-binding protein
CDR20291_0084	rpsE	-2.19910614	2.74E-08	30S ribosomal protein S5
CDR20291_0176	-	2.871003955	3.52E-08	putative oxidoreductase, electron transfer subunit
CDR20291_2325	glyQ	-2.317054576	5.01E-08	glycyl-tRNA synthetase alpha chain
CDR20291_2324	glyS	-1.891496115	5.15E-08	glycyl-tRNA synthetase beta chain

CDR20291_2624	-	-2.211774781	5.98E-08	cell surface protein / cwp14
CDR20291_3103	prdA	2.386296288	6.78E-08	proline reductase subunit proprotein
CDR20291_0756	-	2.476757724	9.10E-08	putative oxidative stress regulatory protein
CDR20291_3527	-	2.347333261	1.14E-07	hypothetical protein
CDR20291_0177	-	1.799402439	1.15E-07	putative oxidoreductase, NAD/FAD binding subunit
CDR20291_0818	speA	2.671600852	1.29E-07	arginine decarboxylase
CDR20291_3393	-	2.09439643	1.40E-07	hypothetical protein
CDR20291_2024	-	2.32543324	1.78E-07	thioredoxin reductase
CDR20291_0074	rplP	-1.924680785	2.23E-07	50S ribosomal protein L16
CDR20291_2826	-	2.170236174	2.23E-07	putative ABC transporter, permease protein
CDR20291_3334	-	2.143100066	3.05E-07	putative bifunctional protein [include tetrapyrrole (Corrin/Porphyrin) methylase and nucleoside triphosphate pyrophosphohydrolase]
CDR20291_3234	uvrB	2.18934535	3.11E-07	excinuclease ABC subunit B
CDR20291_1180	-	-1.774152851	3.97E-07	aspartate aminotransferase
CDR20291_3402	-	1.848040461	4.73E-07	GntR-family transcriptional regulator
CDR20291_0371	etfB1	2.185729802	6.01E-07	electron transfer flavoprotein beta-subunit
CDR20291_3167	tig	1.483111334	9.77E-07	trigger factor
CDR20291_0107	nrdG	2.282035434	1.00E-06	anaerobic ribonucleoside-triphosphate reductase activating protein
CDR20291_2825	-	2.364094037	1.15E-06	ABC transporter, ATP-binding protein
CDR20291_0595		2.302624826	1.24E-06	
CDR20291_0060	rpoB	1.729062202	1.33E-06	DNA-directed RNA polymerase beta chain
CDR20291_0003	serS1	-2.540258814	1.38E-06	seryl-tRNA synthetase
CDR20291_2800	adhE	3.090675998	1.55E-06	aldehyde-alcohol dehydrogenase [includes: alcohol dehydrogenase and pyruvate-formate-lyase deactivase
CDR20291_0175	-	2.173431789	1.57E-06	putative oxidoreductase, acetyl-CoA synthase subunit
CDR20291_0217	purF	2.134395858	1.59E-06	amidophosphoribosyltransferase
CDR20291_3428	pyrAB2	2.17360189	1.59E-06	carbamoyl-phosphate synthase, pyrimidine-specific, large chain
CDR20291_2824	-	1.688735459	1.89E-06	ABC transporter, substrate-binding protein
CDR20291_0083	rplR	-1.909890236	2.13E-06	50S ribosomal protein L18
CDR20291_0077	rplN	-2.035409605	2.13E-06	50S ribosomal protein L14
CDR20291_0978	rnfB	-2.449603378	2.39E-06	electron transport complex protein
CDR20291_1197	cspB	-2.171159944	2.50E-06	putative cold shock protein
CDR20291_3168	-	1.712610489	2.58E-06	putative phosphoesterase
CDR20291_3523	rpsF	-1.846450787	2.83E-06	30S ribosomal protein S6
CDR20291_3461	-	2.442308941	3.14E-06	chloramphenicol o-acetyltransferase
CDR20291_2678	cwp66	-1.885796716	4.20E-06	cell surface protein
CDR20291_1264	-	-2.454347003	4.36E-06	putative ATP-binding protein
CDR20291_2252	-	-1.933642179	4.44E-06	putative sulfonate ABC transporter, solute-binding lipoprotein

CDR20291_0073	rpsC	-1.787199498	4.46E-06	30S ribosomal protein S3
CDR20291_1876	-	2.242698418	5.35E-06	hypothetical protein
CDR20291_0082	rplF	-1.899161763	5.52E-06	50S ribosomal protein L6
CDR20291_0097	rplQ	-1.51979098	5.52E-06	50S ribosomal protein L17
CDR20291_0372	etfA1	2.312860595	5.52E-06	electron transfer flavoprotein alpha-subunit
CDR20291_3521	rpsR	-1.942977352	5.71E-06	30S ribosomal protein S18
CDR20291_0076	rpsQ	-1.988026213	6.18E-06	30S ribosomal protein S17
CDR20291_2152	kamA	-2.231456998	8.88E-06	L-lysine 2,3-aminomutase
CDR20291_0757	rbo	2.056981094	1.02E-05	rubredoxin oxidoreductase (desulfoferrodoxin)
CDR20291_0079	rplE	-1.542209219	1.11E-05	50S ribosomal protein L5
CDR20291_0370	acdB	2.19141721	1.35E-05	acyl-CoA dehydrogenase, short-chain specific
CDR20291_2373	-	-2.287121828	1.36E-05	hypothetical protein
CDR20291_1371	-	-1.496332312	1.42E-05	putative polysaccharide deacetylase
CDR20291_3216	pykF	1.599107378	1.42E-05	pyruvate kinase
CDR20291_2297	-	1.717531176	1.52E-05	putative multidrug efflux pump, membrane protein
CDR20291_0218	purG	2.337989194	1.60E-05	phosphoribosylformylglycinamidine cyclo-ligase
CDR20291_0086	rplO	-2.048428778	1.91E-05	50S ribosomal protein L15
CDR20291_0221	purD	1.978017412	2.03E-05	phosphoribosylamine--glycine ligase
CDR20291_1643	-	1.638246954	2.04E-05	putative 2-hydroxyacyl-CoA dehydratase
CDR20291_0369	hadC	2.276530578	2.07E-05	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase
CDR20291_2866	-	2.154123961	2.30E-05	transcription antiterminator
CDR20291_0687	plfB	1.752791817	2.45E-05	formate acetyltransferase
CDR20291_3026	eno	2.24326915	2.45E-05	enolase
CDR20291_0721	-	-1.918076611	2.45E-05	putative NUDIX-family hydrolase
CDR20291_1115	codY	-1.636905747	2.49E-05	GTP-sensing transcriptional pleiotropic repressor
CDR20291_2369	-	2.126478176	2.68E-05	hypothetical protein
CDR20291_0507	gapN	1.550166015	2.84E-05	NADP-dependent glyceraldehyde-3-phosphate dehydrogenase
CDR20291_3114	valS	-1.857336937	3.29E-05	valyl-tRNA synthetase
CDR20291_1129	-	1.30831169	3.66E-05	metallo beta-lactamase superfamily protein
CDR20291_0689	-	1.872624468	4.10E-05	putative ATP-dependent RNA helicase
CDR20291_0075	rpmC	-1.945282076	4.67E-05	50S ribosomal protein L29
CDR20291_0081	rpsH	-1.629019654	4.67E-05	30S ribosomal protein S8
CDR20291_0078	rplX	-2.016226556	4.69E-05	50S ribosomal protein L24
CDR20291_2733	-	-1.638865227	4.97E-05	nitrilase (carbon-nitrogen hydrolase)
CDR20291_0095	rpsD	-1.336701238	5.01E-05	30S ribosomal protein S4
CDR20291_0129	metE	-1.458144287	5.63E-05	S-adenosylmethionine synthetase
CDR20291_1182	exoA	2.190000337	7.10E-05	putative exodeoxyribonuclease
CDR20291_1493	cysA	-1.463229132	7.10E-05	serine acetyltransferase

CDR20291_0880	potA	1.857061133	7.32E-05	spermidine/putrescine ABC transporter, ATP-binding protein
CDR20291_1794	-	-1.837515971	7.32E-05	hypothetical protein
CDR20291_1395	-	-1.906375203	8.05E-05	putative hemolysin-like membrane protein
CDR20291_0354	-	1.973349959	8.81E-05	hypothetical protein
CDR20291_1151	infB	-1.298721377	0.000101496	translation initiation factor IF-2
CDR20291_0051	-	-1.826359466	0.000110657	elongation factor TU
CDR20291_2601	-	-1.125191546	0.000118043	cell surface protein / cwp22
CDR20291_3299	alr	-1.476734218	0.000167397	alanine racemase
CDR20291_1778	-	1.502052475	0.000178817	hypothetical protein
CDR20291_2671	-	1.344863651	0.000178857	putative glycosyltransferase
CDR20291_2499	-	2.141773002	0.000185031	hypothetical protein
CDR20291_3352	gcaD	1.650758775	0.000215	bifunctional protein [includes: UDP-N-acetylglucosamine pyrophosphorylase and glucosamine-1-phosphate N-acetyltransferase]
CDR20291_0958	-	1.91755497	0.000230462	AraC-family transcriptional regulator
CDR20291_0973	rnfC	-1.383443013	0.000266789	electron transport complex protein
CDR20291_2907	-	-1.996735741	0.000280572	hypothetical protein
CDR20291_0070	rplB	-1.535724796	0.000283539	50S ribosomal protein L2
CDR20291_0222	purL	1.339967006	0.00028537	formylglycinamidine ribonucleotide synthetase
CDR20291_1013	-	1.743442206	0.000305967	hypothetical protein
CDR20291_0005	dnaH	1.806055318	0.000314093	DNA polymerase III subunit gamma/tau
CDR20291_2615	glyA	-1.299707165	0.000327653	putative serine hydroxymethyltransferase
CDR20291_1014	rpmF	1.637203639	0.000348393	50S ribosomal protein L32
CDR20291_2116	-	1.549970234	0.000376288	putative GTP-binding protein
CDR20291_1556	gcvPB	-1.676478144	0.000415736	glycine cleavage system P protein
CDR20291_2571	-	-1.405557746	0.000415736	putative propanediol utilization protein
CDR20291_0656	gcvH	-1.957698572	0.000462216	putative glycine cleavage system H protein
CDR20291_3306	atpA	1.159193991	0.000464959	ATP synthase alpha chain
CDR20291_0094	rpsK	-1.259556349	0.00047317	30S ribosomal protein S11
CDR20291_1545	-	1.678416247	0.000550222	putative iron compound ABC transporter, permease protein
CDR20291_3231	hprK	1.548898277	0.000602139	HPr(Ser) kinase/phosphorylase
CDR20291_3405	veg	-1.657280906	0.000631151	hypothetical protein
CDR20291_1017	fabH	1.4764503	0.000695619	3-oxoacyl-[acyl-carrier-protein] synthase III
CDR20291_1817	-	-1.664961777	0.00071197	hypothetical protein
CDR20291_0184	-	1.33893682	0.000786844	putative cell wall hydrolase
CDR20291_0072	rplV	-1.355106031	0.000792386	50S ribosomal protein L22
CDR20291_1862	accB	-1.815387847	0.000808992	biotin carboxyl carrier protein of acetyl-CoA carboxylase
CDR20291_0881	potB	1.455036232	0.000821859	spermidine/putrescine ABC transporter, permease protein
CDR20291_3396	ftsH2	-1.159431496	0.000821859	cell division protein

CDR20291_0059	-	1.520673839	0.000863058	NADP-dependent 7-alpha-hydroxysteroid dehydrogenase
CDR20291_2087	-	1.57433611	0.000875836	putative aromatic compounds hydrolase
CDR20291_0819	speD	1.483811513	0.000882384	S-adenosylmethionine decarboxylase proenzyme
CDR20291_0626	pheT	-1.434620737	0.00105301	phenylalanyl-tRNA synthetase beta chain
CDR20291_0085	rpmD	-1.402275816	0.001076228	50S ribosomal protein L30
CDR20291_0219	purN	2.083915268	0.001129186	phosphoribosylglycinamide formyltransferase
CDR20291_3518	-	1.307517335	0.001331595	putative RNA/single-stranded DNA exonuclease
CDR20291_0018	-	-1.155679148	0.001372769	hypothetical protein
CDR20291_0087	prlA	-1.82462852	0.001372769	preprotein translocase SecY subunit
CDR20291_1344	-	-1.689284098	0.001378762	putative transcriptional regulator
CDR20291_3307	atpH	1.862843199	0.001378762	ATP synthase subunit delta
CDR20291_0355	-	1.669518323	0.001449219	hypothetical protein
CDR20291_0823	-	1.208568549	0.001449219	iron-dependent hydrogenase
CDR20291_2055	-	1.491534245	0.001548057	hypothetical protein
CDR20291_1644	hgdC	1.339340563	0.001787318	putative (R)-2-hydroxyglutaryl-CoA dehydratase alpha-subunit
CDR20291_0903	-	-1.443247788	0.001788679	cell surface protein / cwp18
CDR20291_2127	-	1.634319251	0.001804825	hypothetical protein
CDR20291_3233	uvrA	1.43813821	0.001828201	excinuclease ABC subunit A
CDR20291_3516	dnaB	1.275516497	0.001829033	replicative DNA helicase
CDR20291_1664	-	1.101879153	0.00184458	hypothetical protein
CDR20291_0711	-	-1.293726546	0.001937983	cytidine/deoxycytidylate deaminase family protein
CDR20291_0685	-	1.383073207	0.001989471	putative signaling protein
CDR20291_0088	adk	-1.357495854	0.002073858	adenylate kinase
CDR20291_2044	rrf	-1.56602337	0.002144309	ribosome recycling factor
CDR20291_0173	fdxA	-2.2706472	0.002673596	ferredoxin
CDR20291_2649	-	1.390592367	0.002737365	putative N-acetylmuramoyl-L-alanine amidase
CDR20291_1328	feoB1	-1.851288392	0.002748461	ferrous iron transport protein B
CDR20291_0329	cbiM	-1.498642313	0.00315094	putative cobalt transport protein
CDR20291_2353	dnaJ	-1.48068629	0.003481518	chaperone protein
CDR20291_3169	rph	1.230748627	0.003571693	ribonuclease Ph
CDR20291_0849	-	-3.158054065	0.003705947	hypothetical protein
CDR20291_0771	-	-1.214461851	0.003806084	putative formyltransferase
CDR20291_2502	-	-1.355428898	0.004106075	isoleucyl-tRNA synthetase
CDR20291_1645	-	-1.247431247	0.004152213	cell surface protein (putative cell surface-associated cysteine protease) / cwp13
CDR20291_2338	rpsU	-1.182054186	0.004282712	putative tRNA binding protein
CDR20291_0089	map1	-1.259394111	0.004353802	methionine aminopeptidase
CDR20291_0725	nadE	1.110227748	0.004730675	NH3-dependent NAD(+) synthetase
CDR20291_3132	-	1.721523213	0.005038174	hypothetical protein

CDR20291_0931	-	-1.311740006	0.005049925	putative aminopeptidase
CDR20291_3522	ssb	-1.082680019	0.005049925	single-strand binding protein
CDR20291_2631	-	-1.209385992	0.005337062	putative hydrolase
CDR20291_3230	-	1.348942546	0.005403558	putative DNA mismatch repair protein
CDR20291_1797	-	1.416930027	0.005534628	putative conjugative transposon DNA recombination protein
CDR20291_2199	cspD	3.134605446	0.005698021	cold shock protein
CDR20291_1639	-	-1.435529003	0.006108998	putative ferrous iron transport protein A
CDR20291_2633	relA	-1.166534574	0.006108998	putative GTP pyrophosphokinase
CDR20291_2256	-	-1.374020591	0.006345264	hypothetical protein
CDR20291_3308	atpF	1.161451172	0.006384009	ATP synthase B chain
CDR20291_0633	-	-1.469601612	0.006407943	putative signaling protein
CDR20291_1933	clpB	-3.015055946	0.006884064	chaperone
CDR20291_3337	prsA	-1.434430388	0.007116447	putative foldase lipoprotein (late stage protein export lipoprotein)
CDR20291_1532	-	1.082274932	0.007256094	tellurium resistance protein
CDR20291_2253	-	-1.263204044	0.007613736	hypothetical protein
CDR20291_0367	hadI	1.235533667	0.007742236	activator of 2-hydroxyisocaproyl-CoA dehydratase
CDR20291_0105	aspC	-1.726289418	0.007798611	aspartate aminotransferase
CDR20291_2561	appB	-1.461575807	0.008276979	oligopeptide ABC transporter, permease protein
CDR20291_1346	-	1.271596824	0.00868289	hypothetical protein
CDR20291_0368	hadB	1.321654717	0.008863437	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase
CDR20291_1373	-	1.113750577	0.00895935	putative rubrerythrin
CDR20291_1788	-	2.293798404	0.009359665	hypothetical protein
CDR20291_1127	fur	1.199906834	0.009464339	ferric uptake regulation protein
CDR20291_0317	-	1.671366245	0.009492622	ArsR-family transcriptional regulator
CDR20291_2155	-	1.325149776	0.010215537	hypothetical protein
CDR20291_0758	-	1.237127109	0.011059989	putative oxidative stress protein
CDR20291_3147	-	1.192181307	0.011109645	putative sodium-dependent phosphate transporter
CDR20291_0187	pyrD	1.41372171	0.011212274	dihydroorotate dehydrogenase, catalytic subunit
CDR20291_1675	-	1.270653279	0.012150281	putative methylase
CDR20291_2777	-	1.350320572	0.01277997	hypothetical protein
CDR20291_0516	-	-1.242931677	0.013487909	putative cation transporting ATPase
CDR20291_0058	rplL	-1.094829595	0.014410162	50S ribosomal protein L7/L12
CDR20291_3530	-	-1.038176291	0.015264595	putative selenocysteine lyase
CDR20291_1760	-	-1.310662804	0.016301657	addiction module antitoxin, relb/dinj family
CDR20291_3048	-	-1.091323359	0.016648357	cell surface protein / cwp21
CDR20291_2520	-	1.058505493	0.017424732	putative GTP-binding protein
CDR20291_1657	-	1.261602444	0.01750047	hypothetical protein
CDR20291_0883	potD	1.021830445	0.018339994	spermidine/putrescine ABC transporter, substrate-binding lipoprotein

CDR20291_1859	accA	-1.165464456	0.018550083	acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha
CDR20291_1038	-	1.347117788	0.018710231	hypothetical protein
CDR20291_2630	-	-1.355909831	0.018710231	putative coproporphyrinogen III oxidase
CDR20291_2406	tdcB	1.040071492	0.019037833	threonine dehydratase catabolic
CDR20291_1663	-	-1.263953628	0.019785084	hypothetical protein
CDR20291_0892	-	-1.042494521	0.020676122	cell surface protein (putative N-acetylmuramoyl-L-alanine amidase) / cwp17
CDR20291_2030	-	1.150831449	0.020676122	putative multiprotein-complex assembly protein
CDR20291_2063	-	-1.39134909	0.020861193	hypothetical protein
CDR20291_3409	-	1.385943867	0.021235157	putative preprotein translocase
CDR20291_1029	fbp	1.227749363	0.021748338	putative fructose-1,6-bisphosphatase
CDR20291_2027	-	-1.097505139	0.023052328	putative 2-nitropropane dioxygenase
CDR20291_0053	secE	1.277306625	0.023654234	preprotein translocase SecE subunit
CDR20291_2283	-	1.316150782	0.024472092	putative transglycosylase
CDR20291_2250	-	-1.010180485	0.024616452	hypothetical protein
CDR20291_0333	pcrA	1.027475186	0.025836971	ATP-dependent DNA helicase
CDR20291_3390	greA	1.275611677	0.026367101	transcription elongation factor grea
CDR20291_0738	-	1.21105463	0.026798519	hypothetical protein
CDR20291_2126	-	1.006885153	0.026798519	putative radical SAM superfamily lipoprotein
CDR20291_1122	alaS	-1.01535807	0.027172728	putative alanyl-tRNA synthetase
CDR20291_2668	pgm2	1.114589618	0.027172728	putative phosphomannomutase/ phosphoglycerate mutase
CDR20291_2708	-	-1.161908151	0.027172728	putative amino acid racemase
CDR20291_0743	-	-1.395935612	0.031457818	putative universal stress protein
CDR20291_1157	rpsO	1.367089311	0.031457818	30S ribosomal protein S15
CDR20291_2874	-	1.245138381	0.03178265	hypothetical protein
CDR20291_1693	-	1.053758965	0.032133163	LysR-family transcriptional regulator
CDR20291_3064	-	1.109786616	0.032714053	hypothetical protein
CDR20291_0071	rpsS	-1.296843784	0.037942467	30S ribosomal protein S19
CDR20291_2744	dltA	1.155368349	0.037942467	D-alanine--poly(phosphoribitol) ligase subunit 1 (D-alanine-activating enzyme)
CDR20291_1481	-	1.120563806	0.038665704	hypothetical protein
CDR20291_3397	-	1.071328011	0.039139417	putative ATPase
CDR20291_0943	-	-1.084411643	0.040632773	ABC transporter, ATP-binding/permease protein
CDR20291_0093	rpsM	-1.210586966	0.041964417	30S ribosomal protein S13
CDR20291_1860	accD	-1.035444259	0.041964417	acetyl-coenzyme A carboxylase carboxyl transferase subunit beta
CDR20291_2835	-	-1.115020555	0.041964417	putative iron ABC transporter, substrate-binding protein
CDR20291_1925	fldX	1.487213466	0.043706791	flavodoxin
CDR20291_0822	cspA	1.184978201	0.045570865	cold shock protein

Appendix table 8. Significantly upregulated bacterial genes which were upregulated during infection at all timepoints.

Gene ID	Gene name	annotation
CDR20291_0005	dnaH	DNA polymerase III subunit gamma/tau
CDR20291_0060	rpoB	DNA-directed RNA polymerase beta chain
CDR20291_0106	nrdD	anaerobic ribonucleoside-triphosphate reductase
CDR20291_0107	nrdG	anaerobic ribonucleoside-triphosphate reductase activating protein
CDR20291_0118	-	phosphoglucomutase/phosphomannomutase mutase
CDR20291_0119	glmS	glucosamine--fructose-6-phosphate aminotransferase [isomerizing]
CDR20291_0122	murA	UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1
CDR20291_0175	-	putative oxidoreductase, acetyl-CoA synthase subunit
CDR20291_0184	-	putative cell wall hydrolase
CDR20291_0187	pyrD	dihydroorotate dehydrogenase, catalytic subunit
CDR20291_0217	purF	amidophosphoribosyltransferase
CDR20291_0218	purG	phosphoribosylformylglycinamidine cyclo-ligase
CDR20291_0219	purN	phosphoribosylglycinamide formyltransferase
CDR20291_0221	purD	phosphoribosylamine--glycine ligase
CDR20291_0354	-	hypothetical protein
CDR20291_0355	-	hypothetical protein
CDR20291_0507	gapN	NADP-dependent glyceraldehyde-3-phosphate dehydrogenase
CDR20291_0595		
CDR20291_0685	-	putative signaling protein
CDR20291_0687	plfB	formate acetyltransferase
CDR20291_0689	-	putative ATP-dependent RNA helicase
CDR20291_0725	nadE	NH3-dependent NAD(+) synthetase
CDR20291_0818	speA	arginine decarboxylase
CDR20291_0958	-	AraC-family transcriptional regulator

CDR20291_1013	-	hypothetical protein
CDR20291_1017	fabH	3-oxoacyl-[acyl-carrier-protein] synthase III
CDR20291_1038	-	hypothetical protein
CDR20291_1119	iscS2	cysteine desulfurase
CDR20291_1120	-	NifU-like protein
CDR20291_1157	rpsO	30S ribosomal protein S15
CDR20291_1182	exoA	putative exodeoxyribonuclease
CDR20291_1346	-	hypothetical protein
CDR20291_1373	-	putative rubrerythrin
CDR20291_1643	-	putative 2-hydroxyacyl-CoA dehydratase
CDR20291_1644	hgdC	putative (R)-2-hydroxyglutaryl-CoA dehydratase alpha-subunit
CDR20291_1657	-	hypothetical protein
CDR20291_1693	-	LysR-family transcriptional regulator
CDR20291_1718	-	putative cations-transporting ATPase
CDR20291_1778	-	hypothetical protein
CDR20291_1876	-	hypothetical protein
CDR20291_2024	-	thioredoxin reductase
CDR20291_2030	-	putative multiprotein-complex assembly protein
CDR20291_2055	-	hypothetical protein
CDR20291_2087	-	putative aromatic compounds hydrolase
CDR20291_2126	-	putative radical SAM superfamily lipoprotein
CDR20291_2127	-	hypothetical protein
CDR20291_2167	-	PTS system, IIbc component
CDR20291_2199	cspD	cold shock protein
CDR20291_2369	-	hypothetical protein
CDR20291_2427	nox	NADH oxidase
CDR20291_2520	-	putative GTP-binding protein
CDR20291_2649	-	putative N-acetylmuramoyl-L-alanine amidase
CDR20291_2668	pgm2	putative phosphomannomutase/phosphoglycerate mutase
CDR20291_2685	cwp10	cell surface protein
CDR20291_2686	-	hypothetical protein

CDR20291_2744	dltA	D-alanine--poly(phosphoribitol) ligase subunit 1 (D-alanine-activating enzyme)
CDR20291_2774	fhuD	putative ferrichrome ABC transporter, substrate-binding protein
CDR20291_2866	-	transcription antiterminator
CDR20291_2874	-	hypothetical protein
CDR20291_3028	tpi	triosephosphate isomerase
CDR20291_3029	pgK	phosphoglycerate kinase
CDR20291_3031	cggR	central glycolytic genes regulator
CDR20291_3132	-	hypothetical protein
CDR20291_3147	-	putative sodium-dependent phosphate transporter
CDR20291_3167	tig	trigger factor
CDR20291_3168	-	putative phosphoesterase
CDR20291_3169	rph	ribonuclease Ph
CDR20291_3217	pfkA	6-phosphofructokinase
CDR20291_3231	hprK	HPr(Ser) kinase/phosphorylase
CDR20291_3234	uvrB	excinuclease ABC subunit B
CDR20291_3306	atpA	ATP synthase alpha chain
CDR20291_3307	atpH	ATP synthase subunit delta
CDR20291_3308	atpF	ATP synthase B chain
CDR20291_3334	-	putative bifunctional protein
CDR20291_3352	gcaD	bifunctional protein
CDR20291_3390	greA	transcription elongation factor grea
CDR20291_3397	-	putative ATPase
CDR20291_3402	-	GntR-family transcriptional regulator
CDR20291_3409	-	putative preprotein translocase
CDR20291_3428	pyrAB2	carbamoyl-phosphate synthase, pyrimidine-specific, large chain
CDR20291_3461	-	chloramphenicol o-acetyltransferase
CDR20291_3527	-	hypothetical protein

Appendix table 9. Significantly downregulated bacterial genes which were downregulated during infection at all timepoints.

Gene ID	Gene name	annotation
CDR20291_0003	serS1	seryl-tRNA synthetase
CDR20291_0039	-	putative dual-specificity prolyl/cysteinyl-tRNA synthetase
CDR20291_0051	-	elongation factor TU
CDR20291_0078	rplX	50S ribosomal protein L24
CDR20291_0081	rpsH	30S ribosomal protein S8
CDR20291_0082	rplF	50S ribosomal protein L6
CDR20291_0083	rplR	50S ribosomal protein L18
CDR20291_0084	rpsE	30S ribosomal protein S5
CDR20291_0085	rpmD	50S ribosomal protein L30
CDR20291_0086	rplO	50S ribosomal protein L15
CDR20291_0087	prlA	preprotein translocase SecY subunit
CDR20291_0088	adk	adenylate kinase
CDR20291_0097	rplQ	50S ribosomal protein L17
CDR20291_0141	-	putative RNA-binding protein
CDR20291_0329	cbiM	putative cobalt transport protein
CDR20291_0445	-	hypothetical protein
CDR20291_0470	-	putative lipoprotein
CDR20291_0516	-	putative cation transporting ATPase
CDR20291_0633	-	putative signaling protein
CDR20291_0656	gcvH	putative glycine cleavage system H protein
CDR20291_0711	-	cytidine/deoxycytidylate deaminase family protein
CDR20291_0721	-	putative NUDIX-family hydrolase
CDR20291_0760	-	hypothetical protein
CDR20291_0815	glgP	glycogen phosphorylase
CDR20291_0847	-	putative aminotransferase
CDR20291_0849	-	hypothetical protein

CDR20291_0892	cwp17	cell surface protein (putative N-acetylmuramoyl-L-alanine amidase)
CDR20291_0903	cwp18	cell surface protein
CDR20291_0944	-	hypothetical protein
CDR20291_0945	-	putative peptidase
CDR20291_0978	rnfB	electron transport complex protein
CDR20291_1100	-	branched chain amino acid transport system carrier protein
CDR20291_1104	-	hypothetical protein
CDR20291_1180	-	aspartate aminotransferase
CDR20291_1197	cspB	putative cold shock protein
CDR20291_1250	-	putative oligopeptide transporter
CDR20291_1264	-	putative ATP-binding protein
CDR20291_1318	cwp20	cell surface protein (putative penicillin-binding protein)
CDR20291_1328	feoB1	ferrous iron transport protein B
CDR20291_1395	-	putative hemolysin-like membrane protein
CDR20291_1493	cysA	serine acetyltransferase
CDR20291_1556	gcvPB	glycine cleavage system P protein
CDR20291_1617	-	putative hydantoinase
CDR20291_1639	-	putative ferrous iron transport protein A
CDR20291_1645	cwp13	cell surface protein (putative cell surface-associated cysteine protease)
CDR20291_1719	cysD	putative O-acetylhomoserine sulfhydrylase
CDR20291_1862	accB	biotin carboxyl carrier protein of acetyl-CoA carboxylase
CDR20291_1963	-	hypothetical protein
CDR20291_2044	rrf	ribosome recycling factor
CDR20291_2063	-	hypothetical protein
CDR20291_2252	-	putative sulfonate ABC transporter, solute-binding lipoprotein
CDR20291_2336	-	putative sigma 54 modulation protein
CDR20291_2340	-	histidine triad nucleotide-binding protein
CDR20291_2373	-	hypothetical protein
CDR20291_2410	leuS	leucyl-tRNA synthetase
CDR20291_2554	crr	PTS system, glucose-specific IIa component

CDR20291_2555		
CDR20291_2561	appB	oligopeptide ABC transporter, permease protein
CDR20291_2571	-	putative propanediol utilization protein
CDR20291_2624	cwp14	cell surface protein
CDR20291_2655	cwp19	cell surface protein
CDR20291_2676	cwp84	cell surface protein (putative cell surface-associated cysteine protease)
CDR20291_2678	cwp66	cell surface protein
CDR20291_2708	-	putative amino acid racemase
CDR20291_2733	-	nitrilase (carbon-nitrogen hydrolase)
CDR20291_3299	alr	alanine racemase
CDR20291_3337	prsA	putative foldase lipoprotein (late stage protein export lipoprotein)
CDR20291_3405	veg	hypothetical protein
CDR20291_3530	-	putative selenocysteine lyase

Appendix table 10. Human differentially expressed genes between 3 h and 24 h infected samples in the small-scale pilot study.

Gene name	log_2(fold change)	padj	annotation
SLC4A4	2.94935493	1.5609E-16	Electrogenic sodium bicarbonate cotransporter 1
AKR1C1	-3.5827632	1.3052E-10	Aldo-keto reductase family 1 member C1
GADD45B	-2.8988759	3.8076E-10	Growth arrest and DNA damage-inducible protein GADD45 beta
CEBPB	-1.9570938	1.8139E-09	CCAAT/enhancer-binding protein beta
CXCL2	-3.7271777	1.8139E-09	C-X-C motif chemokine 2
FASN	-1.9480008	1.8337E-08	Fatty acid synthase
ATF3	-2.8022869	4.8001E-08	Cyclic AMP-dependent transcription factor ATF-3
NDUFA4L2	3.61018618	7.1047E-08	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 4-like 2
EGR1	-2.0603558	7.9863E-08	Early growth response protein 1
RAB3B	-2.3747867	1.2125E-07	Ras-related protein Rab-3B
NQO1	-1.554601	1.5472E-07	NAD(P)H dehydrogenase [quinone] 1
TNFSF9	-3.2008322	4.8552E-07	Tumor necrosis factor ligand superfamily member 9
SLC29A2	-1.487579	6.089E-07	Equilibrative nucleoside transporter 2
ITPKA	-3.1018008	7.882E-07	Inositol-trisphosphate 3-kinase A
HSD17B7	-2.0389454	1.167E-06	3-keto-steroid reductase

ATP1B1	2.27032052	1.1888E-06	Sodium/potassium-transporting ATPase subunit beta-1
PRSS23	1.82439748	2.1137E-06	Serine protease 23
GBP2	2.27826665	2.3977E-06	Guanylate-binding protein 2
RASD1	-3.6909103	3.488E-06	Dexamethasone-induced Ras-related protein 1
SLC29A4	2.61304585	5.0457E-06	Equilibrative nucleoside transporter 4
PCSK9	-3.345592	5.6521E-06	Proprotein convertase subtilisin/kexin type 9
SNAI1	-3.2445076	6.1616E-06	Zinc finger protein SNAI1
ASIC1	2.31290219	6.2556E-06	Acid-sensing ion channel 1
MVD	-2.3878068	6.2556E-06	Diphosphomevalonate decarboxylase
APOL1	3.30772477	8.6936E-06	Apolipoprotein L1
VNN2	3.37117833	1.0351E-05	Vascular non-inflammatory molecule 2
AKR1C3	-2.1460213	1.2189E-05	Aldo-keto reductase family 1 member C3
SLC39A10	1.53198536	1.3421E-05	Zinc transporter ZIP10
RBCK1	-1.6673829	1.4137E-05	RanBP-type and C3HC4-type zinc finger-containing protein 1
ZNF598	-1.3245853	1.5422E-05	E3 ubiquitin-protein ligase ZNF598
LSS	-1.6544901	1.7224E-05	Lanosterol synthase
CYP51A1	-1.743996	1.8287E-05	Lanosterol 14-alpha demethylase
PIR	-1.874565	1.861E-05	Pirin
SLC3A2	-1.8516479	1.9326E-05	4F2 cell-surface antigen heavy chain
MVK	-1.6226005	2.5305E-05	Mevalonate kinase
ABHD4	-2.2968693	3.8245E-05	Protein ABHD4
TRIB3	-2.1540482	3.9219E-05	Tribbles homolog 3
LAMP3	-3.4374098	5.6105E-05	Lysosome-associated membrane glycoprotein 3
NRN1	2.2721842	5.8492E-05	Neuritin
UNC5B	-2.2334802	5.8537E-05	Netrin receptor UNC5B
FOXD1	-2.0222006	6.2469E-05	Forkhead box protein D1
PRKACB	2.69110601	6.8154E-05	cAMP-dependent protein kinase catalytic subunit beta
RNF24	1.50006918	7.6182E-05	RING finger protein 24
CBX4	-1.7763171	9.299E-05	E3 SUMO-protein ligase CBX4
UBA7	2.8048419	9.299E-05	Ubiquitin-like modifier-activating enzyme 7
HSPA6	-3.4356697	9.8718E-05	Heat shock 70 kDa protein 6
NFKBIA	-1.7756067	0.00010204	NF-kappa-B inhibitor alpha
DHCR7	-1.5031409	0.00012326	7-dehydrocholesterol reductase
KITLG	1.70160818	0.00014495	Kit ligand
SLC27A4	-1.3120064	0.00014495	Long-chain fatty acid transport protein 4
CXCL1	-1.8560306	0.00016466	Growth-regulated alpha protein
ATP8B2	-2.408308	0.00019847	Phospholipid-transporting ATPase ID
LONP1	-1.4721459	0.00019847	Lon protease homolog, mitochondrial
MB	3.13762318	0.00019847	Myoglobin
PTPRB	1.47966868	0.00019847	Receptor-type tyrosine-protein phosphatase beta

SGTB	-2.113975	0.00019847	Small glutamine-rich tetratricopeptide repeat-containing protein beta
UGDH	-1.7318827	0.00019847	UDP-glucose 6-dehydrogenase
CLIC3	2.50720145	0.00024881	Chloride intracellular channel protein 3
IKZF2	2.63260096	0.00029715	Zinc finger protein Helios
CA9	2.87565826	0.00029899	Carbonic anhydrase 9
ERN1	-1.7750711	0.00029899	Serine/threonine-protein kinase/endoribonuclease IRE1
LCN15	1.76781828	0.00030866	Lipocalin-15
CACNG4	2.34080254	0.00038997	Voltage-dependent calcium channel gamma-4 subunit
HSPA13	-1.7746588	0.00038997	Heat shock 70 kDa protein 13
PABPC1L	-1.5067792	0.00039311	Polyadenylate-binding protein 1-like
RNF19B	-1.4860002	0.00041249	E3 ubiquitin-protein ligase RNF19B
TBX3	1.41223976	0.00041249	T-box transcription factor TBX3
SMARCD3	2.27478192	0.00041606	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily D member 3
ME1	-2.8304539	0.00041898	Malic enzyme 1
RAB32	-1.670291	0.00041898	Ras-related protein Rab-32
STARD4	-1.538929	0.00041898	StAR-related lipid transfer protein 4
STS	1.80834698	0.00041898	Steryl-sulfatase
PIGZ	2.07227733	0.00042744	GPI mannosyltransferase 4
CCDC183	1.94561672	0.00047929	Coiled-coil domain containing 183
CST1	2.36110392	0.00048472	Cystatin-SN
MMP24	2.19828228	0.00048472	Matrix metalloproteinase-24
KLHDC7A	2.08667477	0.00053932	Kelch domain containing 7A
AKAP12	1.58812549	0.00053948	A-kinase anchor protein 12
IGSF1	1.69268714	0.00053948	Immunoglobulin superfamily member 1
MIA2	-1.4373366	0.0006607	Endoplasmic reticulum export factor CTAGE5
ATAD2	1.74706827	0.00074582	ATPase family AAA domain-containing protein 2
MUC16	3.01066262	0.00074582	Mucin-16
ZFP36	-1.578521	0.00074582	mRNA decay activator protein ZFP36
PTGR1	-1.4616328	0.00097067	Prostaglandin reductase 1
ULBP3	2.05733119	0.00099909	UL16-binding protein 3
BHLHA15	-3.1640017	0.00101182	Class A basic helix-loop-helix protein 15
RNF128	1.33235175	0.00101182	E3 ubiquitin-protein ligase RNF128
CLGN	-3.033103	0.00113787	Calmegin
IFRD1	-1.6471901	0.00113787	Interferon-related developmental regulator 1
HBEGF	-1.7648569	0.0011816	Proheparin-binding EGF-like growth factor
SOCS1	-3.2614555	0.00122181	Suppressor of cytokine signaling 1
FAM177B	2.80297989	0.00123056	Protein FAM177B
BEST1	-2.2583598	0.00126163	Bestrophin-1
AKR1C2	-2.7733029	0.00128345	Aldo-keto reductase family 1 member C2
MCAM	2.21709406	0.00130322	Cell surface glycoprotein MUC18

RSU1	-1.1879741	0.00130322	Ras suppressor protein 1
EPPK1	1.28690969	0.00130636	Epiplakin
ERO1B	-2.3330135	0.00134298	ERO1-like protein beta
GPR137B	1.44427721	0.00134298	Integral membrane protein GPR137B
KLC2	-1.2302935	0.00134298	Kinesin light chain 2
MMAB	-1.1300912	0.00161414	Cob(I)yrinic acid a,c-diamide adenosyltransferase, mitochondrial
ACAT2	-1.8141392	0.00167594	Acetyl-CoA acetyltransferase, cytosolic
AHNAK	1.74637816	0.00167594	Neuroblast differentiation-associated protein AHNAK
DDIT3	-1.6841927	0.00167594	DNA damage-inducible transcript 3 protein
TNS4	2.33720537	0.00167594	Tensin-4
CAPN13	1.59304482	0.00171211	Calpain-13
CREB3L2	-1.848452	0.00171211	Cyclic AMP-responsive element-binding protein 3-like protein 2
PPP1R15A	-1.959386	0.00172801	Protein phosphatase 1 regulatory subunit 15A
RRM2	2.0408249	0.0017407	Ribonucleoside-diphosphate reductase subunit M2
ESAM	1.49889272	0.00176577	Endothelial cell-selective adhesion molecule
MAFA	-3.1748946	0.00179061	Transcription factor MafA
CYP3A5	1.01658836	0.00198181	Cytochrome P450 3A5
NT5DC4	-2.4666691	0.00202219	5'-nucleotidase domain containing 4
IDH1	-1.2967397	0.00225072	Isocitrate dehydrogenase 1, cytosolic
CXCL8	-2.3076646	0.00248129	Interleukin-8
EIF5	-1.3418706	0.00253747	Eukaryotic translation initiation factor 5
AXIN1	-1.0024555	0.00256241	Axin-1
KREMEN2	2.51302505	0.00278282	Kremen protein 2
WDR18	-1.0562453	0.00279679	WD repeat-containing protein 18
FAM129A	-2.8728264	0.00287577	Protein Niban
CYP4F3	-2.385897	0.00316032	Docosahexaenoic acid omega-hydroxylase CYP4F3
NUCB2	-1.7046413	0.00316032	Nucleobindin-2
SLC29A1	1.27443286	0.00321371	Equilibrative nucleoside transporter 1
G6PD	-1.9756974	0.00322295	Glucose-6-phosphate 1-dehydrogenase
GMPPA	-1.2278084	0.00373671	Mannose-1-phosphate guanyltransferase alpha
TRAK1	1.48674185	0.00402576	Trafficking kinesin-binding protein 1
MIIP	-1.4981054	0.00410504	Migration and invasion-inhibitory protein
FDPS	-1.4663444	0.00425673	Farnesyl pyrophosphate synthase
BAG2	-1.4534776	0.00429338	BAG family molecular chaperone regulator 2
CXCL3	-2.3376101	0.00429338	C-X-C motif chemokine 3
MXD1	-1.7033252	0.00430873	Max dimerization protein 1
TNIK	1.12243219	0.00430873	TRAF2 and NCK-interacting protein kinase
ACER2	1.73989407	0.00439517	Alkaline ceramidase 2
LDLRAD1	2.98384794	0.00439517	Low density lipoprotein receptor class A domain containing 1
NAMPT	-1.2620093	0.00445352	Nicotinamide phosphoribosyltransferase

SLC33A1	-1.3612725	0.00449354	Acetyl-coenzyme A transporter 1
C4orf3	1.85839207	0.00458828	Uncharacterized protein C4orf3
INO80C	-2.1420212	0.00460696	Chromosome 18 open reading frame 37, isoform CRA_a
ADCY7	-2.1448185	0.0046321	Adenylate cyclase type 7
FAM3D	2.68431874	0.0046321	Protein FAM3D
PIGA	-1.5418706	0.0046321	Phosphatidylinositol N-acetylglucosaminyltransferase subunit A
WIPI1	-1.4331154	0.00476832	WD repeat domain phosphoinositide-interacting protein 1
ATP6V1B1	2.96224293	0.00531101	V-type proton ATPase subunit B, kidney isoform
FEN1	1.11470447	0.00549245	Flap endonuclease 1
TP53I3	-1.2867232	0.00562253	Quinone oxidoreductase PIG3
RNASET2	1.79391208	0.00565933	Ribonuclease T2
RASGRF2	-1.6448199	0.00574999	Ras-specific guanine nucleotide-releasing factor 2
HMGCR	-1.383326	0.00579609	3-hydroxy-3-methylglutaryl-coenzyme A reductase
LIG1	1.37872761	0.00579609	DNA ligase 1
SIDT2	1.30545876	0.00579609	SID1 transmembrane family member 2
SCIN	1.91173223	0.00581445	Adseverin
CDKN2C	2.02764027	0.00589219	Cyclin-dependent kinase 4 inhibitor C
TLR4	2.49983956	0.00610212	Toll-like receptor 4
TSPYL2	-1.2999012	0.00617286	Testis-specific Y-encoded-like protein 2
PPARD	1.45211077	0.00624352	Peroxisome proliferator-activated receptor delta
ACSS2	-2.3658413	0.00655677	Acetyl-coenzyme A synthetase, cytoplasmic
EIF2AK3	-1.3214641	0.00656692	Eukaryotic translation initiation factor 2-alpha kinase 3
CDCA5	1.40622322	0.00657118	Sororin
BZRAP1	1.90184449	0.00744958	Benzodiazepine receptor (peripheral) associated protein 1
CNFN	2.08726883	0.00744958	Cornifelin
RDH11	-1.2661981	0.00744958	Retinol dehydrogenase 11
ITM2C	1.22867026	0.00754178	Integral membrane protein 2C
HEXIM1	-1.1893718	0.00787316	Protein HEXIM1
CLCN6	-1.1937022	0.00801061	Chloride transport protein 6
GFPT1	-1.3158219	0.00801061	Glutamine--fructose-6-phosphate aminotransferase [isomerizing] 1
MROH1	-1.0230796	0.00833777	Maestro heat like repeat family member 1
STMN1	1.49933208	0.00835593	Stathmin
GBP3	1.34436547	0.00839978	Guanylate-binding protein 3
RAB7B	2.45709217	0.00839978	Ras-related protein Rab-7b
NPC1	-1.5290868	0.00865268	Niemann-Pick C1 protein
CRABP2	2.28582349	0.00871344	Cellular retinoic acid-binding protein 2
NCR3LG1	-1.8720556	0.00871344	Natural cytotoxicity triggering receptor 3 ligand 1
PDZK1IP1	2.13452357	0.00871344	PDZK1-interacting protein 1

TRMT1	-1.049826	0.00871344	tRNA (guanine(26)-N(2))-dimethyltransferase
ARL6IP5	1.34240347	0.00875108	PRA1 family protein 3
HHLA2	2.26789666	0.00911301	HERV-H LTR-associating protein 2
TK1	1.65920801	0.00934567	Thymidine kinase, cytosolic
DVL2	-1.2621462	0.00954503	Segment polarity protein dishevelled homolog DVL-2
CAV1	2.19574733	0.00960817	Caveolin-1
CTSB	1.04506421	0.00960817	Cathepsin B
POC1A	1.35204602	0.00960817	POC1 centriolar protein homolog A
TESK1	-1.0603817	0.00960817	Dual specificity testis-specific protein kinase 1
GLA	-1.4768568	0.00963743	Galactosidase alpha
NDRG2	1.82076044	0.00963743	Protein NDRG2
PGK1	1.47588961	0.00963743	Phosphoglycerate kinase 1
ENTPD3	2.79472615	0.00963995	Ectonucleoside triphosphate diphosphohydrolase 3
TSPAN15	1.26788367	0.00963995	Tetraspanin-15
HSD17B12	-1.1305533	0.00967064	Very-long-chain 3-oxoacyl-CoA reductase
IKBKG	-1.6346706	0.00987048	NF-kappa-B essential modulator
HRCT1	2.81702808	0.00996366	Histidine rich carboxyl terminus 1
SPIRE1	-1.5570037	0.00998224	Protein spire homolog 1
PTRF	1.28980809	0.01000304	Caveolae-associated protein 1
ZFAND2A	-1.650983	0.01005209	Zinc finger AN1-type containing 2A
CCNYL1	-1.6218066	0.01038598	Cyclin-Y-like protein 1
RORA	-2.6196583	0.01038598	Nuclear receptor ROR-alpha
MICALL1	-1.1135816	0.01047901	MICAL-like protein 1
FXYD3	1.72420144	0.01061093	FXYD domain-containing ion transport regulator 3
SYT8	1.96334493	0.01069746	Synaptotagmin-8
WDR54	1.90958723	0.01070198	WD repeat-containing protein 54
FANCE	-1.5179216	0.01084558	Fanconi anemia group E protein
HOMER1	-1.3991157	0.01084558	Homer protein homolog 1
ABHD3	-1.2814496	0.01102335	Phospholipase ABHD3
SMAP2	1.34017054	0.01102335	Stromal membrane-associated protein 2
DDIT4	-1.0855691	0.0111955	DNA damage-inducible transcript 4 protein
ELK3	1.91918941	0.01119622	ETS domain-containing protein Elk-3
FKBP11	-1.4076012	0.01133318	Peptidyl-prolyl cis-trans isomerase FKBP11
NEK6	1.56748093	0.01139972	Serine/threonine-protein kinase Nek6
ZDHHC11B	2.32469433	0.01139972	Probable palmitoyltransferase ZDHHC11B
HKDC1	-1.7524055	0.01172378	Putative hexokinase HKDC1
ARG2	-2.2667069	0.01192528	Arginase-2, mitochondrial
IDI1	-1.3060234	0.01216747	Isopentenyl-diphosphate Delta-isomerase 1
ADAM28	2.21293783	0.01239625	Disintegrin and metalloproteinase domain-containing protein 28
GSTA4	-1.2887897	0.01250719	Glutathione S-transferase A4
ZNF75D	-1.4223247	0.01250719	Zinc finger protein 75D

BTN3A2	1.49287089	0.01263422	Butyrophilin subfamily 3 member A2
CDC45	1.71171581	0.0128686	Cell division control protein 45 homolog
TSPAN3	1.15438902	0.01305074	Tetraspanin-3
MKI67	1.23102736	0.01342154	Proliferation marker protein Ki-67
EIF4EBP1	-1.7552738	0.01346205	Eukaryotic translation initiation factor 4E-binding protein 1
UCP2	1.17880706	0.01351141	Mitochondrial uncoupling protein 2
BAMBI	-1.3527186	0.01375445	BMP and activin membrane-bound inhibitor homolog
PRSS2	-2.7347699	0.01407737	Trypsin-2
PDZD7	1.63098224	0.01418281	PDZ domain-containing protein 7
RPL22L1	-1.2232017	0.01438104	Ribosomal protein L22 like 1
JSRP1	-2.5547113	0.01438485	Junctional sarcoplasmic reticulum protein 1
DSCC1	1.6142658	0.01441584	Sister chromatid cohesion protein DCC1
PGM3	-1.4474143	0.01451474	Phosphoacetylglucosamine mutase
ACSS1	1.78944872	0.01470219	Acetyl-coenzyme A synthetase 2-like, mitochondrial
APBA3	-1.2872093	0.01470219	Amyloid-beta A4 precursor protein-binding family A member 3
BTG1	-1.2029012	0.01470219	Protein BTG1
CCL28	1.56858714	0.01470219	C-C motif chemokine 28
SRXN1	-1.8171151	0.01470219	Sulfiredoxin-1
TNFAIP2	1.92434072	0.01476791	Tumor necrosis factor alpha-induced protein 2
ARPIN	1.24021344	0.01481683	Arpin
TTC30A	1.61449815	0.01605557	Tetratricopeptide repeat protein 30A
MYO7A	-1.8046946	0.01608513	Unconventional myosin-VIIa
SPOCD1	2.48240788	0.01627881	Protein phosphatase 1 regulatory subunits
ERMP1	1.049112	0.01639811	Endoplasmic reticulum metallopeptidase 1
LBHD1	-1.3984396	0.01639811	LBH domain-containing protein 1
IL22RA1	1.16876389	0.01642884	Interleukin-22 receptor subunit alpha-1
TMEM97	-1.4653434	0.01706458	Sigma intracellular receptor 2
ZNF367	1.62263953	0.01721461	Zinc finger protein 367
VWA1	1.54719659	0.01722318	Von Willebrand factor A domain-containing protein 1
ZNF165	-1.5606505	0.01726534	Zinc finger protein 165
OSGIN1	-1.7275056	0.01752775	Oxidative stress-induced growth inhibitor 1
DUSP5	-1.6133407	0.01755526	Dual specificity protein phosphatase 5
GCH1	-1.6674826	0.01755526	GTP cyclohydrolase 1
LPIN1	-1.3259738	0.01755526	Phosphatidate phosphatase LPIN1
SMCO4	1.21789779	0.01755526	Single-pass membrane protein with coiled-coil domains 4
ST6GALNAC3	-2.0852411	0.01755526	Alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase 3
UPP1	-1.6043184	0.01773073	Uridine phosphorylase 1
BCO1	2.0029326	0.01782408	Beta,beta-carotene 15,15'-dioxygenase
ELL2	-1.3264201	0.01821005	RNA polymerase II elongation factor ELL2

FARP2	1.0866604	0.01848739	FERM, ARHGEF and pleckstrin domain-containing protein 2
SLC44A3	1.21318658	0.01864396	Choline transporter-like protein 3
CTTNBP2	2.51024197	0.01899818	Cortactin-binding protein 2
REXO2	-1.0737898	0.01899818	Oligoribonuclease, mitochondrial
ZNF816	1.19803097	0.01899818	Zinc finger protein 816
IRS2	-1.2919784	0.01908452	Insulin receptor substrate 2
EMP3	-1.6317882	0.01920245	Epithelial membrane protein 3
FGD2	2.47017878	0.01920245	FYVE, RhoGEF and PH domain-containing protein 2
GHITM	-1.0318716	0.01920245	Growth hormone-inducible transmembrane protein
CKMT1A	1.96013731	0.01971602	Creatine kinase U-type, mitochondrial
TRAK2	-1.0457877	0.01971602	Trafficking kinesin-binding protein 2
ATP10B	1.86222051	0.0199172	Probable phospholipid-transporting ATPase VB
IDE	-1.1066339	0.02008867	Insulin-degrading enzyme
GPRC5C	-1.0786415	0.02091416	G-protein coupled receptor family C group 5 member C
MGST1	-1.3086008	0.02091416	Microsomal glutathione S-transferase 1
SDPR	2.73281129	0.02119218	Caveolae-associated protein 2
FANCI	1.51708412	0.02152563	Fanconi anemia group I protein
ZDHHC14	-1.4958399	0.02172256	Probable palmitoyltransferase ZDHHC14
TRIM24	-1.2069222	0.02183189	Transcription intermediary factor 1-alpha
GPAT3	-1.5343398	0.02266371	Glycerol-3-phosphate acyltransferase 3
TMEM135	-1.4598288	0.02266371	Transmembrane protein 135
AP5Z1	-1.4909495	0.02269182	AP-5 complex subunit zeta-1
CBX7	1.25722751	0.02269182	Chromobox protein homolog 7
RGS2	1.01037348	0.02269182	Regulator of G-protein signaling 2
OPLAH	-2.0650325	0.02275327	5-oxoprolinase
PCYOX1L	1.12714508	0.02275327	Prenylcysteine oxidase-like
PFDN6	2.7219581	0.02275327	Prefoldin subunit 6
RNF166	-1.1839052	0.02275327	Ring finger protein 166
XPOT	-1.5302228	0.02291706	Exportin-T
ANKRD36C	1.47141184	0.02319802	Ankyrin repeat domain-containing protein 36C
INPP1	-1.0763441	0.02347102	Inositol polyphosphate-1-phosphatase
FTL	-1.1971289	0.02392054	Ferritin light chain
PRELID3A	-1.8938587	0.02454542	PRELI domain containing protein 3A
SPSB1	-1.7700937	0.02467115	SPRY domain-containing SOCS box protein 1
PTGS1	2.34344743	0.0247159	Prostaglandin G/H synthase 1
SPICE1	1.0537163	0.0247159	Spindle and centriole-associated protein 1
TNFRSF11A	1.5712272	0.02490075	Tumor necrosis factor receptor superfamily member 11A
SLC25A36	-1.4566301	0.02583834	Solute carrier family 25 member 36
FUCA1	1.11230385	0.02587358	Tissue alpha-L-fucosidase
BDKRB2	2.13972229	0.02588675	B2 bradykinin receptor

ADK	-1.1760481	0.02595818	Adenosine kinase
STEAP1	-1.0765693	0.02595818	Metalloreductase STEAP1
BBC3	-1.4036836	0.02637879	Bcl-2-binding component 3
RARRES3	1.73460421	0.02637879	Retinoic acid receptor responder protein 3
VIPR1	1.33143782	0.02637879	Vasoactive intestinal polypeptide receptor 1
TAGLN2	1.82276954	0.02647151	Transgelin-2
KIF11	1.22735425	0.02727839	Kinesin-like protein KIF11
C1orf95	2.24815765	0.02748912	Protein stum homolog
YPEL2	1.02405276	0.02748912	Protein yippee-like 2
TULP3	-1.3251351	0.02794287	Tubby-related protein 3
GPI	1.18707924	0.02834053	Glucose-6-phosphate isomerase
LIPE	1.82865216	0.02848869	Hormone-sensitive lipase
CCPG1	-1.114795	0.02917011	Cell cycle progression protein 1
HM13	-1.0014668	0.02917011	Histocompatibility minor 13
ACSL4	-1.0348614	0.02976647	Long-chain-fatty-acid--CoA ligase 4
DUSP1	-1.648853	0.03010818	Dual specificity protein phosphatase 1
CDKN2D	1.45271135	0.03110609	Cyclin-dependent kinase 4 inhibitor D
MPC1	-1.7934791	0.03288939	Mitochondrial pyruvate carrier 1
BRSK1	-1.3411429	0.03314883	Serine/threonine-protein kinase BRSK1
ITPKC	-1.470021	0.03314883	Inositol-trisphosphate 3-kinase C
MAN2B2	1.20779986	0.03323273	Epididymis-specific alpha-mannosidase
PGM1	1.03883618	0.03334266	Phosphoglucomutase-1
AKNA	-1.2333908	0.03387208	AT-hook-containing transcription factor
GPR155	2.04530446	0.03387208	Integral membrane protein GPR155
HCFC1R1	1.4591255	0.03387208	Host cell factor C1 regulator 1
MYCBP	1.24109924	0.03387208	c-Myc-binding protein
PODXL	1.81287641	0.03387208	Podocalyxin
FHAD1	2.30985728	0.03410165	Forkhead-associated domain-containing protein 1
PKD1	-1.0468938	0.03410165	Polycystin-1
TFE3	-1.1219534	0.03410165	Transcription factor E3
ASF1B	1.48825513	0.03424347	Histone chaperone ASF1B
ZNF354A	-1.2529258	0.03424347	Zinc finger protein 354A
HSPA9	-1.1544622	0.03475997	Stress-70 protein, mitochondrial
CPEB4	-1.2077394	0.0354978	Cytoplasmic polyadenylation element-binding protein 4
CCDC64	-1.5350208	0.03563314	BICD family-like cargo adapter 1
CDKN1A	-1.1408441	0.03591448	Cyclin-dependent kinase inhibitor 1
THBS1	1.28896905	0.03614375	Thrombospondin-1
EPT1	-1.3204681	0.03646195	Ethanolaminephosphotransferase 1
AQP5	1.88973363	0.03706801	Aquaporin-5
CLMN	1.51708105	0.03706801	Calmin
MLXIPL	-1.7498756	0.03706801	Carbohydrate-responsive element-binding protein

PDIA5	-1.0719846	0.03750019	Protein disulfide-isomerase A5
BCAS1	2.21176851	0.03754167	Breast carcinoma amplified sequence 1
CCND3	1.41355191	0.03754167	G1/S-specific cyclin-D3
CDK2AP2	-1.0594077	0.03754167	Cyclin-dependent kinase 2-associated protein 2
MPV17L2	-1.0862253	0.03757517	Mpv17-like protein 2
MSX1	-2.5197616	0.03781037	Homeobox protein MSX-1
NXPH4	1.94959245	0.03786942	Neurexophilin-4
DMGDH	-2.4809887	0.03788515	Dimethylglycine dehydrogenase, mitochondrial
NCF2	-2.3625203	0.03790419	Neutrophil cytosol factor 2
AZGP1	1.6826537	0.0380166	Zinc-alpha-2-glycoprotein
BIRC7	1.23883132	0.0380166	Baculoviral IAP repeat-containing protein 7
NFE2L1	-1.3581337	0.0380166	Nuclear factor erythroid 2-related factor 1
TSPAN14	1.04785115	0.0380166	Tetraspanin-14
GULP1	-1.3133122	0.03853416	PTB domain-containing engulfment adapter protein 1
TNFRSF14	1.64580832	0.03875729	Tumor necrosis factor receptor superfamily member 14
VIMP	-1.3262874	0.03875729	Selenoprotein S
ENO2	1.45263029	0.03897777	Gamma-enolase
ACP7	-2.3354311	0.03918529	Iron/zinc purple acid phosphatase-like protein
AGA	-1.1037376	0.03918529	N(4)-(beta-N-acetylglucosaminyl)-L-asparaginase
PLS1	1.02225193	0.0398763	Plastin-1
GPAT2	2.51710544	0.04049434	Glycerol-3-phosphate acyltransferase 2, mitochondrial
CLCF1	-1.0170746	0.04076689	Cardiotrophin-like cytokine factor 1
DNAJB9	-1.6283524	0.04076689	DnaJ homolog subfamily B member 9
EPB41L2	1.25173004	0.04076689	Band 4.1-like protein 2
GALNT18	1.77516463	0.04076689	Polypeptide N-acetylgalactosaminyltransferase 18
GNG7	2.28937972	0.04076689	Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-7
JAK3	1.77422043	0.04076689	Tyrosine-protein kinase JAK3
JMJD6	-1.228932	0.04076689	Bifunctional arginine demethylase and lysyl-hydroxylase JMJD6
KREMEN1	1.13452728	0.04076689	Kremen protein 1
LRRC59	-1.3003194	0.04076689	Leucine-rich repeat-containing protein 59
MS4A8	2.05806845	0.04076689	Membrane-spanning 4-domains subfamily A member 8
PLEKHB1	1.19783488	0.04076689	Pleckstrin homology domain-containing family B member 1
SSBP2	1.73284186	0.04076689	Single stranded DNA binding protein 2
TOP2A	1.24544961	0.04076689	DNA topoisomerase 2-alpha
SCRIB	2.13107893	0.04101671	Protein scribble homolog
SPC24	1.6247927	0.04101671	Kinetochore protein Spc24
CHPF	-1.2679855	0.04240742	Chondroitin sulfate synthase 2
ADGRF1	2.03745477	0.04251519	Adhesion G-protein coupled receptor F1

C6orf48	2.49181294	0.04251519	Protein G8
ADD3	1.03435119	0.042998	Gamma-adducin
TMEM41B	-1.1816583	0.0439311	Transmembrane protein 41B
MSMO1	-1.2022029	0.04399134	Methylsterol monooxygenase 1
TMCC3	1.97515684	0.04403955	Transmembrane and coiled-coil domain containing
SPTBN5	-1.1626611	0.04468792	Spectrin beta, non-erythrocytic 5
TUBB	2.38528658	0.04468792	Tubulin beta chain
NUSAP1	1.06653811	0.04480427	Nucleolar and spindle-associated protein 1
SETD1B	-1.1065257	0.04487637	Histone-lysine N-methyltransferase SETD1B
LIMCH1	2.01213415	0.04526884	LIM and calponin homology domains 1
NTF4	2.11733253	0.04544685	Neurotrophin-4
FAM111B	1.98484909	0.04645227	Protein FAM111B
DHFRL1	1.30374549	0.04714182	Dihydrofolate reductase 2, mitochondrial
LDHA	1.37157635	0.04714182	Lactate dehydrogenase A
GTPBP2	-1.2789313	0.04731042	GTP binding protein 2
CYSRT1	1.92521095	0.04740395	Cysteine rich tail 1

Appendix table 11. Human differentially expressed genes between uninfected controls and infected samples at 3 h post infection.

Gene name	**log_2(fold change)**	**padj**	**annotation**
EPC2	-1.5315562	3.55E-14	Enhancer of polycomb homolog 2
TMEM169	3.09764036	4.80E-14	Transmembrane protein 169
TNFRSF11B	-2.3793284	1.43E-13	Tumor necrosis factor receptor superfamily member 11B
KAT5	-1.3089709	6.19E-13	Histone acetyltransferase KAT5
FGG	-1.8976023	6.63E-13	Fibrinogen gamma chain
KEAP1	-1.2644894	1.49E-12	Kelch-like ECH-associated protein 1
EPC1	-1.394593	3.14E-12	Enhancer of polycomb homolog 1
ZCCHC3	-1.0026955	7.05E-12	Zinc finger CCHC-type containing 3
JARID2	1.07119071	8.76E-11	Protein Jumonji
TMEM185B	-1.117556	8.76E-11	Transmembrane protein 185B
DUS3L	-1.7298327	1.11E-10	tRNA-dihydrouridine(47) synthase [NAD(P)(+)]-like
XBP1	-1.0952191	1.11E-10	X-box-binding protein 1
LUZP1	1.08255482	3.39E-10	Leucine zipper protein 1
BEND3	-1.3734481	8.24E-10	BEN domain-containing protein 3
SRSF6	-1.0915863	8.24E-10	Serine/arginine-rich splicing factor 6
FEM1A	-1.2249469	1.08E-09	Protein fem-1 homolog A
UBIAD1	-1.7505287	1.15E-09	UbiA prenyltransferase domain-containing protein 1

LANCL2	-1.6265345	1.24E-09	LanC-like protein 2
NCOA5	-1.3556052	1.34E-09	Nuclear receptor coactivator 5
FAM101A	-2.0707479	1.42E-09	Refilin-A
UGT2B17	-2.6683489	2.03E-09	UDP-glucuronosyltransferase 2B17
ARMC5	-1.8981879	2.52E-09	Armadillo repeat-containing protein 5
SLC3A2	-1.2073558	4.34E-09	4F2 cell-surface antigen heavy chain
FKRP	-1.5353278	5.18E-09	Fukutin-related protein
PTCH1	1.26400476	6.02E-09	Protein patched homolog 1
TRIB3	-1.4058143	6.14E-09	Tribbles homolog 3
FGB	-2.0361994	9.19E-09	Fibrinogen beta chain
TANC1	1.07305701	1.17E-08	Protein TANC1
BCL6	1.08706584	1.23E-08	B-cell lymphoma 6 protein
NCEH1	1.19503544	2.27E-08	Neutral cholesterol ester hydrolase 1
FGA	-2.099032	2.49E-08	Fibrinogen alpha chain
NFKB1	-1.3457026	4.91E-08	Nuclear factor NF-kappa-B p105 subunit
ING3	-1.341835	5.07E-08	Inhibitor of growth protein 3
ZNF608	1.37944921	7.08E-08	Zinc finger protein 608
C9orf156	-1.2101141	7.90E-08	tRNA (adenine(37)-N6)-methyltransferase
GVQW3	1.82722387	1.04E-07	annotation not available
RABIF	-1.2944558	1.31E-07	Guanine nucleotide exchange factor MSS4
MCRS1	-1.0058878	1.40E-07	Microspherule protein 1
TOE1	-1.498355	2.10E-07	Target of EGR1 protein 1
TMEM9B	-1.1101945	2.51E-07	Transmembrane protein 9B
HOXA1	1.77078234	2.55E-07	Homeobox protein Hox-A1
CNN2	1.20441162	3.10E-07	Calponin-2
LRRC20	1.66857639	3.10E-07	Leucine-rich repeat-containing protein 20
TRIM38	1.48031614	4.92E-07	E3 ubiquitin-protein ligase TRIM38
YRDC	-1.137583	6.89E-07	YrdC domain-containing protein, mitochondrial
GSDMB	1.91736727	7.95E-07	Gasdermin-B
TSPYL4	1.06812713	8.47E-07	Testis-specific Y-encoded-like protein 4
USP35	1.32825226	9.14E-07	Ubiquitin specific peptidase 35
SRSF2	-1.4197949	9.14E-07	Serine/arginine-rich splicing factor 2
ARMC6	-1.6061946	1.20E-06	Armadillo repeat containing 6
HK1	1.9236167	1.27E-06	Hexokinase-1
PDX1	1.44095394	1.27E-06	Pancreas/duodenum homeobox protein 1
TBC1D9	1.25011735	1.42E-06	TBC1 domain family member 9
ZNF202	-1.3708302	1.51E-06	Zinc finger protein 202
GIPR	2.65275773	1.58E-06	Gastric inhibitory polypeptide receptor
MTHFD2	-1.5881027	1.58E-06	Bifunctional methylenetetrahydrofolate dehydrogenase/cyclohydrolase, mitochondrial

AP5S1	-1.4585057	1.58E-06	AP-5 complex subunit sigma-1
IBA57	-1.2311805	1.58E-06	Putative transferase CAF17, mitochondrial
STEAP2	1.19085832	1.69E-06	Metalloreductase STEAP2
MRPL15	-1.3565599	2.12E-06	Mitochondrial ribosomal protein L15
PUS1	-1.2139277	2.28E-06	tRNA pseudouridine synthase A
THSD4	-1.9305159	2.34E-06	Thrombospondin type-1 domain-containing protein 4
ZNF689	-1.0575266	2.67E-06	Zinc finger protein 689
DMXL2	1.26872916	2.84E-06	DmX-like protein 2
AMOT	1.23014298	3.18E-06	Angiomotin
CECR2	1.77670977	3.18E-06	Cat eye syndrome critical region protein 2
DBR1	-1.1507891	3.18E-06	Lariat debranching enzyme
TUBA1C	-1.1467719	3.18E-06	Tubulin alpha-1C chain
NMNAT1	-1.8887652	3.36E-06	Nicotinamide/nicotinic acid mononucleotide adenylyltransferase 1
DUSP2	-2.0321483	4.00E-06	Dual specificity protein phosphatase 2
BTN2A2	2.11560253	4.00E-06	Butyrophilin subfamily 2 member A2
SRGAP2	1.91649385	4.33E-06	SLIT-ROBO Rho GTPase-activating protein 2
GFOD2	1.0304406	4.61E-06	Glucose-fructose oxidoreductase domain-containing protein 2
ZFP64	-1.0888105	4.61E-06	Zinc finger protein 64 homolog, isoforms 1 and 2
FBXL3	-1.1224656	5.12E-06	F-box/LRR-repeat protein 3
PPRC1	-1.1569603	5.14E-06	Peroxisome proliferator-activated receptor gamma coactivator-related protein 1
C19orf26	1.19297312	6.16E-06	Voltage-dependent calcium channel beta subunit-associated regulatory protein
NOP16	-1.4739016	6.19E-06	NOP16 nucleolar protein
RBM3	-1.0525859	6.19E-06	RNA-binding protein 3
ZCCHC24	1.9298218	6.92E-06	Zinc finger CCHC-type containing 24
ZNF784	-1.9448308	7.53E-06	Zinc finger protein 784
COA7	-1.2812635	7.55E-06	Cytochrome c oxidase assembly factor 7
MYC	-1.1734082	7.59E-06	Myc proto-oncogene protein
SERPINB8	-2.4380788	8.67E-06	Serpin B8
TSHZ1	1.42608203	8.90E-06	Teashirt homolog 1
HOXB13	1.75729028	9.37E-06	Homeobox protein Hox-B13
CHMP4C	-1.054015	9.48E-06	Charged multivesicular body protein 4c
MEPCE	-1.3332797	9.91E-06	7SK snRNA methylphosphate capping enzyme
NDUFAF4	-1.5832766	1.05E-05	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex assembly factor 4
LMCD1	1.53605998	1.06E-05	LIM and cysteine-rich domains protein 1
AAR2	-1.1323541	1.11E-05	Protein AAR2 homolog
CCDC94	-1.1254111	1.14E-05	Coiled-coil domain containing 94
EPM2A	-1.6730545	1.23E-05	Laforin

L3MBTL3	1.07361322	1.27E-05	Lethal(3)malignant brain tumor-like protein 3
MVK	-1.0189971	1.32E-05	Mevalonate kinase
PLEKHM3	1.02132606	1.34E-05	Pleckstrin homology domain containing M3
ZNF518A	-1.2301038	1.35E-05	Zinc finger protein 518A
SLC7A1	-1.4234323	1.58E-05	High affinity cationic amino acid transporter 1
UBXN8	-1.3849848	1.58E-05	UBX domain-containing protein 8
METTL1	-1.5283668	1.90E-05	tRNA (guanine-N(7)-)-methyltransferase
SULF2	1.37419453	1.90E-05	Extracellular sulfatase Sulf-2
TAF5L	-1.2385981	1.90E-05	TAF5-like RNA polymerase II p300/CBP-associated factor-associated factor 65 kDa subunit 5L
KCTD15	-1.105059	1.95E-05	BTB/POZ domain-containing protein KCTD15
ATXN7L2	-1.801636	1.98E-05	Ataxin-7-like protein 2
CYTH2	-1.1699634	2.06E-05	Cytohesin-2
MAP2K4	-1.0567188	2.06E-05	Dual specificity mitogen-activated protein kinase kinase 4
KIAA1377	1.04099415	2.18E-05	Centrosomal protein of 126 kDa
TMEM63C	1.45683429	2.24E-05	Calcium permeable stress-gated cation channel 1
NR6A1	1.04605075	2.30E-05	Nuclear receptor subfamily 6 group A member 1
ZNF181	-1.1484566	2.30E-05	Zinc finger protein 181
MRPL54	-1.0902838	2.30E-05	Mitochondrial ribosomal protein L54
SLC2A8	-1.1191743	2.33E-05	Solute carrier family 2, facilitated glucose transporter member 8
GADD45A	1.49341659	2.34E-05	Growth arrest and DNA damage-inducible protein GADD45 alpha
HOXA13	1.35753347	2.38E-05	Homeobox protein Hox-A13
TFAP2C	2.7614553	2.38E-05	Transcription factor AP-2 gamma
NIPAL3	1.09963592	2.43E-05	NIPA-like protein 3
PRDM15	-1.0918093	2.90E-05	PR domain zinc finger protein 15
ZNF180	-1.0126877	3.09E-05	Zinc finger protein 180
DHX37	-1.1952011	3.10E-05	Probable ATP-dependent RNA helicase DHX37
JUNB	-1.4158213	3.21E-05	Transcription factor jun-B
BCORL1	1.10545656	3.24E-05	BCL-6 corepressor-like protein 1
ESAM	1.61505589	3.26E-05	Endothelial cell-selective adhesion molecule
CYP20A1	-1.0767633	3.46E-05	Cytochrome P450 family 20 subfamily A member 1
EME2	-1.3175066	3.47E-05	Probable crossover junction endonuclease EME2
RBMS2	1.11170804	3.55E-05	RNA-binding motif, single-stranded-interacting protein 2
DIEXF	-1.0405504	3.62E-05	Digestive organ expansion factor homolog
NSMAF	-1.0925482	3.76E-05	Protein FAN
POLG2	1.20908851	3.79E-05	DNA polymerase subunit gamma-2, mitochondrial
PRDM10	-1.1327401	3.79E-05	PR domain zinc finger protein 10

MBD3	-1.1261289	3.79E-05	Methyl-CpG-binding domain protein 3
MED12L	1.9726729	3.84E-05	Mediator of RNA polymerase II transcription subunit 12-like protein
SERPINB9	1.42862712	4.15E-05	Serpin B9
ZNF566	-1.1222879	4.15E-05	Zinc finger protein 566
UAP1L1	1.09494369	5.10E-05	UDP-N-acetylhexosamine pyrophosphorylase-like protein 1
LNX1	1.19335437	5.35E-05	E3 ubiquitin-protein ligase LNX
MMP1	-3.3782354	5.55E-05	Interstitial collagenase
URB2	-1.4296969	6.12E-05	URB2 ribosome biogenesis 2 homolog
ZDHHC23	-1.4241707	6.38E-05	Palmitoyltransferase ZDHHC23
C1orf122	-1.3792896	7.49E-05	Uncharacterized protein C1orf122
PHLDB2	1.16270831	8.46E-05	Pleckstrin homology-like domain family B member 2
RRS1	-1.3269283	8.51E-05	Ribosome biogenesis regulatory protein homolog
SLC25A29	1.65380137	8.57E-05	Mitochondrial basic amino acids transporter
CCDC136	1.91519069	8.93E-05	Coiled-coil domain-containing protein 136
ZNF555	1.15702712	8.93E-05	Zinc finger protein 555
CCDC144A	1.02882574	9.07E-05	Coiled-coil domain containing 144A
FIZ1	-1.1008899	9.16E-05	Flt3-interacting zinc finger protein 1
DPH1	-1.0127973	9.16E-05	2-(3-amino-3-carboxypropyl)histidine synthase subunit 1
NUMBL	1.29630061	9.29E-05	Numb-like protein
WDR19	1.15116221	9.39E-05	WD repeat-containing protein 19
KLHL21	-1.2190175	0.00010074	Kelch-like protein 21
FAIM	-1.3529565	0.00010462	Fas apoptotic inhibitory molecule 1
SLC6A12	2.01504247	0.00010631	Sodium- and chloride-dependent betaine transporter
TMED1	-1.0692503	0.00010739	Transmembrane emp24 domain-containing protein 1
ZXDB	-1.0310063	0.00011525	Zinc finger X-linked protein ZXDB
ZBTB25	1.21778018	0.00011666	Zinc finger and BTB domain-containing protein 25
FABP2	-1.684705	0.00011898	Fatty acid-binding protein, intestinal
FDXACB1	-2.0417677	0.00011929	Ferredoxin-fold anticodon-binding domain-containing protein 1
FAM160A1	1.0199726	0.00012286	Protein FAM160A1
IRF8	1.22944021	0.00012355	Interferon regulatory factor 8
PSTK	-1.5612471	0.00012425	L-seryl-tRNA(Sec) kinase
RASA2	-1.5813302	0.0001328	Ras GTPase-activating protein 2
TMEM184A	-1.14239	0.0001328	Transmembrane protein 184A
HFE	1.24432207	0.00013644	Hereditary hemochromatosis protein
PPM1K	1.30117377	0.00013644	Protein phosphatase 1K, mitochondrial
NDRG4	1.58201035	0.00013736	Protein NDRG4
SRSF7	-1.1342227	0.00014458	Serine/arginine-rich splicing factor 7
EFNA3	2.62914657	0.00016014	Ephrin-A3

PLD6	-1.3671918	0.00016381	Mitochondrial cardiolipin hydrolase
SIRT1	-1.0858921	0.00016526	NAD-dependent protein deacetylase sirtuin-1
METTL25	-1.3229905	0.00017151	Methyltransferase-like protein 25
NINJ1	-1.2743104	0.00017675	Ninjurin-1
PRR22	-1.828226	0.00018032	Proline-rich protein 22
CDH24	1.06509436	0.00018417	Cadherin-24
LCA5	1.24190997	0.00018417	Lebercilin
PHF21A	1.3831329	0.00018417	PHD finger protein 21A
ENPP5	1.54014322	0.00018435	Ectonucleotide pyrophosphatase/phosphodiesterase family member 5
ZNF41	-1.0985221	0.00018807	Zinc finger protein 41
RASSF4	1.50409126	0.00018835	Ras association domain-containing protein 4
USP2	1.85630323	0.0002023	Ubiquitin carboxyl-terminal hydrolase 2
STPG1	1.42647618	0.00020333	O(6)-methylguanine-induced apoptosis 2
SNAI3	2.59527952	0.00022208	Zinc finger protein SNAI3
RILPL2	1.50397539	0.00022444	RILP-like protein 2
GDAP1	-1.0665528	0.00022444	Ganglioside-induced differentiation-associated protein 1
CD3EAP	-1.62297	0.00023089	DNA-directed RNA polymerase I subunit RPA34
ZNF221	1.4505877	0.00023668	Zinc finger protein 221
FOXA1	-1.383907	0.00024264	Hepatocyte nuclear factor 3-alpha
PLEKHF1	1.9088662	0.00024896	Pleckstrin homology domain-containing family F member 1
MEIS2	1.1831094	0.00025	Homeobox protein Meis2
ZNF385A	1.62165339	0.00025396	Zinc finger protein 385A
DPH2	-1.4183907	0.00025988	2-(3-amino-3-carboxypropyl)histidine synthase subunit 2
HSPA4L	1.26793923	0.00026145	Heat shock 70 kDa protein 4L
MBOAT2	1.33502593	0.00026145	Lysophospholipid acyltransferase 2
SLC2A13	1.42164618	0.00026327	Proton myo-inositol cotransporter
CACNG8	2.40633926	0.00026998	Voltage-dependent calcium channel gamma-8 subunit
FASTKD2	-1.1565758	0.00028056	FAST kinase domain-containing protein 2, mitochondrial
PNO1	-1.2451663	0.00028257	RNA-binding protein PNO1
CPLX1	-1.6304361	0.00029019	Complexin-1
PRR11	1.23795318	0.00029537	Proline-rich protein 11
RAB3A	3.43026657	0.00029879	Ras-related protein Rab-3A
ANKZF1	1.29378538	0.00030574	Ankyrin repeat and zinc finger domain-containing protein 1
SUFU	1.01930287	0.00030763	Suppressor of fused homolog
C1orf106	-1.045107	0.00032402	Innate immunity activator protein
SP2	-1.1275907	0.00032605	Transcription factor Sp2
PDE5A	2.01319597	0.00034455	cGMP-specific 3',5'-cyclic phosphodiesterase

ANKRD42	1.09866347	0.0003597	Ankyrin repeat domain-containing protein 42
CCND2	1.0263409	0.00036053	G1/S-specific cyclin-D2
NDNF	1.60914458	0.00036053	Protein NDNF
C9orf91	-1.2322623	0.00036348	Transmembrane protein 268
KIF3C	1.40846169	0.00037634	Kinesin-like protein KIF3C
ZC3H12C	-1.0795311	0.0003994	Probable ribonuclease ZC3H12C
MAP1A	1.69966102	0.00043123	Microtubule-associated protein 1A
CLCF1	-2.0760567	0.00043387	Cardiotrophin-like cytokine factor 1
WDR4	-1.2069691	0.00044281	tRNA (guanine-N(7)-)-methyltransferase non-catalytic subunit WDR4
EPN3	1.89252452	0.00044281	Epsin-3
DOHH	-1.0413415	0.00044281	Deoxyhypusine hydroxylase
MEGF9	1.01504483	0.00044513	Multiple epidermal growth factor-like domains protein 9
ZC3HC1	-1.113271	0.00046481	Nuclear-interacting partner of ALK
DIMT1	-1.2084396	0.00046908	Probable dimethyladenosine transferase
TP53RK	-1.067324	0.00046908	TP53-regulating kinase
CITED2	1.37103921	0.00050637	Cbp/p300-interacting transactivator 2
PAQR8	1.15425074	0.00052549	Membrane progestin receptor beta
FGFBP1	-2.8189441	0.00052709	Fibroblast growth factor-binding protein 1
TFB2M	-1.2145596	0.00053087	Dimethyladenosine transferase 2, mitochondrial
ZSWIM3	-1.1380394	0.00053087	Zinc finger SWIM-type containing 3
NOXO1	-1.8073548	0.00054491	NADPH oxidase organizer 1
SH3PXD2B	1.38100862	0.00057045	SH3 and PX domain-containing protein 2B
MAML2	1.06734745	0.0005855	Mastermind-like protein 2
BRMS1L	-1.3219032	0.00058785	Breast cancer metastasis-suppressor 1-like protein
HCAR1	1.30111514	0.00059188	Hydroxycarboxylic acid receptor 1
ADAMTS6	1.87622314	0.00060171	ADAM metallopeptidase with thrombospondin type 1 motif 6
ARL4D	1.50912233	0.00060171	ADP-ribosylation factor-like protein 4D
N6AMT1	-1.0049587	0.00060207	HemK methyltransferase family member 2
PSAPL1	-1.4539153	0.00063177	Proactivator polypeptide-like 1
AATK	-1.1288403	0.00063914	Serine/threonine-protein kinase LMTK1
CTTNBP2NL	1.05306379	0.00068878	CTTNBP2 N-terminal-like protein
CALCOCO1	1.26455412	0.0006909	Calcium-binding and coiled-coil domain-containing protein 1
SH2B2	1.57776499	0.00074112	SH2B adapter protein 2
RFX2	1.73740996	0.00075106	DNA-binding protein RFX2
AADAC	-1.1846016	0.00079363	Arylacetamide deacetylase
SH3BP1	1.53430865	0.0007966	SH3 domain-binding protein 1
SRM	-1.4624939	0.00083301	Spermidine synthase
TMEM216	-1.145314	0.00083612	Transmembrane protein 216
ZNF485	-1.2984322	0.00085052	Zinc finger protein 485

IRS2	-1.0021489	0.00085163	Insulin receptor substrate 2
SPTBN5	-1.6553505	0.0008522	Spectrin beta, non-erythrocytic 5
IER2	-1.0236873	0.00086125	Immediate early response gene 2 protein
SGK3	-1.0042459	0.00087442	Serum/glucocorticoid regulated kinase family, member 3
SCN1B	2.1999662	0.00087825	Sodium channel subunit beta-1
LRIF1	-1.0493685	0.00088769	Ligand-dependent nuclear receptor-interacting factor 1
DMRTA2	-1.2572668	0.00093349	Doublesex- and mab-3-related transcription factor A2
ST3GAL5	2.21640857	0.0009568	Lactosylceramide alpha-2,3-sialyltransferase
AP5Z1	-1.0069439	0.0009724	AP-5 complex subunit zeta-1
PPIL1	-1.1667733	0.00097484	Peptidyl-prolyl cis-trans isomerase-like 1
NFIA	1.16486375	0.00100675	Nuclear factor 1 A-type
PLCG2	1.82210029	0.00103291	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase gamma-2
EOGT	-1.3138516	0.00103847	EGF domain-specific O-linked N-acetylglucosamine transferase
LRP4	1.4606949	0.00106449	Low-density lipoprotein receptor-related protein 4
FBXL16	1.60876788	0.00107149	F-box/LRR-repeat protein 16
TSEN15	-1.1796331	0.00107175	tRNA-splicing endonuclease subunit Sen15
HPS3	1.08976507	0.00109028	Hermansky-Pudlak syndrome 3 protein
TMEM217	-2.3302895	0.00111208	Transmembrane protein 217
ADH4	-1.3557481	0.00111671	Alcohol dehydrogenase 4, pi polypeptide
PLCXD2	1.28875631	0.00111671	PI-PLC X domain-containing protein 2
MXI1	1.35941301	0.00113663	Max-interacting protein 1
STARD13	1.10253557	0.00115106	StAR-related lipid transfer protein 13
ZNF439	1.09688598	0.00126346	Zinc finger protein 439
CDON	1.28940486	0.00128072	Cell adhesion molecule-related/down-regulated by oncogenes
FAM110A	1.38208355	0.00131172	Protein FAM110A
HNF1A	-1.0038686	0.00131274	Hepatocyte nuclear factor 1-alpha
SIX4	1.36094128	0.00133028	Homeobox protein SIX4
SALL2	1.72221766	0.0013583	Sal-like protein 2
ETV5	1.21519684	0.00136752	ETS translocation variant 5
PPP1R9B	1.04388315	0.00136752	Neurabin-2
HIP1	1.42709574	0.00137477	Huntingtin-interacting protein 1
BAI2	2.39474114	0.00138117	Adhesion G protein-coupled receptor B2
DGKA	1.25713269	0.00140362	Diacylglycerol kinase alpha
TCEANC2	1.2573215	0.00140362	Transcription elongation factor A N-terminal and central domain containing 2
ZNF628	-1.2578376	0.00140396	Zinc finger protein 628
CEACAM18	-1.1502734	0.00144255	Carcinoembryonic antigen related cell adhesion molecule family
MAK16	-1.1321909	0.00148842	Protein MAK16 homolog

HOXC6	1.07022607	0.00151509	Homeobox protein Hox-C6
NOL6	-1.0568011	0.00151904	Nucleolar protein 6
GPR157	1.28975017	0.00151904	G-protein coupled receptor 157
YDJC	-1.0602388	0.00159939	Carbohydrate deacetylase
HIST1H4J	-1.8557767	0.00161409	Histone cluster 1 H4 family member j
LDLR	-1.0236325	0.00163118	Low-density lipoprotein receptor
SHF	1.58759294	0.00165988	SH2 domain-containing adapter protein F
SLC7A11	-1.4925692	0.00166027	Cystine/glutamate transporter
STXBP5	-1.0184178	0.0016658	Syntaxin-binding protein 5
GEM	1.40188367	0.00167164	GTP-binding protein GEM
MCAM	1.39547781	0.00167164	Cell surface glycoprotein MUC18
KIAA0513	1.03759897	0.00168532	Uncharacterized protein KIAA0513
RCN1	-1.13995	0.00171704	Reticulocalbin-1
PDCL3	-1.0568043	0.00171704	Phosducin-like protein 3
TMEM5	-1.0574836	0.00176188	UDP-D-xylose:ribitol-5-phosphate beta1,4-xylosyltransferase
MAPK8IP1	1.24117159	0.0017755	C-Jun-amino-terminal kinase-interacting protein 1
AEN	-1.0984547	0.00179007	Apoptosis-enhancing nuclease
DBP	1.75355757	0.00180375	D site-binding protein
GATA2	1.67857626	0.00181107	Endothelial transcription factor GATA-2
NOP56	-1.0800552	0.00184581	Nucleolar protein 56
RRP9	-1.0060744	0.00185981	U3 small nucleolar RNA-interacting protein 2
SMAD7	1.08084843	0.00188364	Mothers against decapentaplegic homolog 7
KLHL35	-1.3058523	0.00189398	Kelch-like protein 35
TMEM80	-1.0753183	0.00189398	Transmembrane protein 80
GINS3	-1.414754	0.00189757	GINS complex subunit 3
TNFRSF12A	-1.0254135	0.00189837	Tumor necrosis factor receptor superfamily member 12A
ZNF792	1.00870513	0.00192646	Zinc finger protein 792
PGBD2	-1.080563	0.00197059	piggyBac transposable element derived 2
SHROOM4	1.3697806	0.00203561	Protein Shroom4
SLC39A3	-1.0448337	0.00208696	Zinc transporter ZIP3
SERTAD1	-1.1222217	0.00215165	SERTA domain-containing protein 1
FGF9	1.98001588	0.00228066	Fibroblast growth factor 9
AHDC1	1.19093068	0.00237163	AT-hook DNA-binding motif-containing protein 1
CSRNP3	1.9775869	0.00241767	Cysteine/serine-rich nuclear protein 3
SPN	1.41634014	0.00243768	Leukosialin
FAM149A	1.29909494	0.00250825	Protein FAM149A
FHDC1	1.27102974	0.00252952	FH2 domain containing 1
FAM131B	2.08790225	0.00255423	Protein FAM131B
PREX1	2.17889684	0.00256244	Phosphatidylinositol 3,4,5-trisphosphate-dependent Rac exchanger 1 protein

PKNOX1	-1.1051188	0.00266366	Homeobox protein PKNOX1
BCAS4	1.29391921	0.0027552	Breast carcinoma amplified sequence 4
IFT81	1.04685834	0.00285701	Intraflagellar transport protein 81 homolog
UTP20	-1.383028	0.00295601	Small subunit processome component 20 homolog
PER1	1.13013517	0.00299162	Period circadian protein homolog 1
PHEX	1.02562133	0.00300858	Phosphate-regulating neutral endopeptidase
LBH	1.25147947	0.0030174	Protein LBH
NBN	-1.1136822	0.0030445	Nibrin
NCR3LG1	-1.0374955	0.0030659	Natural cytotoxicity triggering receptor 3 ligand 1
PLCH1	1.01052769	0.00306758	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase eta-1
FOXQ1	-1.126116	0.00309713	Forkhead box protein Q1
CBFA2T3	1.65938373	0.00313471	Protein CBFA2T3
ATP1B2	1.35450753	0.00331868	Sodium/potassium-transporting ATPase subunit beta-2
C3orf80	-2.1196712	0.00345796	Uncharacterized membrane protein C3orf80
NRGN	1.29686864	0.00358051	Neurogranin
FAM122C	1.34136359	0.00360585	Protein FAM122C
ADRA1D	-2.165562	0.00371377	Alpha-1D adrenergic receptor
VPS37D	1.47471135	0.00375835	Vacuolar protein sorting-associated protein 37D
RPP40	-1.4086713	0.00377109	Ribonuclease P protein subunit p40
EGR1	-1.3375478	0.00384272	Early growth response protein 1
PELI3	1.17755135	0.00391534	E3 ubiquitin-protein ligase pellino homolog 3
AHNAK	1.02088392	0.00391807	Neuroblast differentiation-associated protein AHNAK
FYN	1.31400839	0.00391807	Tyrosine-protein kinase Fyn
GOLGA8B	1.00330432	0.003934	Golgin subfamily A member 8B
F3	-1.1400009	0.00394521	Tissue factor
MAP4K2	1.34716441	0.00397176	Mitogen-activated protein kinase kinase kinase kinase 2
GNAZ	1.23498897	0.00403052	Guanine nucleotide-binding protein G(z) subunit alpha
ZNF414	-1.1530038	0.00413649	Zinc finger protein 414
GATA3	1.01408707	0.00416535	Trans-acting T-cell-specific transcription factor GATA-3
KREMEN2	1.97741387	0.00426527	Kremen protein 2
SLC41A1	1.33042922	0.00442724	Solute carrier family 41 member 1
FOXO4	1.30792224	0.00448402	Forkhead box protein O4
IRF1	1.3443794	0.00453529	Interferon regulatory factor 1
TMEM136	1.47306496	0.00458142	Transmembrane protein 136
CACNB1	1.50648376	0.00484073	Voltage-dependent L-type calcium channel subunit beta-1
WDR31	1.14325752	0.00484073	WD repeat-containing protein 31
B4GALNT1	1.07225171	0.00518059	Beta-1,4 N-acetylgalactosaminyltransferase 1

RIMKLA	1.29686172	0.00537983	N-acetylaspartylglutamate synthase A
SLC2A3	1.83936386	0.0054069	Solute carrier family 2, facilitated glucose transporter member 3
SLC25A35	1.31531332	0.0054104	Solute carrier family 25 member 35
HIVEP2	1.26764093	0.00542105	Transcription factor HIVEP2
PSRC1	1.19708096	0.00548999	Proline/serine-rich coiled-coil protein 1
AQP7	1.2967626	0.005509	Aquaporin-7
SLC4A8	1.14460551	0.00554203	Electroneutral sodium bicarbonate exchanger 1
MON1A	-1.163814	0.00556909	Vacuolar fusion protein MON1 homolog A
SLFN5	2.02083824	0.00556909	Schlafen family member 5
ERBB4	2.5098041	0.00571873	Receptor tyrosine-protein kinase erbB-4
ABCE1	-1.0770091	0.00576693	ATP-binding cassette sub-family E member 1
KCTD17	1.03689928	0.00577104	BTB/POZ domain-containing protein KCTD17
PGBD5	1.07676785	0.00581799	PiggyBac transposable element-derived protein 5
GCNT3	-1.0928594	0.00584454	Beta-1,3-galactosyl-O-glycosyl-glycoprotein beta-1,6-N-acetylglucosaminyltransferase 3
SPRED3	1.59798153	0.00584454	Sprouty-related, EVH1 domain-containing protein 3
KIAA0020	-1.0980486	0.005936	Pumilio homolog 3
NFIB	1.1651468	0.00602395	Nuclear factor 1 B-type
RABEPK	-1.1521131	0.00607468	Rab9 effector protein with kelch motifs
JAG2	1.0501942	0.0061129	Protein jagged-2
POU2F2	1.46756152	0.00614702	POU domain, class 2, transcription factor 2
RRAGD	1.56326471	0.00619715	Ras-related GTP-binding protein D
HABP2	-1.2391533	0.0062039	Hyaluronan-binding protein 2
FER1L6	-1.9261819	0.00650659	Fer-1-like protein 6
SP8	2.0178492	0.00654486	Transcription factor Sp8
SH2B3	-1.0388976	0.00668383	SH2B adapter protein 3
PLAG1	1.03199383	0.00685847	Zinc finger protein PLAG1
FAM214B	1.316953	0.00702221	Protein FAM214B
MAP3K12	1.81297928	0.00702221	Mitogen-activated protein kinase kinase kinase 12
LRFN1	1.1221081	0.00705507	Leucine-rich repeat and fibronectin type III domain-containing protein 1
TUBB6	-2.0821355	0.00706877	Tubulin beta-6 chain
ALKBH6	1.03016863	0.00712888	Alpha-ketoglutarate-dependent dioxygenase alkB homolog 6
CCDC113	1.45493923	0.00713249	Coiled-coil domain-containing protein 113
TRAF5	1.1948174	0.00713249	TNF receptor-associated factor 5
KIAA0922	1.40044443	0.00741968	Transmembrane protein 131-like
KIAA1467	1.02114543	0.0074516	Protein FAM234B
SAMD14	1.30993267	0.00763762	Sterile alpha motif domain containing 14
KANK3	2.2822107	0.00787359	KN motif and ankyrin repeat domain-containing protein 3

ST6GALNAC4	1.05634496	0.00795199	ST6 N-acetylgalactosaminide alpha-2,6-sialyltransferase 4
SLC30A4	1.13902997	0.00798766	Zinc transporter 4
MBL2	-1.3611093	0.00807486	Mannose-binding protein C
SKIDA1	1.67995545	0.0085215	SKI/DACH domain containing 1
SLC6A9	-1.5102304	0.00874928	Sodium- and chloride-dependent glycine transporter 1
PDSS1	-1.0883347	0.00885145	Decaprenyl-diphosphate synthase subunit 1
GPATCH4	-1.0099583	0.00898959	G-patch domain containing 4
EDN1	-1.116709	0.00905133	Endothelin-1
DUSP19	1.24525374	0.00921062	Dual specificity protein phosphatase 19
RAB6B	1.5588704	0.00929305	Ras-related protein Rab-6B
RBP5	1.61477457	0.00929327	Retinol-binding protein 5
CACNB3	1.1613482	0.009439	Voltage-dependent L-type calcium channel subunit beta-3
ARHGAP8	1.32345298	0.00979264	Rho GTPase activating protein 8
MYBPC2	2.57573459	0.00979264	Myosin-binding protein C, fast-type
HSPA8	-1.1029792	0.00999659	Heat shock cognate 71 kDa protein
KIRREL2	1.22590622	0.01004454	Kin of IRRE-like protein 2
CEBPD	-1.0912328	0.0100729	CCAAT/enhancer-binding protein delta
XPOT	-1.0523005	0.01042644	Exportin-T
ATG16L2	1.15160177	0.01057456	Autophagy-related protein 16-2
SLC10A2	-1.3765161	0.01059619	Ileal sodium/bile acid cotransporter
VAMP1	1.40887306	0.01084551	Vesicle-associated membrane protein 1
SEC31B	1.28930797	0.01084928	Protein transport protein Sec31B
RAG1	1.24071915	0.01089025	V(D)J recombination-activating protein 1
ZNF565	-1.0227002	0.01099831	Zinc finger protein 565
MLLT11	1.25507482	0.01103558	Protein AF1q
FAM171A2	1.41454356	0.01108935	Protein FAM171A2
ATF5	-1.0566786	0.01116928	Cyclic AMP-dependent transcription factor ATF-5
NKX3-1	1.34312115	0.01139224	Homeobox protein Nkx-3.1
WARS	-1.0267485	0.01152425	Tryptophan--tRNA ligase, cytoplasmic
MRC2	-1.3034635	0.01153293	C-type mannose receptor 2
CYB561D2	-1.0513542	0.01154556	Cytochrome b561 family member D2
ARHGAP44	1.3181863	0.01177605	Rho GTPase-activating protein 44
SWSAP1	-1.1168359	0.01188204	ATPase SWSAP1
FBXO16	1.4144818	0.01190929	F-box only protein 16
RNF186	2.45214477	0.01191721	Ring finger protein 186
FOSL1	-1.3185317	0.01219692	Fos-related antigen 1
GBP1	1.33446981	0.01226058	Guanylate-binding protein 1
DPF3	1.58281486	0.01226928	Zinc finger protein DPF3
OVOL2	1.70897981	0.0122845	Transcription factor Ovo-like 2
LOX	1.09862152	0.01238511	Protein-lysine 6-oxidase

PALLD	1.309389	0.01266727	Palladin
CXCL3	-2.5982603	0.01277126	C-X-C motif chemokine 3
NEURL1B	1.18262141	0.01291136	E3 ubiquitin-protein ligase NEURL1B
KLF15	1.61865053	0.01300856	Krueppel-like factor 15
CCDC88A	1.32075822	0.01309622	Girdin
PSAT1	-1.1458782	0.01335363	Phosphoserine aminotransferase
SEC14L2	1.20890919	0.01375525	SEC14-like protein 2
DKK1	-1.8537276	0.01395257	Dickkopf-related protein 1
AGFG2	1.10981513	0.01397812	Arf-GAP domain and FG repeat-containing protein 2
ICA1L	-1.6634048	0.01398435	Islet cell autoantigen 1-like protein
DLG2	1.57227036	0.014375	Disks large homolog 2
FAM134B	1.05053348	0.014406	Reticulophagy regulator 1
CCDC33	1.82443563	0.01449491	Coiled-coil domain containing 33
DLG4	1.31316801	0.01457782	Disks large homolog 4
FSTL3	1.16630523	0.01470085	Follistatin-related protein 3
MINA	-1.0057315	0.01496209	Ribosomal oxygenase 2
ASPHD1	1.06183307	0.01496209	Aspartate beta-hydroxylase domain containing 1
GPRC6A	-1.7059284	0.01556562	G-protein coupled receptor family C group 6 member A
LYPD6	1.05033796	0.01567821	Ly6/PLAUR domain-containing protein 6
ZDHHC11B	1.68609157	0.01575934	Probable palmitoyltransferase ZDHHC11B
WDR91	1.22079938	0.0158201	WD repeat-containing protein 91
ALOX15	1.4238128	0.01591201	Arachidonate 15-lipoxygenase
MYCL	1.54619542	0.01604585	Protein L-Myc
C5orf34	1.49514677	0.01660667	Uncharacterized protein C5orf34
NME1-NME2	-1.4029822	0.01734816	Nucleoside diphosphate kinase B
GUCY1A2	-1.7535576	0.01742171	Guanylate cyclase soluble subunit alpha-2
LRRC24	-1.3830699	0.01795066	Leucine-rich repeat-containing protein 24
LETM2	1.57336834	0.01797957	LETM1 domain-containing protein LETM2, mitochondrial
RBM24	2.47569181	0.01862077	RNA-binding protein 24
COL16A1	1.67213115	0.01886423	Collagen alpha-1(XVI) chain
EIF4EBP1	-1.0860945	0.01902282	Eukaryotic translation initiation factor 4E-binding protein 1
PBX1	1.00777965	0.01909915	Pre-B-cell leukemia transcription factor 1
DHRS9	-1.3423079	0.0192336	Dehydrogenase/reductase SDR family member 9
NR4A3	2.13735497	0.01938574	Nuclear receptor subfamily 4 group A member 3
BVES	1.21090206	0.01941979	Blood vessel epicardial substance
CNPY4	1.18880854	0.01982654	Protein canopy homolog 4
LIN7B	1.34081374	0.01996625	Protein lin-7 homolog B
TMEM159	1.16936087	0.02029197	Promethin

CDK5R1	-1.091212	0.02033327	Cyclin-dependent kinase 5 activator 1
C17orf67	1.05962633	0.02129654	Uncharacterized protein C17orf67
FZD4	1.04113908	0.02129654	Frizzled-4
NFE2	1.49058124	0.02139782	Transcription factor NF-E2 45 kDa subunit
TMTC2	1.54949485	0.02155464	Transmembrane and tetratricopeptide repeat containing 2
GNB3	1.40729881	0.02186567	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-3
PROCA1	1.45282258	0.02200671	Protein interacting with cyclin A1
SLC17A1	1.56865623	0.02207195	Sodium-dependent phosphate transport protein 1
SLC2A14	1.28615369	0.02207195	Solute carrier family 2, facilitated glucose transporter member 14
DSC3	-1.1926407	0.02216227	Desmocollin-3
MEF2C	1.21810051	0.02285308	Myocyte-specific enhancer factor 2C
DNAJC6	1.00149923	0.02286776	Putative tyrosine-protein phosphatase auxilin
BCAT1	-1.2969336	0.02335508	Branched-chain-amino-acid aminotransferase, cytosolic
DLX3	1.32221961	0.02349903	Homeobox protein DLX-3
PIK3CD	1.62556423	0.02350419	Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit delta isoform
SLC2A10	1.17507681	0.02450494	Solute carrier family 2, facilitated glucose transporter member 10
SLC19A3	1.03126922	0.02451608	Thiamine transporter 2
TMEM107	1.2587519	0.02507668	Transmembrane protein 107
WIPF3	1.30036966	0.02524257	WAS/WASL-interacting protein family member 3
HSPA6	-5.2785543	0.02524804	Heat shock 70 kDa protein 6
C14orf28	-1.1451589	0.0260525	Uncharacterized protein C14orf28
MAP1LC3B2	-1.0069231	0.02626231	Microtubule-associated proteins 1A/1B light chain 3 beta 2
NOC3L	-1.1586814	0.02658166	Nucleolar complex protein 3 homolog
SEPT3	1.42973652	0.02684664	Neuronal-specific septin-3
ZNF563	1.15100396	0.02708554	Zinc finger protein 563
GAL	-1.109946	0.02714956	Galanin peptides
FAM227A	1.37089167	0.02758441	Protein FAM227A
CHAC1	-1.5936899	0.02769408	Glutathione-specific gamma-glutamylcyclotransferase 1
SCLY	1.16026997	0.0278444	Selenocysteine lyase
AZIN2	1.30233887	0.02820345	Antizyme inhibitor 2
CCDC80	1.24438945	0.02909055	Coiled-coil domain-containing protein 80
G0S2	-1.1585522	0.0292801	G0/G1 switch protein 2
AGPAT4	1.79948502	0.02940676	1-acyl-sn-glycerol-3-phosphate acyltransferase delta
MYZAP	1.00381076	0.02990323	Myocardial zonula adherens protein
CCDC30	1.4666145	0.03035374	Coiled-coil domain containing 30
HIST1H4K	-1.5161211	0.0304457	Histone cluster 1 H4 family member k
AK7	1.09671648	0.03077237	Adenylate kinase 7
C1S	1.03175636	0.03082014	Complement C1s subcomponent

CCBL1	1.26883372	0.03092825	Kynurenine--oxoglutarate transaminase 1
HILPDA	1.47563375	0.03092825	Hypoxia-inducible lipid droplet-associated protein
DNAJB9	-1.021807	0.03098156	DnaJ homolog subfamily B member 9
CDA	1.03634684	0.03098156	Cytidine deaminase
BMF	1.01961729	0.03104973	Bcl-2-modifying factor
NR4A2	1.17925194	0.03104973	Nuclear receptor subfamily 4 group A member 2
WBP1	1.05104469	0.03128572	WW domain binding protein 1
NEDD9	1.54504033	0.03144178	Enhancer of filamentation 1
CAV1	1.35642949	0.03183785	Caveolin-1
MAB21L3	2.08433891	0.03183785	Protein mab-21-like 3
KCNJ8	1.11401426	0.03207864	ATP-sensitive inward rectifier potassium channel 8
PXK	1.12445442	0.03387836	PX domain-containing protein kinase-like protein
ASTN2	1.05720164	0.03409318	Astrotactin-2
DNAJB5	1.26619691	0.03472248	DnaJ heat shock protein family member B5
TTLL6	1.09393775	0.03489288	Tubulin polyglutamylase TTLL6
CEP19	1.45651711	0.03526527	Centrosomal protein of 19 kDa
TBX20	-1.2995672	0.03552491	T-box transcription factor TBX20
TMEM105	1.36551573	0.03582737	Transmembrane protein 105
DFNB31	1.7834282	0.03604839	Whirlin
RHPN1	1.02561896	0.03617919	Rhophilin-1
ALPPL2	1.24328694	0.03637221	Alkaline phosphatase, placental like 2
PGM2L1	1.11501646	0.03653536	Glucose 1,6-bisphosphate synthase
RELB	-1.1016776	0.03670212	Transcription factor RelB
FAM46C	1.40208158	0.03670212	Putative nucleotidyltransferase FAM46C
ZNF430	-1.0216191	0.03683773	Zinc finger protein 430
ITGAM	1.17312559	0.03724024	Integrin alpha-M
ZNF575	-1.1476698	0.03742155	Zinc finger protein 575
MPP2	1.67729709	0.03806887	MAGUK p55 subfamily member 2
INSIG2	1.083259	0.03807226	Insulin-induced gene 2 protein
BTN1A1	1.14516174	0.03813746	Butyrophilin subfamily 1 member A1
NOVA1	1.14223153	0.03836965	RNA-binding protein Nova-1
CYP19A1	1.55956781	0.03935527	Aromatase
HOXC10	1.07487959	0.03936795	Homeobox protein Hox-C10
SENP8	-1.3631695	0.03945151	Sentrin-specific protease 8
CDKN2D	1.37382653	0.04019941	Cyclin-dependent kinase 4 inhibitor D
ESRRG	1.02725911	0.04136386	Estrogen-related receptor gamma
ABR	-1.5786221	0.04293049	Active breakpoint cluster region-related protein
TMCC3	1.3712361	0.04482452	Transmembrane and coiled-coil domain containing
RAD54B	1.11463809	0.04508742	DNA repair and recombination protein RAD54B

VSTM5	-1.2334114	0.0451735	V-set and transmembrane domain-containing protein 5
SH3BP5	1.27782324	0.04551108	SH3 domain-binding protein 5
SLC39A2	1.11909237	0.04697375	Zinc transporter ZIP2
CASZ1	1.02232296	0.04914708	Zinc finger protein castor homolog 1
AK1	1.08014711	0.04966817	Adenylate kinase isoenzyme 1

Appendix table 12. Human differentially expressed genes between uninfected controls and infected samples at 6 h post infection.

Gene name	log_2(fold change)	padj	annotation
FGG	-1.7353744	3.39E-10	Fibrinogen gamma chain
TNIK	1.04609136	2.10E-06	TRAF2 and NCK-interacting protein kinase
UGT2B17	-2.4121841	2.10E-06	UDP-glucuronosyltransferase 2B17
LRRC20	1.76291526	2.45E-06	Leucine-rich repeat-containing protein 20
EME1	1.3014727	3.14E-05	Crossover junction endonuclease EME1
HIP1	2.0012788	3.35E-05	Huntingtin-interacting protein 1
NIPAL4	-1.4967801	3.35E-05	Magnesium transporter NIPA4
EPB41L2	1.04930436	3.42E-05	Band 4.1-like protein 2
FGB	-1.6820799	5.79E-05	Fibrinogen beta chain
FAM184A	1.16561658	7.43E-05	Protein FAM184A
HK1	1.81936296	0.00010127	Hexokinase-1
GATA6	-1.1979182	0.00026919	Transcription factor GATA-6
BTN2A2	2.07840185	0.00037501	Butyrophilin subfamily 2 member A2
HSPB1	1.54592834	0.00041931	Heat shock protein beta-1
USP43	-1.3063781	0.00041931	Ubiquitin carboxyl-terminal hydrolase 43
ADAM19	1.278649	0.00045163	Disintegrin and metalloproteinase domain-containing protein 19
MRPL54	-1.0361828	0.00049981	Mitochondrial ribosomal protein L54
NDRG4	1.62582655	0.00060809	Protein NDRG4
AAED1	-1.0102022	0.00060809	Thioredoxin-like protein AAED1
TNRC18	-1.1123319	0.00060809	Trinucleotide repeat-containing gene 18 protein
L3MBTL3	1.02329345	0.00060809	Lethal(3)malignant brain tumor-like protein 3
ADAMTS6	2.23953443	0.00087962	ADAM metallopeptidase with thrombospondin type 1 motif 6
STEAP2	1.01872578	0.0009391	Metalloreductase STEAP2
NIPAL3	1.02387495	0.00096273	NIPA-like protein 3
AADAC	-1.3213942	0.00100452	Arylacetamide deacetylase
WIPF1	1.08746069	0.00100452	WAS/WASL-interacting protein family member 1
MBOAT2	1.39940654	0.00121934	Lysophospholipid acyltransferase 2

NR3C1	-1.1512995	0.0014564	Glucocorticoid receptor
TMEM63C	1.27598004	0.0014564	Calcium permeable stress-gated cation channel 1
MB21D2	-1.0698681	0.0014564	Protein MB21D2
CECR2	1.42010103	0.00152162	Cat eye syndrome critical region protein 2
SCAMP5	1.25449602	0.00157863	Secretory carrier-associated membrane protein 5
TMEM169	1.5994724	0.00172965	Transmembrane protein 169
WDR31	1.49329149	0.00175227	WD repeat-containing protein 31
LYRM9	1.28067422	0.00175227	LYR motif containing 9
C1orf122	-1.2334011	0.00175986	Uncharacterized protein C1orf122
MEF2C	1.78976076	0.00195451	Myocyte-specific enhancer factor 2C
MED12L	1.65737251	0.00195451	Mediator of RNA polymerase II transcription subunit 12-like protein
FGA	-1.4390903	0.00195451	Fibrinogen alpha chain
MAML2	1.11387766	0.00195451	Mastermind-like protein 2
ATP2B4	1.11260951	0.00216229	Plasma membrane calcium-transporting ATPase 4
B4GALNT1	1.25953401	0.00216229	Beta-1,4 N-acetylgalactosaminyltransferase 1
DMXL2	1.00246154	0.00223621	DmX-like protein 2
NFIB	1.40480959	0.00223621	Nuclear factor 1 B-type
ARHGEF4	1.2323472	0.00233093	Rho guanine nucleotide exchange factor 4
PPARG	1.35442	0.00252496	Peroxisome proliferator-activated receptor gamma
MSX2	-1.1358802	0.00270515	Homeobox protein MSX-2
IRX5	-1.34171	0.00274523	Iroquois-class homeodomain protein IRX-5
FOXQ1	-1.2626591	0.00287356	Forkhead box protein Q1
GIPR	2.07042761	0.00313138	Gastric inhibitory polypeptide receptor
AKAP7	1.5411756	0.00332089	A-kinase anchor protein 7 isoform gamma
PAIP2B	1.43219315	0.00341054	Polyadenylate-binding protein-interacting protein 2B
KIAA0319	1.18901501	0.00387246	Dyslexia-associated protein KIAA0319
ESAM	1.32474797	0.00406201	Endothelial cell-selective adhesion molecule
DNAH3	1.13116143	0.00406201	Dynein heavy chain 3, axonemal
TM4SF20	1.17734434	0.00406201	Transmembrane 4 L6 family member 20
C20orf96	1.07567836	0.00406201	Uncharacterized protein C20orf96
GVQW3	1.29513847	0.00449755	annotation not available
CREB3L3	1.33773707	0.00535211	Cyclic AMP-responsive element-binding protein 3-like protein 3
ARHGEF25	1.26828202	0.00535211	Rho guanine nucleotide exchange factor 25
SPTBN5	-1.371332	0.00535266	Spectrin beta, non-erythrocytic 5
PLCH1	1.13839431	0.00558806	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase eta-1
PALLD	1.59250463	0.00702097	Palladin
TNNC1	1.28728924	0.0073335	Troponin C, slow skeletal and cardiac muscles
HABP2	-1.3723385	0.00748143	Hyaluronan-binding protein 2
DIRAS1	2.07969089	0.00748143	GTP-binding protein Di-Ras1

KIAA1467	1.19332152	0.00757384	Protein FAM234B
TRAF5	1.37183334	0.00779296	TNF receptor-associated factor 5
TMCC3	2.00920414	0.00783612	Transmembrane and coiled-coil domain containing
CYP19A1	2.376754	0.00783612	Aromatase
ALOX15	1.70002489	0.00814348	Arachidonate 15-lipoxygenase
RAB31	1.51355504	0.00814348	Ras-related protein Rab-31
HPDL	-1.0120652	0.00821355	4-hydroxyphenylpyruvate dioxygenase-like protein
MCAM	1.32837978	0.00839019	Cell surface glycoprotein MUC18
AMN	-1.0401404	0.0084306	Protein amnionless
PPP1R3G	1.84749895	0.00863076	Protein phosphatase 1 regulatory subunit 3G
CATSPER2	1.28331928	0.00895031	Cation channel sperm-associated protein 2
ATP1B2	1.37204323	0.00934315	Sodium/potassium-transporting ATPase subunit beta-2
FAM101A	-1.0940138	0.00934315	Refilin-A
ALPK1	1.11323406	0.00959738	Alpha-protein kinase 1
LPHN1	1.11245116	0.0096271	Adhesion G protein-coupled receptor L1
LGALS1	1.47629601	0.01014901	Galectin-1
RNF144A	1.46599263	0.01118568	E3 ubiquitin-protein ligase RNF144A
NR2F6	-1.0640524	0.01118568	Nuclear receptor subfamily 2 group F member 6
AZGP1	1.47848312	0.0117249	Zinc-alpha-2-glycoprotein
FAM122C	1.50921306	0.01195586	Protein FAM122C
ANKZF1	1.08267244	0.01342582	Ankyrin repeat and zinc finger domain-containing protein 1
BLVRA	1.40013704	0.01348939	Biliverdin reductase A
APOL2	1.1201107	0.01348939	Apolipoprotein L2
PORCN	1.86125926	0.01424079	Protein-serine O-palmitoleoyltransferase porcupine
ITIH2	-1.2867769	0.01436061	Inter-alpha-trypsin inhibitor heavy chain H2
TDP2	1.05310326	0.01520368	Tyrosyl-DNA phosphodiesterase 2
CCDC88A	1.49266422	0.01525878	Girdin
NFIA	1.03764461	0.01525878	Nuclear factor 1 A-type
FABP2	-1.3008386	0.01638041	Fatty acid-binding protein, intestinal
SEPT6	1.21212144	0.01645844	Septin-6
FYN	1.30085948	0.0166411	Tyrosine-protein kinase Fyn
PIPOX	1.04710584	0.01763303	Peroxisomal sarcosine oxidase
FOXA1	-1.0931889	0.01788469	Hepatocyte nuclear factor 3-alpha
ZBED8	1.42630243	0.01788469	Protein ZBED8
HOXB13	1.20353209	0.01800522	Homeobox protein Hox-B13
CDA	1.28345673	0.01841286	Cytidine deaminase
MAP1A	1.33240021	0.01876919	Microtubule-associated protein 1A
SEPT3	1.79231268	0.01943844	Neuronal-specific septin-3
CNPY4	1.4110512	0.01952969	Protein canopy homolog 4
ZFHX4	1.7166174	0.02063604	Zinc finger homeobox protein 4

EML6	1.11326907	0.02088053	Echinoderm microtubule-associated protein-like 6
NOXO1	-1.4357165	0.02144577	NADPH oxidase organizer 1
ANKRD37	1.87367081	0.02152015	Ankyrin repeat domain-containing protein 37
MFAP3L	1.08003312	0.02152015	Microfibrillar-associated protein 3-like
ENTPD1	1.53518771	0.0223412	Ectonucleoside triphosphate diphosphohydrolase 1
HILPDA	1.75856506	0.02250276	Hypoxia-inducible lipid droplet-associated protein
CCDC80	1.48714543	0.02252718	Coiled-coil domain-containing protein 80
FBXO16	1.63568285	0.02273088	F-box only protein 16
RSPH1	1.4082552	0.02293413	Radial spoke head 1 homolog
SRGAP2	1.22223229	0.0237914	SLIT-ROBO Rho GTPase-activating protein 2
FAM13A	1.91126029	0.02420738	Protein FAM13A
SLC25A35	1.24051437	0.02423816	Solute carrier family 25 member 35
CXCL8	-1.902278	0.0250175	Interleukin-8
ST3GAL5	1.91875001	0.02532524	Lactosylceramide alpha-2,3-sialyltransferase
FAM171A2	1.48889291	0.02568808	Protein FAM171A2
TBC1D32	1.04189196	0.02720141	Protein broad-minded
DOK3	-1.0877915	0.02747035	Docking protein 3
SOHLH1	-1.0759591	0.02850043	Spermatogenesis- and oogenesis-specific basic helix-loop-helix-containing protein 1
NCMAP	1.68056488	0.02970443	Noncompact myelin-associated protein
S100A1	1.46228715	0.03005323	Protein S100-A1
ZNF577	1.15632271	0.03016807	Zinc finger protein 577
RNF186	2.46857565	0.03045733	Ring finger protein 186
MNS1	1.18558545	0.03284187	Meiosis-specific nuclear structural protein 1
BVES	1.22099892	0.03302579	Blood vessel epicardial substance
ZNF385A	1.18623151	0.03348419	Zinc finger protein 385A
CSRNP3	1.68589358	0.03392758	Cysteine/serine-rich nuclear protein 3
RHOBTB1	1.17044137	0.03401286	Rho-related BTB domain-containing protein 1
ZFP14	1.01353224	0.03411053	Zinc finger protein 14 homolog
SH3PXD2B	1.07418033	0.03411053	SH3 and PX domain-containing protein 2B
CRYBG3	1.04553466	0.03538463	Very large A-kinase anchor protein
COL4A5	-1.0575882	0.03578634	Collagen alpha-5(IV) chain
SLC2A3	1.67817279	0.03603782	Solute carrier family 2, facilitated glucose transporter member 3
MPP2	1.95140294	0.03603782	MAGUK p55 subfamily member 2
TNFAIP2	1.68675596	0.03603782	Tumor necrosis factor alpha-induced protein 2
PLCG2	1.4060057	0.03735669	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase gamma-2
FOSL1	-1.1565385	0.0374129	Fos-related antigen 1
MLLT11	1.21606538	0.0374129	Protein AF1q
SLC25A27	2.14249977	0.03777013	Mitochondrial uncoupling protein 4
TMCC1	1.01585868	0.03830274	Transmembrane and coiled-coil domain containing
HMOX1	-1.6934867	0.04136916	Heme oxygenase 1

RAB6B	1.50022089	0.04196671	Ras-related protein Rab-6B
CCL20	-2.8186678	0.0427418	C-C motif chemokine 20
ROBO1	1.26171061	0.04279273	Roundabout homolog 1
CCDC136	1.27689184	0.04288982	Coiled-coil domain-containing protein 136
CRABP2	1.45330357	0.04332087	Cellular retinoic acid-binding protein 2
SLC6A12	1.33519645	0.04494911	Sodium- and chloride-dependent betaine transporter
ZNF287	1.02582511	0.04504058	Zinc finger protein 287
RASSF4	1.05154395	0.04699085	Ras association domain-containing protein 4
FBXL16	1.19696515	0.04733043	F-box/LRR-repeat protein 16
GNB3	1.28545363	0.047752	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-3
ARHGAP33	1.02092682	0.04792836	Rho GTPase-activating protein 33
NRGN	1.05998641	0.04824144	Neurogranin
ITGB6	1.36592678	0.04891502	Integrin beta-6
MTURN	1.41849159	0.04907918	Maturin
HSH2D	1.53985567	0.0491621	Hematopoietic SH2 domain-containing protein
KIAA1407	1.07650843	0.04940807	Coiled-coil domain-containing protein 191

Appendix table 13. Human differentially expressed genes between uninfected controls and infected samples at 12 h post infection.

Gene name	log_2(fold change)	padj	annotation
HSPB1	2.36056938	2.10E-11	Heat shock protein beta-1
GPT2	1.0020111	5.96E-10	Alanine aminotransferase 2
ANXA6	1.66217484	8.45E-09	Annexin A6
TM4SF20	1.96434946	2.19E-08	Transmembrane 4 L6 family member 20
FGG	-1.5705264	2.41E-08	Fibrinogen gamma chain
TNFRSF12A	-1.5584119	3.18E-07	Tumor necrosis factor receptor superfamily member 12A
CREB3L3	2.09264695	3.85E-07	Cyclic AMP-responsive element-binding protein 3-like protein 3
EPB41L2	1.16657895	5.60E-07	Band 4.1-like protein 2
LGALS1	2.34945594	5.60E-07	Galectin-1
AKAP7	2.11921019	7.28E-06	A-kinase anchor protein 7 isoform gamma
FAM13A	3.18035253	9.48E-06	Protein FAM13A
SLC6A12	2.45346743	1.29E-05	Sodium- and chloride-dependent betaine transporter
EGR1	-2.004161	1.83E-05	Early growth response protein 1
STC2	1.53717258	2.01E-05	Stanniocalcin-2
SULF2	1.43339237	2.74E-05	Extracellular sulfatase Sulf-2
RPS6KA4	-1.1390668	2.76E-05	Ribosomal protein S6 kinase alpha-4
HIP1	1.96964661	2.93E-05	Huntingtin-interacting protein 1

KIFC3	-1.1158051	2.98E-05	Kinesin-like protein KIFC3
EME1	1.15365158	3.38E-05	Crossover junction endonuclease EME1
CLU	1.0537914	3.40E-05	Clusterin
BTN3A2	1.33400706	3.42E-05	Butyrophilin subfamily 3 member A2
PIPOX	1.69265976	3.44E-05	Peroxisomal sarcosine oxidase
ANKZF1	1.51605054	5.40E-05	Ankyrin repeat and zinc finger domain-containing protein 1
CDA	1.92873796	6.42E-05	Cytidine deaminase
PTPRS	1.81673326	7.25E-05	Receptor-type tyrosine-protein phosphatase S
PDK1	1.6718722	8.10E-05	[Pyruvate dehydrogenase (acetyl-transferring)] kinase isozyme 1, mitochondrial
CACNA1H	1.52620775	9.31E-05	Voltage-dependent T-type calcium channel subunit alpha-1H
FAM184A	1.05496575	9.31E-05	Protein FAM184A
PLOD2	1.86565213	9.99E-05	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 2
HILPDA	2.50496753	0.00011285	Hypoxia-inducible lipid droplet-associated protein
NDRG1	3.9141664	0.00013009	Protein NDRG1
FOSL1	-1.7975244	0.00015883	Fos-related antigen 1
KIF20A	1.22834724	0.00018064	Kinesin-like protein KIF20A
ATP1B2	1.82923948	0.00020362	Sodium/potassium-transporting ATPase subunit beta-2
TDP2	1.43907558	0.00020362	Tyrosyl-DNA phosphodiesterase 2
BNIP3L	1.87693369	0.00020978	BCL2/adenovirus E1B 19 kDa protein-interacting protein 3-like
LRP4	1.83306083	0.00020978	Low-density lipoprotein receptor-related protein 4
SPTBN5	-1.6245759	0.00020978	Spectrin beta, non-erythrocytic 5
EPHA2	-1.2189044	0.00030822	Ephrin type-A receptor 2
ARHGEF25	1.54784009	0.00031615	Rho guanine nucleotide exchange factor 25
NRGN	1.7271506	0.00032987	Neurogranin
WNT11	2.15223138	0.00032987	Protein Wnt-11
NDRG4	1.67255565	0.00033028	Protein NDRG4
LRP2	1.48986887	0.000337	Low-density lipoprotein receptor-related protein 2
P4HA1	1.53447907	0.00037677	Prolyl 4-hydroxylase subunit alpha-1
APOA4	2.10151207	0.0003793	Apolipoprotein A-IV
PPFIA4	3.8382608	0.0003793	Liprin-alpha-4
ANKLE1	1.38760404	0.00038104	Ankyrin repeat and LEM domain-containing protein 1
HAVCR1	1.52487923	0.00038104	Hepatitis A virus cellular receptor 1
CCDC183	2.41325885	0.00038363	Coiled-coil domain containing 183
SH2D3A	-1.1893468	0.00040421	SH2 domain-containing protein 3A
LCAT	-1.7534069	0.0004678	Phosphatidylcholine-sterol acyltransferase
CECR2	1.4873269	0.00048644	Cat eye syndrome critical region protein 2
MOGAT2	4.39667495	0.00048644	2-acylglycerol O-acyltransferase 2
ZHX2	1.18591682	0.00048644	Zinc fingers and homeoboxes protein 2

PPARG	1.45676387	0.00050822	Peroxisome proliferator-activated receptor gamma
APOL6	1.01747027	0.00051363	Apolipoprotein L6
LDHA	1.1756039	0.00051363	Lactate dehydrogenase A
CXCL8	-2.6487032	0.00051827	Interleukin-8
ITGA5	3.21956136	0.00051928	Integrin alpha-5
HPDL	-1.1868906	0.00055379	4-hydroxyphenylpyruvate dioxygenase-like protein
UGT2B17	-1.7518983	0.00055379	UDP-glucuronosyltransferase 2B17
DNAH3	1.25429259	0.00056884	Dynein heavy chain 3, axonemal
SLC36A2	3.2743743	0.00057247	Proton-coupled amino acid transporter 2
REEP2	1.78322038	0.00061877	Receptor expression-enhancing protein 2
GATM	1.50190144	0.00066663	Glycine amidinotransferase, mitochondrial
CTSV	1.10740786	0.00073892	Cathepsin L2
PHEX	1.28674893	0.00073892	Phosphate-regulating neutral endopeptidase
ANXA4	1.0875356	0.00075546	Annexin A4
CCDC136	2.07701644	0.00084075	Coiled-coil domain-containing protein 136
CLCF1	-1.8461846	0.0008496	Cardiotrophin-like cytokine factor 1
PREX1	2.64091886	0.00098889	Phosphatidylinositol 3,4,5-trisphosphate-dependent Rac exchanger 1 protein
SLC2A3	2.31729977	0.00102059	Solute carrier family 2, facilitated glucose transporter member 3
SYNM	1.60643986	0.00102059	Synemin
SERTAD1	-1.2975896	0.00104772	SERTA domain-containing protein 1
ADA	2.13798568	0.00112236	Adenosine deaminase
DEPTOR	1.92909287	0.00115753	DEP domain-containing mTOR-interacting protein
HSDL2	1.09342951	0.00123603	Hydroxysteroid dehydrogenase-like protein 2
ADAM19	1.11045332	0.00143714	Disintegrin and metalloproteinase domain-containing protein 19
VSNL1	2.42201441	0.00145998	Visinin-like protein 1
CBFA2T3	1.97785252	0.00154383	Protein CBFA2T3
NFE2	1.99910859	0.00159547	Transcription factor NF-E2 45 kDa subunit
DIMT1	-1.1488683	0.0017238	Probable dimethyladenosine transferase
SEPT6	1.49169362	0.00196063	Septin-6
HK1	1.45230957	0.00201177	Hexokinase-1
PRR15L	1.27953555	0.00201177	Proline-rich protein 15-like protein
SERPINA5	1.00610283	0.00201177	Plasma serine protease inhibitor
PGBD5	1.30161825	0.0022253	PiggyBac transposable element-derived protein 5
AMICA1	1.34573242	0.00234218	Junctional adhesion molecule-like
CRABP2	1.96244019	0.00234218	Cellular retinoic acid-binding protein 2
FAM114A1	1.21184519	0.00234218	Protein NOXP20
LGI4	3.18731507	0.00234218	Leucine rich repeat LGI family member 4
OAS1	1.10563178	0.00235344	2'-5'-oligoadenylate synthase 1
HMGCS2	2.22618103	0.00238368	Hydroxymethylglutaryl-CoA synthase, mitochondrial
ANKRD37	2.18865482	0.00246924	Ankyrin repeat domain-containing protein 37

PLCG2	1.89462065	0.00252224	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase gamma-2
BTN3A1	1.22565728	0.0027635	Butyrophilin subfamily 3 member A1
RAB6B	2.03986145	0.00286433	Ras-related protein Rab-6B
KDM3A	1.32434826	0.00287991	Lysine-specific demethylase 3A
CD3G	1.88477489	0.00326712	T-cell surface glycoprotein CD3 gamma chain
TTLL6	1.70990556	0.00329853	Tubulin polyglutamylase TTLL6
RNF183	2.00179765	0.00344935	Probable E3 ubiquitin-protein ligase RNF183
CACNA1G	2.02976471	0.00347683	Voltage-dependent T-type calcium channel subunit alpha-1G
TRABD2B	2.12579063	0.00367406	Metalloprotease TIKI2
LPPR3	2.12301898	0.00385928	Lipid phosphate phosphatase-related protein type 3
RGCC	-1.6280554	0.00392117	Regulator of cell cycle RGCC
ASPA	1.59533425	0.00415629	Aspartoacylase
SLC25A6	-2.2475594	0.00460836	ADP/ATP translocase 3
CSF1R	2.02225755	0.00460847	Macrophage colony-stimulating factor 1 receptor
EIF5A2	-1.3071344	0.00487376	Eukaryotic translation initiation factor 5A-2
ALOX15	1.75463357	0.00496758	Arachidonate 15-lipoxygenase
CDCA2	1.16429275	0.00496758	Cell division cycle-associated protein 2
SLC28A1	2.53600671	0.00496758	Sodium/nucleoside cotransporter 1
DIRAS1	2.20815473	0.0050859	GTP-binding protein Di-Ras1
MCAM	1.37101899	0.0050859	Cell surface glycoprotein MUC18
PTPRR	1.49936418	0.0050859	Receptor-type tyrosine-protein phosphatase R
SORBS1	1.14579645	0.00509288	Sorbin and SH3 domain-containing protein 1
ZNF395	1.29057953	0.00528378	Zinc finger protein 395
CDHR5	-3.7939526	0.00541832	Cadherin-related family member 5
SULT1C2	1.03015613	0.00541832	Sulfotransferase 1C2
FOXQ1	-1.153391	0.00559658	Forkhead box protein Q1
TSHZ1	1.09221675	0.0059337	Teashirt homolog 1
CACNB3	1.38068105	0.00616492	Voltage-dependent L-type calcium channel subunit beta-3
FAM173A	-1.1159487	0.00616492	Protein FAM173A
SCAMP5	1.11009955	0.00619139	Secretory carrier-associated membrane protein 5
SLC6A8	2.43648525	0.00619139	Sodium- and chloride-dependent creatine transporter 1
RNF186	2.95296432	0.00632857	Ring finger protein 186
MEF2C	1.57068562	0.00678021	Myocyte-specific enhancer factor 2C
PDZK1	1.76992259	0.00678021	Na(+)/H(+) exchange regulatory cofactor NHE-RF3
IRX5	-1.2259042	0.00709329	Iroquois-class homeodomain protein IRX-5
CADM1	1.04687584	0.00718249	Cell adhesion molecule 1
ACOX2	1.50049635	0.00730765	Peroxisomal acyl-coenzyme A oxidase 2
BNIP3	1.36122308	0.00730765	BCL2/adenovirus E1B 19 kDa protein-interacting protein 3
RGS11	2.4342335	0.00732039	Regulator of G-protein signaling 11
FGGY	1.04934679	0.00757819	FGGY carbohydrate kinase domain containing

SALL2	1.74000261	0.0078182	Sal-like protein 2
TNFRSF6B	-1.2581879	0.00782074	Tumor necrosis factor receptor superfamily member 6B
GALNT16	2.20118344	0.00791821	Polypeptide N-acetylgalactosaminyltransferase 16
BUB1B	1.00502676	0.00805142	Mitotic checkpoint serine/threonine-protein kinase BUB1 beta
HOXB13	1.26964243	0.00829909	Homeobox protein Hox-B13
SLC6A4	1.01917259	0.00829909	Sodium-dependent serotonin transporter
TTC7B	1.21635693	0.00836413	Tetratricopeptide repeat protein 7B
ACOT13	1.00214207	0.00840684	Acyl-coenzyme A thioesterase 13
PLEKHB1	1.20446094	0.00840684	Pleckstrin homology domain-containing family B member 1
C1QL1	2.03723302	0.00841067	C1q-related factor
UGT1A6	1.06680146	0.00847373	UDP-glucuronosyltransferase 1-6
TNS1	2.61327958	0.00855923	Tensin-1
ANKRD1	-1.590978	0.00877901	Ankyrin repeat domain-containing protein 1
PRR22	-1.2218001	0.00904457	Proline-rich protein 22
TRPV2	1.46936308	0.00934944	Transient receptor potential cation channel subfamily V member 2
FBXL16	1.48957527	0.0093888	F-box/LRR-repeat protein 16
SLC51A	1.96679222	0.0093888	Organic solute transporter subunit alpha
ANPEP	1.09191477	0.00972265	Aminopeptidase N
TMEM45A	1.89117609	0.01009851	Transmembrane protein 45A
FGB	-1.1323342	0.01018856	Fibrinogen beta chain
ST8SIA6	2.33920084	0.01019454	Alpha-2,8-sialyltransferase 8F
ALPK2	1.37475224	0.01020774	Alpha-protein kinase 2
ENTPD2	1.54956677	0.01041604	Ectonucleoside triphosphate diphosphohydrolase 2
ONECUT1	2.24766676	0.01091694	Hepatocyte nuclear factor 6
SLC4A3	2.06312385	0.01105326	Anion exchange protein 3
JAG2	1.0612505	0.01110664	Protein jagged-2
TMEM169	1.30247309	0.01110664	Transmembrane protein 169
PLIN2	1.09389531	0.01132026	Perilipin-2
SORL1	1.21412973	0.01132026	Sortilin-related receptor
CERS1	1.4274336	0.01145573	Embryonic growth/differentiation factor 1
PITX2	1.03052429	0.01198189	Pituitary homeobox 2
DOC2A	1.24868322	0.01203409	Double C2-like domain-containing protein alpha
CA12	1.12902113	0.01209236	Carbonic anhydrase 12
FAM101A	-1.0324736	0.01216958	Refilin-A
SEPT3	1.93223889	0.01216958	Neuronal-specific septin-3
GPR111	1.93262236	0.0125269	Adhesion G-protein coupled receptor F2
TMCC1	1.12233512	0.0128906	Transmembrane and coiled-coil domain containing
TRABD2A	1.10244469	0.01293382	Metalloprotease TIKI1
TCP11L2	1.44011335	0.01297262	T-complex protein 11-like protein 2
LOX	1.22600199	0.01393991	Protein-lysine 6-oxidase

TMCC3	1.99530365	0.01430422	Transmembrane and coiled-coil domain containing
PFKFB3	1.71290541	0.0145675	6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 3
TIAM1	1.19731744	0.01539962	T-lymphoma invasion and metastasis-inducing protein 1
CSRP2	1.40529778	0.01550555	Cysteine and glycine-rich protein 2
ALDH1L2	1.59993092	0.0155302	Mitochondrial 10-formyltetrahydrofolate dehydrogenase
ENO2	2.33651241	0.0155302	Gamma-enolase
HEATR4	1.8674093	0.01571736	HEAT repeat-containing protein 4
LAPTM5	1.09665206	0.01571736	Lysosomal-associated transmembrane protein 5
AKR1D1	1.2558455	0.01580099	3-oxo-5-beta-steroid 4-dehydrogenase
PPP1R3G	1.57339829	0.0159387	Protein phosphatase 1 regulatory subunit 3G
OLFM2	1.69984111	0.01627021	Noelin-2
CHRNA4	1.96700347	0.01634443	Neuronal acetylcholine receptor subunit alpha-4
ACSM3	1.07389533	0.01658908	Acyl-coenzyme A synthetase ACSM3, mitochondrial
PLIN4	2.34881315	0.01658908	Perilipin-4
TLE2	1.00328014	0.01658908	Transducin-like enhancer protein 2
LEPREL2	1.05739176	0.01663069	Prolyl 3-hydroxylase 3
PLCH1	1.01154697	0.01663933	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase eta-1
EMP1	-1.2915656	0.01674989	Epithelial membrane protein 1
SLC2A14	1.50167682	0.01755301	Solute carrier family 2, facilitated glucose transporter member 14
BVES	1.35412285	0.01769272	Blood vessel epicardial substance
TNFAIP2	1.86304175	0.01783225	Tumor necrosis factor alpha-induced protein 2
PADI2	1.17242083	0.01820779	Protein-arginine deiminase type-2
UBE2L6	1.8610356	0.01822439	Ubiquitin/ISG15-conjugating enzyme E2 L6
SULT1A1	1.10965418	0.01823496	Sulfotransferase 1A1
SLC38A4	2.55373817	0.01832772	Sodium-coupled neutral amino acid transporter 4
ERBB4	3.06449783	0.01861157	Receptor tyrosine-protein kinase erbB-4
MXI1	1.14332443	0.01892356	Max-interacting protein 1
NREP	1.11933066	0.01892356	Neuronal regeneration-related protein
ST6GALNAC4	1.05668506	0.01904396	ST6 N-acetylgalactosaminide alpha-2,6-sialyltransferase 4
FAM131B	2.20263281	0.01930448	Protein FAM131B
ADAMTSL4	1.79826292	0.01966118	ADAMTS-like protein 4
KIAA0319	1.01882011	0.01970825	Dyslexia-associated protein KIAA0319
AGBL5	1.01944419	0.02002488	Cytosolic carboxypeptidase-like protein 5
CYP4F12	1.10218878	0.02055802	Cytochrome P450 4F12
RPS9	2.00320425	0.02061493	Ribosomal protein S9
CYP19A1	1.87160542	0.02067178	Aromatase
NFIL3	1.04693872	0.0211791	Nuclear factor interleukin-3-regulated protein
TNNT1	1.36754654	0.02129057	Troponin T, slow skeletal muscle
IKBKE	-1.5411338	0.02143469	Inhibitor of nuclear factor kappa-B kinase subunit epsilon

MYB	1.61373602	0.02143469	Transcriptional activator Myb
CP	1.16450685	0.02152289	Ceruloplasmin
NR1I3	1.99208699	0.02157679	Nuclear receptor subfamily 1 group I member 3
INSIG2	1.32263034	0.02203336	Insulin-induced gene 2 protein
SLC2A13	1.06842893	0.02242503	Proton myo-inositol cotransporter
BTN2A2	1.45421292	0.02267956	Butyrophilin subfamily 2 member A2
GADD45G	-1.4219499	0.02267956	Growth arrest and DNA damage-inducible protein GADD45 gamma
NRN1	2.62382145	0.02287528	Neuritin
ALS2CL	-1.2002829	0.02308855	ALS2 C-terminal-like protein
BMF	1.20669183	0.02326468	Bcl-2-modifying factor
MFAP3L	1.07918626	0.02339751	Microfibrillar-associated protein 3-like
HHEX	1.0566407	0.02402677	Hematopoietically-expressed homeobox protein HHEX
PDE6A	1.73436571	0.02402677	Rod cGMP-specific 3',5'-cyclic phosphodiesterase subunit alpha
MYLK	1.02409655	0.02425422	Myosin light chain kinase, smooth muscle
CASP16	1.99430671	0.02477642	Caspase 16, apoptosis-related cysteine peptidase (putative)
DPYSL5	2.10817871	0.02477642	Dihydropyrimidinase-related protein 5
CAPG	1.1293492	0.02495308	Macrophage-capping protein
DKK1	-1.9674225	0.02495769	Dickkopf-related protein 1
RIMKLA	1.20019703	0.02510854	N-acetylaspartylglutamate synthase A
NDNF	1.19883759	0.0258011	Protein NDNF
ZNF467	2.19611686	0.02586665	Zinc finger protein 467
UNC93A	1.19056348	0.02591705	Unc-93 homolog A
SYT17	1.163685	0.02651109	Synaptotagmin-17
TTYH2	1.1990418	0.02657061	Protein tweety homolog 2
KIRREL2	1.1916437	0.02722404	Kin of IRRE-like protein 2
ZNF618	1.30112318	0.02724738	Zinc finger protein 618
NOVA1	1.35233753	0.02741987	RNA-binding protein Nova-1
TNNC1	1.02982313	0.0282114	Troponin C, slow skeletal and cardiac muscles
SLC35E4	-1.0592955	0.02821703	Solute carrier family 35 member E4
DACH1	1.12184185	0.02830995	Dachshund homolog 1
GATA2	1.33072156	0.02863951	Endothelial transcription factor GATA-2
HDAC11	1.03627969	0.02874569	Histone deacetylase 11
PAIP2B	1.08709963	0.02896538	Polyadenylate-binding protein-interacting protein 2B
F12	1.08032596	0.02929618	Coagulation factor XII
FGA	-1.0476624	0.02933998	Fibrinogen alpha chain
APOBEC1	2.24566371	0.03029859	C->U-editing enzyme APOBEC-1
DNAH17	1.68021713	0.03046179	Dynein heavy chain 17, axonemal
MR1	1.08211457	0.03046179	Major histocompatibility complex class I-related gene protein
TSPAN18	1.39905342	0.03046179	Tetraspanin-18
RGS9	2.1003938	0.03046859	Regulator of G-protein signaling 9

FAM84A	1.99986416	0.03058409	Protein FAM84A
B3GNT7	1.28035684	0.03076871	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 7
IKBKG	-1.1445755	0.03082956	NF-kappa-B essential modulator
RHOV	1.60298042	0.03082956	Rho-related GTP-binding protein RhoV
RRAGD	1.49136006	0.03082956	Ras-related GTP-binding protein D
HMGN5	1.23446928	0.03221588	High mobility group nucleosome-binding domain-containing protein 5
MME	1.89591187	0.03302484	Neprilysin
ARHGAP44	1.20900608	0.03309731	Rho GTPase-activating protein 44
SFRP5	1.1695917	0.03353534	Secreted frizzled-related protein 5
ADRA1D	-1.5427394	0.03411001	Alpha-1D adrenergic receptor
EPN3	1.46868788	0.034154	Epsin-3
WIPF3	1.47298227	0.03428631	WAS/WASL-interacting protein family member 3
GUCY2D	1.82025551	0.03520564	Retinal guanylyl cyclase 1
C1orf53	-1.3048208	0.03526229	Uncharacterized protein C1orf53
EIF3C	-1.3614854	0.03542649	Eukaryotic translation initiation factor 3 subunit C
SCML4	2.15398516	0.03587105	Sex comb on midleg-like protein 4
CKB	1.05544595	0.03629987	Creatine kinase B-type
NR1H4	1.09973663	0.03629987	Bile acid receptor
CCDC80	1.51539299	0.03688038	Coiled-coil domain-containing protein 80
GNB3	1.18445663	0.03743836	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-3
NAV2	1.46895491	0.03824773	Neuron navigator 2
EFNB3	1.42468673	0.03826709	Ephrin-B3
NFATC1	1.55561021	0.03826709	Nuclear factor of activated T-cells, cytoplasmic 1
SV2B	2.51870235	0.03897892	Synaptic vesicle glycoprotein 2B
SYTL5	1.12860262	0.03897892	Synaptotagmin-like protein 5
TBX6	1.42631545	0.03981198	T-box transcription factor TBX6
ATP8B2	2.56671328	0.04059726	Phospholipid-transporting ATPase ID
MIR1199	1.39761158	0.04091748	Uncharacterized protein MISP3
SYNPO	1.07855175	0.04091748	Synaptopodin
AQP7	1.23838257	0.0413267	Aquaporin-7
SLC22A7	2.09351838	0.04153774	Solute carrier family 22 member 7
PMM1	1.04912756	0.04155773	Phosphomannomutase 1
S100A1	1.34268362	0.04194949	Protein S100-A1
CLIC3	1.32706785	0.04255273	Chloride intracellular channel protein 3
MPP2	1.89402992	0.04305344	MAGUK p55 subfamily member 2
IQSEC3	1.35776479	0.04342793	IQ motif and SEC7 domain-containing protein 3
AZGP1	1.15478166	0.04378275	Zinc-alpha-2-glycoprotein
CIART	1.2447061	0.04401734	Circadian-associated transcriptional repressor
SULT2B1	2.29560377	0.04404245	Sulfotransferase family cytosolic 2B member 1
CDON	1.01031901	0.04462369	Cell adhesion molecule-related/down-regulated by oncogenes
PPP1R14C	1.43303337	0.04470084	Protein phosphatase 1 regulatory subunit 14C

RASL11A	1.19496063	0.04538676	Ras-like protein family member 11A
VLDLR	1.35761353	0.04589815	Very low-density lipoprotein receptor
BLVRA	1.13670314	0.04595355	Biliverdin reductase A
MNS1	1.02450288	0.04636519	Meiosis-specific nuclear structural protein 1
ADAM23	-1.1064376	0.04717926	Disintegrin and metalloproteinase domain-containing protein 23
ZNF695	1.34792799	0.04761861	Zinc finger protein 695
DSC3	-1.2123436	0.04860591	Desmocollin-3
PDZRN3	1.27847004	0.04862421	E3 ubiquitin-protein ligase PDZRN3
SRXN1	-1.2238188	0.0493329	Sulfiredoxin-1
BSND	1.51287391	0.04954363	Barttin

Appendix table 14. Human differentially expressed genes between uninfected controls and infected samples at 24 h post infection.

Gene name	log_2(fold change)	padj	annotation
RBM3	-1.4573367	9.40E-10	RNA-binding protein 3
TM4SF20	2.25274744	9.40E-10	Transmembrane 4 L6 family member 20
APOA4	2.99113736	4.88E-08	Apolipoprotein A-IV
SLC6A12	2.84209984	2.03E-07	Sodium- and chloride-dependent betaine transporter
AKR1C1	1.24752684	6.87E-06	Aldo-keto reductase family 1 member C1
CREB3L3	1.84261482	6.37E-05	Cyclic AMP-responsive element-binding protein 3-like protein 3
PTPRS	1.92851987	6.37E-05	Receptor-type tyrosine-protein phosphatase S
ANPEP	1.60090135	0.00010279	Aminopeptidase N
EGR1	-1.9195402	0.00014659	Early growth response protein 1
HSPB1	1.56282758	0.00039161	Heat shock protein beta-1
LGALS1	1.92299236	0.00056212	Galectin-1
AKR1D1	1.87968051	0.00056212	3-oxo-5-beta-steroid 4-dehydrogenase
AQP3	1.58192301	0.00056212	Aquaporin-3
AMACR	1.15991841	0.00056212	Alpha-methylacyl-CoA racemase
TNFRSF12A	-1.1921195	0.00122472	Tumor necrosis factor receptor superfamily member 12A
PSAPL1	-1.534337	0.00122472	Proactivator polypeptide-like 1
TLN2	-1.0031113	0.00139838	Talin-2
TMEM63C	1.52272713	0.00213564	Calcium permeable stress-gated cation channel 1
HTR1D	1.58282728	0.00213564	5-hydroxytryptamine receptor 1D
KIF20A	1.11119455	0.00213868	Kinesin-like protein KIF20A
PLA2G2A	2.30226346	0.00247779	Phospholipase A2, membrane associated
SLC2A13	1.5594337	0.00248979	Proton myo-inositol cotransporter
MBL2	2.35574326	0.00248979	Mannose-binding protein C
CDH16	1.33971673	0.00248979	Cadherin-16

CDA	1.7435047	0.00257381	Cytidine deaminase
RAPGEFL1	-1.0260722	0.003042	Rap guanine nucleotide exchange factor-like 1
DACH1	1.58185303	0.003042	Dachshund homolog 1
CP	1.54361086	0.00327224	Ceruloplasmin
TBC1D1	1.02102674	0.00377048	TBC1 domain family member 1
HMGCS2	2.26884183	0.00424972	Hydroxymethylglutaryl-CoA synthase, mitochondrial
PIPOX	1.40190416	0.00450494	Peroxisomal sarcosine oxidase
CDK14	1.03661772	0.00457045	Cyclin-dependent kinase 14
HCN2	1.37537633	0.00457045	Potassium/sodium hyperpolarization-activated cyclic nucleotide-gated channel 2
CYP4F3	1.62735711	0.00693697	Docosahexaenoic acid omega-hydroxylase CYP4F3
CREG2	-1.4178413	0.00839847	Protein CREG2
GUCA2A	1.03129795	0.0084813	Guanylin
LRRC31	1.0531158	0.00936214	Leucine-rich repeat-containing protein 31
TTC7B	1.35090828	0.00962842	Tetratricopeptide repeat protein 7B
SH2D5	-1.4563388	0.0102394	SH2 domain-containing protein 5
NR1H4	1.58912913	0.01029858	Bile acid receptor
MFAP3L	1.41257711	0.01039049	Microfibrillar-associated protein 3-like
OAS1	1.05810229	0.0106891	2'-5'-oligoadenylate synthase 1
PHEX	1.11646237	0.0106891	Phosphate-regulating neutral endopeptidase
SLC38A4	2.31466234	0.0106891	Sodium-coupled neutral amino acid transporter 4
NFE2	1.7164499	0.01215148	Transcription factor NF-E2 45 kDa subunit
BTD	1.06128197	0.01246828	Biotinidase
ZNF117	1.24980564	0.01375101	Zinc finger protein 117
B3GNT7	1.64457203	0.01422908	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 7
FOSL1	-1.458734	0.01422908	Fos-related antigen 1
UGT1A6	1.15380667	0.0153192	UDP-glucuronosyltransferase 1-6
SULF2	1.03863272	0.0153192	Extracellular sulfatase Sulf-2
HIP1	1.3201257	0.01574583	Huntingtin-interacting protein 1
PGBD5	1.18848031	0.01574583	PiggyBac transposable element-derived protein 5
ADRA1D	-2.6450757	0.01583184	Alpha-1D adrenergic receptor
ADAM23	-1.5035947	0.01677284	Disintegrin and metalloproteinase domain-containing protein 23
TBX4	1.77416351	0.01771758	T-box transcription factor TBX4
ITPR1	1.54475623	0.01855909	Inositol 1,4,5-trisphosphate receptor type 1
GAS2	1.02654513	0.01978439	Growth arrest-specific protein 2
HIST2H2BF	1.1979795	0.01978439	Histone H2B type 2-F
PADI2	1.31988643	0.01979361	Protein-arginine deiminase type-2
CACNA1G	1.89334077	0.01983583	Voltage-dependent T-type calcium channel subunit alpha-1G
DAB1	1.09829845	0.02013313	Disabled homolog 1
CYP19A1	2.99436492	0.02118196	Aromatase
CD3G	1.68227827	0.02151844	T-cell surface glycoprotein CD3 gamma chain

PFKFB1	1.51583108	0.02427014	6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 1
HIST1H2BF	1.10460295	0.0294735	Histone cluster 1 H2B family member f
ATP1B2	1.64454248	0.03177459	Sodium/potassium-transporting ATPase subunit beta-2
TMCC3	2.43538673	0.03218809	Transmembrane and coiled-coil domain containing
SLC31A2	1.12613912	0.03218809	Probable low affinity copper uptake protein 2
MUC17	1.13740613	0.0322699	Mucin-17
MOGAT2	2.71020155	0.03780809	2-acylglycerol O-acyltransferase 2
ABAT	1.1054448	0.03809522	4-aminobutyrate aminotransferase, mitochondrial
RHBG	1.50974705	0.03822678	Ammonium transporter Rh type B
FSCN1	1.14840571	0.04423455	Fascin
SH2B3	-1.0694087	0.04423455	SH2B adapter protein 3
CYP4F12	1.18930089	0.04423455	Cytochrome P450 4F12
ZNF350	1.89408842	0.04490392	Zinc finger protein 350
FAM84A	2.65731031	0.04603839	Protein FAM84A
PCK1	2.10086997	0.04773825	Phosphoenolpyruvate carboxykinase, cytosolic [GTP]
P2RY6	1.65481395	0.04852702	P2Y purinoceptor 6
MUC13	1.62192447	0.04852702	Mucin-13
SI	-1.0899279	0.04896973	Sucrase-isomaltase, intestinal

Appendix table 15. Human differentially expressed genes between the 3 h and 24 h infected samples, with the genes differentially expressed between the 3 h and 24 h uninfected control samples removed.

Gene name	log_2(fold change)	padj	annotation
ARL4C	-1.8818047	7.45E-21	ADP-ribosylation factor-like protein 4C
GNRH2	-3.1352049	7.45E-21	Progonadoliberin-2
CCND3	-2.4467493	1.67E-18	G1/S-specific cyclin-D3
REEP2	-2.9422557	2.10E-17	Receptor expression-enhancing protein 2
MSR1	-1.9842792	5.15E-16	Macrophage scavenger receptor types I and II
PDZD7	-2.2446353	1.30E-13	PDZ domain-containing protein 7
EHD1	-1.3856707	1.09E-11	EH domain-containing protein 1
MYLIP	-1.627914	1.34E-11	E3 ubiquitin-protein ligase MYLIP
BCAS4	-2.468792	1.51E-11	Breast carcinoma amplified sequence 4
VDR	-1.6674314	4.92E-11	Vitamin D3 receptor
LPCAT2	-1.324178	1.59E-10	Lysophosphatidylcholine acyltransferase 2
C19orf26	-1.7022516	1.74E-10	Voltage-dependent calcium channel beta subunit-associated regulatory protein
PCSK5	-1.1909273	3.24E-10	Proprotein convertase subtilisin/kexin type 5

FSTL3	-2.588119	3.69E-10	Follistatin-related protein 3
ANKRD33B	-1.9035797	5.19E-10	Ankyrin repeat domain-containing protein 33B
ASUN	1.45705624	6.60E-10	Integrator complex subunit 13
SLC22A11	2.89008526	1.02E-09	Solute carrier family 22 member 11
ATP8A1	-1.4820933	1.05E-09	Phospholipid-transporting ATPase IA
ANKRD42	-2.1395525	1.10E-09	Ankyrin repeat domain-containing protein 42
LIMK2	-1.1906356	1.22E-09	LIM domain kinase 2
USP2	-2.3701397	2.22E-09	Ubiquitin carboxyl-terminal hydrolase 2
CHTF8	1.34805429	2.29E-09	Chromosome transmission fidelity protein 8 homolog
C2CD4A	3.23444802	2.70E-09	C2 calcium-dependent domain-containing protein 4A
CYP2U1	-2.0639237	3.45E-09	Cytochrome P450 2U1
GATS	-1.6847278	4.48E-09	Putative protein GATS
MAP3K14	-1.6060019	5.09E-09	Mitogen-activated protein kinase kinase kinase 14
GALNT12	-1.8615027	7.01E-09	Polypeptide N-acetylgalactosaminyltransferase 12
ATP1B2	-2.458784	8.50E-09	Sodium/potassium-transporting ATPase subunit beta-2
NOS2	2.17146903	9.15E-09	Nitric oxide synthase, inducible
ENDOD1	-1.2350487	9.39E-09	Endonuclease domain-containing 1 protein
CITED2	-1.6049787	1.02E-08	Cbp/p300-interacting transactivator 2
ADAMTS15	-2.859615	1.29E-08	ADAM metallopeptidase with thrombospondin type 1 motif 15
DOCK9	1.02054805	1.63E-08	Dedicator of cytokinesis protein 9
RGL1	-1.8146466	2.12E-08	Ral guanine nucleotide dissociation stimulator-like 1
PHLDB2	-1.269922	2.24E-08	Pleckstrin homology-like domain family B member 2
KCNAB2	-1.5457254	2.90E-08	Voltage-gated potassium channel subunit beta-2
CDC14A	-1.7082542	4.46E-08	Dual specificity protein phosphatase CDC14A
SLC4A8	-2.1979652	5.73E-08	Electroneutral sodium bicarbonate exchanger 1
KREMEN1	-1.5201724	5.84E-08	Kremen protein 1
HSD17B7	1.219508	6.35E-08	3-keto-steroid reductase
ETNK2	-1.0514953	7.36E-08	Ethanolamine kinase 2
C1orf21	-1.2603685	7.36E-08	Uncharacterized protein C1orf21
ID2	-1.4412867	7.36E-08	DNA-binding protein inhibitor ID-2
HOXC4	-1.8086286	1.37E-07	Homeobox protein Hox-C4
NAAA	-1.3948038	1.40E-07	N-acylethanolamine-hydrolyzing acid amidase
C6	-2.2129324	1.58E-07	Complement component C6
TMEM170B	-1.0849564	1.84E-07	Transmembrane protein 170B
TMEM169	-2.1037581	1.98E-07	Transmembrane protein 169
CITED4	-2.2128589	2.02E-07	Cbp/p300-interacting transactivator 4
LBH	-1.8927602	2.11E-07	Protein LBH
RCAN1	-1.2375728	2.35E-07	Calcipressin-1
SMAD7	-1.7106733	2.90E-07	Mothers against decapentaplegic homolog 7
C10orf54	-1.2954157	2.92E-07	V-type immunoglobulin domain-containing suppressor of T-cell activation

GPRC6A	2.71719529	3.59E-07	G-protein coupled receptor family C group 6 member A
CASKIN1	-1.9248976	4.22E-07	Caskin-1
GLTP	-1.2817117	4.57E-07	Glycolipid transfer protein
THNSL1	1.18885477	4.63E-07	Threonine synthase like 1
EBP	1.00094167	4.71E-07	3-beta-hydroxysteroid-Delta(8),Delta(7)-isomerase
GDF7	-1.5089768	4.71E-07	Growth/differentiation factor 7
SLC22A31	-1.6618919	5.44E-07	Putative solute carrier family 22 member 31
SLC23A3	-2.0130421	5.44E-07	Solute carrier family 23 member 3
RBM14	1.0895154	5.48E-07	RNA-binding protein 14
PRDM10	1.42867673	7.12E-07	PR domain zinc finger protein 10
GADD45A	-1.9579893	7.38E-07	Growth arrest and DNA damage-inducible protein GADD45 alpha
ITGA7	-1.5573965	7.68E-07	Integrin alpha-7
TCEA2	-1.6768391	7.82E-07	Transcription elongation factor A protein 2
ARHGDIB	-1.9740083	8.26E-07	Rho GDP-dissociation inhibitor 2
FLVCR2	-1.6194282	8.86E-07	Feline leukemia virus subgroup C receptor-related protein 2
TAGLN	-1.7810544	9.07E-07	Transgelin
EXT1	-1.2794001	1.01E-06	Exostosin-1
SARM1	-1.358182	1.14E-06	Sterile alpha and TIR motif-containing protein 1
ACADSB	-1.262606	1.19E-06	Short/branched chain specific acyl-CoA dehydrogenase, mitochondrial
RAPGEFL1	-1.3900872	1.19E-06	Rap guanine nucleotide exchange factor-like 1
HAPLN3	-2.0333731	1.21E-06	Hyaluronan and proteoglycan link protein 3
BACE1	-1.6602408	1.38E-06	Beta-secretase 1
GPRC5B	-1.6314611	1.63E-06	G-protein coupled receptor family C group 5 member B
FZD4	-1.3461186	1.81E-06	Frizzled-4
EVI5L	-1.6475238	2.53E-06	EVI5-like protein
HGNC:9982	1.00029177	2.71E-06	MHC class II regulatory factor RFX1
KIF26A	-1.4053334	2.71E-06	Kinesin-like protein KIF26A
TSPAN18	-1.7409533	3.00E-06	Tetraspanin-18
ANXA6	-1.1415084	3.04E-06	Annexin A6
C16orf87	-1.3884225	3.09E-06	UPF0547 protein C16orf87
CLDN19	-2.0390353	3.38E-06	Claudin-19
RAB12	-1.1138428	3.55E-06	Ras-related protein Rab-12
NDRG4	-2.0128898	4.20E-06	Protein NDRG4
NPC1L1	1.44607768	4.28E-06	Niemann-Pick C1-like protein 1
TNNC1	-1.611234	4.34E-06	Troponin C, slow skeletal and cardiac muscles
PARVB	1.19173415	4.36E-06	Beta-parvin
B4GALNT1	1.55168543	4.76E-06	Beta-1,4 N-acetylgalactosaminyltransferase 1
MKKS	1.2126993	4.76E-06	McKusick-Kaufman/Bardet-Biedl syndromes putative chaperonin
NQO1	1.13483414	4.76E-06	NAD(P)H dehydrogenase [quinone] 1
AHNAK	-1.4655254	5.53E-06	Neuroblast differentiation-associated protein AHNAK

FAM126B	-1.1103928	5.70E-06	Protein FAM126B
LCA5	-1.271671	6.42E-06	Lebercilin
STARD13	-1.2538583	6.64E-06	StAR-related lipid transfer protein 13
OXT	-1.5826701	6.76E-06	Oxytocin-neurophysin 1
CRYL1	-1.0043154	7.39E-06	Lambda-crystallin homolog
HSD17B14	-1.6682413	7.43E-06	17-beta-hydroxysteroid dehydrogenase 14
SLX4IP	1.24382896	8.32E-06	SLX4 interacting protein
FGFBP3	1.31353961	8.48E-06	Fibroblast growth factor-binding protein 3
FASTKD1	1.19265237	8.91E-06	FAST kinase domain-containing protein 1, mitochondrial
NUMBL	-1.2799119	9.06E-06	Numb-like protein
WIPF3	-1.9961971	9.06E-06	WAS/WASL-interacting protein family member 3
CYS1	-1.8474926	9.54E-06	Cystin-1
OGFRL1	-1.144457	1.01E-05	Opioid growth factor receptor like 1
LEAP2	-1.818212	1.01E-05	Liver-expressed antimicrobial peptide 2
IYD	-1.3294117	1.07E-05	Iodotyrosine deiodinase 1
HOXA1	-1.7656016	1.13E-05	Homeobox protein Hox-A1
CD99	-2.753025	1.13E-05	CD99 antigen
ENPP3	-1.8657815	1.16E-05	Ectonucleotide pyrophosphatase/phosphodiesterase family member 3
PRX	-1.7929967	1.20E-05	Periaxin
SFXN4	1.1624572	1.21E-05	Sideroflexin-4
TUFT1	-1.1385616	1.21E-05	Tuftelin
PRSS35	-1.654604	1.21E-05	Serine protease 35
CRY1	-1.0891391	1.22E-05	Cryptochrome-1
SPTLC3	-1.9217304	1.35E-05	Serine palmitoyltransferase 3
BAIAP3	-1.9822217	1.38E-05	BAI1 associated protein 3
ANK2	-1.5463319	1.39E-05	Ankyrin-2
CDK18	-1.4525672	1.40E-05	Cyclin-dependent kinase 18
PARM1	-1.425407	1.41E-05	Prostate androgen-regulated mucin-like protein 1
IGFBP1	2.96517433	1.50E-05	Insulin-like growth factor-binding protein 1
LIPC	-1.5096887	1.54E-05	Hepatic triacylglycerol lipase
SPSB1	-1.3432463	1.56E-05	SPRY domain-containing SOCS box protein 1
GPX2	2.08316243	1.61E-05	Glutathione peroxidase 2
ZNF608	-1.1880128	1.61E-05	Zinc finger protein 608
TIAM1	-1.7767724	1.66E-05	T-lymphoma invasion and metastasis-inducing protein 1
RHOBTB2	-1.1253224	1.77E-05	Rho-related BTB domain-containing protein 2
KHK	-1.4911674	1.79E-05	Ketohexokinase
CCSER1	1.341502	1.81E-05	Coiled-coil serine rich protein 1
KIAA0040	-1.9446771	1.84E-05	Uncharacterized protein KIAA0040
MRPL15	1.10932595	1.89E-05	Mitochondrial ribosomal protein L15
ELMSAN1	-1.2863673	1.89E-05	ELM2 and Myb/SANT domain containing 1

ARHGEF4	-1.7578004	1.95E-05	Rho guanine nucleotide exchange factor 4
DYNC2LI1	-1.3430559	2.29E-05	Cytoplasmic dynein 2 light intermediate chain 1
PIR	1.39164773	2.36E-05	Pirin
BAIAP2	-1.1331674	2.47E-05	Brain-specific angiogenesis inhibitor 1-associated protein 2
CEACAM20	-2.6760726	2.51E-05	Carcinoembryonic antigen-related cell adhesion molecule 20
JDP2	-1.2318617	2.53E-05	Jun dimerization protein 2
C15orf48	-1.4971469	2.53E-05	Normal mucosa of esophagus-specific gene 1 protein
XPA	-1.5194486	2.62E-05	DNA repair protein complementing XP-A cells
TTYH2	-2.1302297	2.66E-05	Protein tweety homolog 2
KIFC3	-1.0237442	2.73E-05	Kinesin-like protein KIFC3
FZD5	-1.0246988	2.75E-05	Frizzled-5
NKAIN4	-2.5910981	2.83E-05	Sodium/potassium-transporting ATPase subunit beta-1-interacting protein 4
NPNT	-1.2686346	3.03E-05	Nephronectin
TMEM185A	-1.2266788	3.06E-05	Transmembrane protein 185A
SORCS2	-1.9447403	3.15E-05	Sortilin related VPS10 domain containing receptor 2
ALDH3B1	-1.0535988	3.17E-05	Aldehyde dehydrogenase family 3 member B1
RRAGD	-2.4424981	3.38E-05	Ras-related GTP-binding protein D
SSX2IP	-1.0107935	3.39E-05	Afadin- and alpha-actinin-binding protein
XPOT	1.28829083	3.43E-05	Exportin-T
JPH2	-3.0430025	3.43E-05	Junctophilin-2
KIF7	-1.6026047	3.49E-05	Kinesin-like protein KIF7
NME1	1.32555927	3.74E-05	Nucleoside diphosphate kinase A
ECM2	-1.5418145	3.76E-05	Extracellular matrix protein 2
SH3BGRL	-1.0615383	3.77E-05	SH3 domain-binding glutamic acid-rich-like protein
PPP2R3A	-1.1678969	3.83E-05	Serine/threonine-protein phosphatase 2A regulatory subunit B'' subunit alpha
DAO	-1.6371074	3.83E-05	D-amino-acid oxidase
RILPL2	-1.6683634	3.83E-05	RILP-like protein 2
ZNF697	-1.0072905	3.98E-05	Zinc finger protein 697
AQP7	-1.424884	4.12E-05	Aquaporin-7
TMEM25	-1.2289558	4.53E-05	Transmembrane protein 25
AIMP2	1.09481294	4.76E-05	Aminoacyl tRNA synthase complex-interacting multifunctional protein 2
WTIP	-1.2365689	4.79E-05	Wilms tumor protein 1-interacting protein
POTEF	-1.3599869	5.05E-05	POTE ankyrin domain family member F
CDK20	-1.7180239	5.06E-05	Cyclin-dependent kinase 20
TLN2	-1.0424114	5.17E-05	Talin-2
COL12A1	-1.6185669	5.17E-05	Collagen alpha-1(XII) chain
KCNE3	1.2243797	5.63E-05	Potassium voltage-gated channel subfamily E member 3
CFAP69	-1.8125977	5.95E-05	Cilia- and flagella-associated protein 69
WWC1	-1.0979523	6.70E-05	Protein KIBRA
SFXN3	-1.1926077	6.70E-05	Sideroflexin-3

FADS6	-1.3448373	6.95E-05	Fatty acid desaturase 6
SGCB	-1.0992769	7.04E-05	Beta-sarcoglycan
JPH3	-1.3551178	7.14E-05	Junctophilin-3
PLA2G15	-1.0696017	7.46E-05	Group XV phospholipase A2
CTNNBIP1	-1.415049	7.95E-05	Beta-catenin-interacting protein 1
TIPARP	-1.7454898	8.29E-05	TCDD-inducible poly [ADP-ribose] polymerase
RELL1	-1.1079357	8.43E-05	RELT-like protein 1
SLC8B1	-1.045306	8.82E-05	Mitochondrial sodium/calcium exchanger protein
ZNF275	-1.237236	8.94E-05	Zinc finger protein 275
SERPINI1	-1.8403257	8.96E-05	Neuroserpin
PTPRU	-1.2290075	9.11E-05	Receptor-type tyrosine-protein phosphatase U
NCKAP5L	-1.1230695	9.26E-05	Nck-associated protein 5-like
DPCD	-1.1827642	9.42E-05	Protein DPCD
TNFRSF12A	-1.0406714	9.49E-05	Tumor necrosis factor receptor superfamily member 12A
CACNG4	-1.8663904	9.71E-05	Voltage-dependent calcium channel gamma-4 subunit
MRO	-2.0500521	9.75E-05	Maestro heat like repeat containing
GRIN3B	-1.9464108	0.00010905	Glutamate receptor ionotropic, NMDA 3B
FCHSD1	-1.0987043	0.00011326	FCH and double SH3 domains 1
VGLL3	-1.4924031	0.00011326	Transcription cofactor vestigial-like protein 3
CCDC85C	-1.0426722	0.0001165	Coiled-coil domain-containing protein 85C
SLC5A11	-1.2398915	0.00012684	Sodium/myo-inositol cotransporter 2
TPST2	-1.0479785	0.00012695	Protein-tyrosine sulfotransferase 2
SLC25A29	-1.6959682	0.00012695	Mitochondrial basic amino acids transporter
NEURL1B	-1.5456433	0.00013068	E3 ubiquitin-protein ligase NEURL1B
DPY19L1	1.1113195	0.00013643	Probable C-mannosyltransferase DPY19L1
SGTB	-1.0560631	0.00013987	Small glutamine-rich tetratricopeptide repeat-containing protein beta
RHBG	-1.3215962	0.00014239	Ammonium transporter Rh type B
IDE	1.0900806	0.00014434	Insulin-degrading enzyme
PARD6B	-1.0452973	0.00015048	Partitioning defective 6 homolog beta
ACOX3	-1.1200633	0.00015456	Peroxisomal acyl-coenzyme A oxidase 3
PLAC8	-1.4680201	0.00015657	Placenta-specific gene 8 protein
BTN2A2	-1.8309443	0.00015799	Butyrophilin subfamily 2 member A2
PLEKHH3	-1.2714893	0.00017537	Pleckstrin homology, MyTH4 and FERM domain containing H3
SLC35G1	-1.0673771	0.00018284	Solute carrier family 35 member G1
SLC27A1	-1.3111115	0.00018533	Long-chain fatty acid transport protein 1
CCDC126	-1.0819645	0.00018849	Coiled-coil domain containing 126
OTUD1	-1.4515092	0.00019093	OTU domain-containing protein 1
AMOTL1	-1.0326823	0.00019389	Angiomotin-like protein 1
TMEM5	1.14774732	0.00019608	UDP-D-xylose:ribitol-5-phosphate beta1,4-xylosyltransferase
UAP1L1	-1.0592756	0.00019674	UDP-N-acetylhexosamine pyrophosphorylase-like protein 1

LAPTM5	-1.4259238	0.00019674	Lysosomal-associated transmembrane protein 5
EPPK1	-1.3833206	0.00019755	Epiplakin
MCAM	-1.6673994	0.00019755	Cell surface glycoprotein MUC18
SYNGR1	-1.6949426	0.00020009	Synaptogyrin-1
PITPNM3	-1.5279971	0.00020189	Membrane-associated phosphatidylinositol transfer protein 3
WDR19	-1.1706088	0.00020715	WD repeat-containing protein 19
ASPA	-1.1510394	0.0002089	Aspartoacylase
MAP3K8	-1.2071267	0.00021423	Mitogen-activated protein kinase kinase kinase 8
ABCA5	-1.0708747	0.00022197	ATP-binding cassette sub-family A member 5
AQP11	-1.6808145	0.00022853	Aquaporin-11
BCAS1	-1.8206671	0.00023025	Breast carcinoma amplified sequence 1
PELI3	-1.3863271	0.00023419	E3 ubiquitin-protein ligase pellino homolog 3
CHCHD7	-1.0739805	0.00023928	Mitochondrial coiled-coil-helix-coiled-coil-helix domain containing proteins
ACAT2	1.09989878	0.00024225	Acetyl-CoA acetyltransferase, cytosolic
ARMC6	1.15429063	0.00027034	Armadillo repeat containing 6
ITGB2	-1.2914544	0.00027705	Integrin beta-2
IL21R	2.23058391	0.00028591	Interleukin-21 receptor
RND1	-1.6350323	0.00028591	Rho-related GTP-binding protein Rho6
GDPD2	-1.2444074	0.00029363	Glycerophosphoinositol inositolphosphodiesterase GDPD2
SLC30A10	-1.2923476	0.0003054	Zinc transporter 10
ACOT7	-1.0520559	0.00032632	Cytosolic acyl coenzyme A thioester hydrolase
KLB	-2.2385377	0.00032632	Beta-klotho
TMEM176B	-1.0749819	0.00034763	Transmembrane protein 176B
BOLA3	1.05833708	0.00036376	BolA-like protein 3
GRID1	-1.1662642	0.00036481	Glutamate receptor ionotropic, delta-1
PKIG	-1.3857587	0.00036685	cAMP-dependent protein kinase inhibitor gamma
LAMA1	-2.2264097	0.00036685	Laminin subunit alpha-1
AGR2	1.28453141	0.00036695	Anterior gradient protein 2 homolog
NFKB2	-1.059878	0.00037255	Nuclear factor NF-kappa-B p100 subunit
CCDC141	-1.7696666	0.00040867	Coiled-coil domain-containing protein 141
SLC25A42	-1.3838295	0.00040948	Mitochondrial coenzyme A transporter SLC25A42
FADS2	1.13450429	0.00041554	Fatty acid desaturase 2
PLEKHM3	-1.0459657	0.00042829	Pleckstrin homology domain containing M3
NETO2	1.74507678	0.00044559	Neuropilin and tolloid-like protein 2
TIFA	-1.2602829	0.00046607	TRAF-interacting protein with FHA domain-containing protein A
LYPD6	-1.3446446	0.00048058	Ly6/PLAUR domain-containing protein 6
IGFBP4	-1.2883381	0.00049204	Insulin-like growth factor-binding protein 4
AHRR	-1.7585463	0.0005084	Aryl hydrocarbon receptor repressor
GLDN	-1.8494596	0.0005091	Gliomedin
AKAP5	-1.4492432	0.0005309	A-kinase anchor protein 5
PMEL	-1.5168429	0.00054297	Melanocyte protein PMEL

PADI2	-1.2641368	0.00054585	Protein-arginine deiminase type-2
OSBPL7	-1.1595815	0.00055021	Oxysterol-binding protein-related protein 7
GBP2	-1.2749168	0.00056132	Guanylate-binding protein 2
BIRC3	-1.8342995	0.00056334	Baculoviral IAP repeat-containing protein 3
LOX	-1.3720791	0.00056988	Protein-lysine 6-oxidase
SUGCT	-1.4242055	0.00056988	Succinate--hydroxymethylglutarate CoA-transferase
TSHZ1	-1.1316912	0.0006019	Teashirt homolog 1
LYRM9	-1.2066057	0.00062684	LYR motif containing 9
CADM1	-1.2522712	0.00063788	Cell adhesion molecule 1
LRRC75A	-1.0838052	0.00064085	Leucine-rich repeat-containing protein 75A
ZNF555	-1.1217013	0.00064085	Zinc finger protein 555
CCDC77	1.03781628	0.00066403	Coiled-coil domain containing 77
LHPP	-1.1050045	0.00067259	Phospholysine phosphohistidine inorganic pyrophosphate phosphatase
STC1	-1.5409402	0.00071879	Stanniocalcin-1
BBOX1	-2.1137174	0.00072876	Gamma-butyrobetaine dioxygenase
KIAA0319	-1.0405709	0.00073784	Dyslexia-associated protein KIAA0319
DBP	-1.519139	0.00074586	D site-binding protein
PBX1	-1.3891645	0.00074942	Pre-B-cell leukemia transcription factor 1
COL1A1	-1.5363321	0.00074974	Collagen alpha-1(I) chain
IFT22	-1.0443673	0.00075899	Intraflagellar transport protein 22 homolog
TNFSF4	-1.6215963	0.00081174	Tumor necrosis factor ligand superfamily member 4
TMEM151A	-2.1294809	0.00083056	Transmembrane protein 151A
DEPDC7	-1.0925488	0.00083338	DEP domain containing 7
C1QL1	-2.0656501	0.00084461	C1q-related factor
CCDC136	-1.6880451	0.00084921	Coiled-coil domain-containing protein 136
TBC1D4	1.09673939	0.00085198	TBC1 domain family member 4
ABAT	-1.1610266	0.0008827	4-aminobutyrate aminotransferase, mitochondrial
CYLD	-1.0560343	0.00088442	Ubiquitin carboxyl-terminal hydrolase CYLD
KCTD15	1.00642953	0.00088985	BTB/POZ domain-containing protein KCTD15
CD320	1.20270863	0.00089635	CD320 antigen
SORL1	-1.4436314	0.00089635	Sortilin-related receptor
RHOD	-1.2631996	0.00090373	Rho-related GTP-binding protein RhoD
IQCE	-1.0485623	0.00092608	IQ domain-containing protein E
TESK2	-1.0958937	0.0009295	Dual specificity testis-specific protein kinase 2
NOL4L	-1.0079071	0.00096912	Nucleolar protein 4 like
AJUBA	-1.0896475	0.00099511	LIM domain-containing protein ajuba
ERAP2	-1.2512991	0.00101475	Endoplasmic reticulum aminopeptidase 2
MEGF6	-1.594306	0.0010617	Multiple epidermal growth factor-like domains protein 6
ZNF554	-1.3448472	0.00107773	Zinc finger protein 554
PGAP3	-1.1431998	0.00108984	Post-GPI attachment to proteins factor 3
BAG2	1.24693823	0.00109251	BAG family molecular chaperone regulator 2

METTL3	1.12759855	0.00109682	N6-adenosine-methyltransferase catalytic subunit
MICALL2	-1.0436106	0.00112307	MICAL-like protein 2
SLC6A9	1.34274919	0.00112543	Sodium- and chloride-dependent glycine transporter 1
TRIM2	1.01721665	0.00114	Tripartite motif-containing protein 2
AHSA1	1.01774034	0.00117968	Activator of 90 kDa heat shock protein ATPase homolog 1
REL	-1.006604	0.0012245	Proto-oncogene c-Rel
SPEG	-1.572918	0.00125128	Striated muscle preferentially expressed protein kinase
KCNK15	-1.3991203	0.00126163	Potassium channel subfamily K member 15
HOXA3	-1.2011476	0.00128923	Homeobox protein Hox-A3
FRMD4B	-1.0990673	0.00130129	FERM domain-containing protein 4B
NRN1	-2.09856	0.00136394	Neuritin
RASSF4	-1.4995272	0.00138337	Ras association domain-containing protein 4
CEACAM1	1.52212668	0.00138471	Carcinoembryonic antigen-related cell adhesion molecule 1
TSPYL4	-1.0178095	0.00143494	Testis-specific Y-encoded-like protein 4
IRF1	-1.0740759	0.00144754	Interferon regulatory factor 1
CPNE8	1.04509977	0.00145343	Copine-8
TRIM62	-1.4826115	0.00145827	E3 ubiquitin-protein ligase TRIM62
KIAA0226L	-1.4915693	0.00146229	Protein RUBCNL-like
MARK4	-1.4487599	0.00150106	MAP/microtubule affinity-regulating kinase 4
SCRN2	-1.0371869	0.00151385	Secernin-2
GAS2	1.07955068	0.00155458	Growth arrest-specific protein 2
SRM	1.3046008	0.00157761	Spermidine synthase
SLC20A1	-1.1847485	0.00158529	Sodium-dependent phosphate transporter 1
HSD3B7	-1.182952	0.00160273	3 beta-hydroxysteroid dehydrogenase type 7
TICRR	1.14863868	0.00166663	Treslin
GVQW3	-1.1127624	0.0016924	annotation not available
RAB40B	-1.1076384	0.00169459	Ras-related protein Rab-40B
HSPB1	1.31666511	0.00173166	Heat shock protein beta-1
HABP4	-1.2908415	0.0017383	Intracellular hyaluronan-binding protein 4
EBF4	-1.5557388	0.00174754	Transcription factor COE4
ERCC6L	1.57472077	0.00175403	DNA excision repair protein ERCC-6-like
EFR3B	-1.155818	0.00175967	Protein EFR3 homolog B
CACHD1	-1.073539	0.00181942	VWFA and cache domain-containing protein 1
ADH4	1.41153234	0.00183424	Alcohol dehydrogenase 4, pi polypeptide
ACSS1	-1.3134439	0.00183734	Acetyl-coenzyme A synthetase 2-like, mitochondrial
CCDC170	-1.6924805	0.00184269	Coiled-coil domain containing 170
FOSB	-1.5689686	0.00186539	Protein fosB
ISG20	1.54793974	0.00188992	Interferon-stimulated gene 20 kDa protein
ABCC9	-1.7430808	0.00190721	ATP-binding cassette sub-family C member 9
CNTN4	-1.3673373	0.00196213	Contactin-4
CTSS	-1.18304	0.00197342	Cathepsin S

CPLX1	1.17677008	0.00198394	Complexin-1
IER5	-1.1813799	0.00204309	Immediate early response gene 5 protein
AKR1B10	2.80949263	0.00205699	Aldo-keto reductase family 1 member B10
KAZN	-1.0234784	0.00205699	Kazrin
BSN	-1.3744986	0.002069	Protein bassoon
CYP51A1	1.31851833	0.00211617	Lanosterol 14-alpha demethylase
RABEPK	1.31166883	0.00214036	Rab9 effector protein with kelch motifs
PLCXD2	-1.20428	0.00214036	PI-PLC X domain-containing protein 2
HS6ST2	1.1367835	0.00214213	Heparan-sulfate 6-O-sulfotransferase 2
CYR61	-1.6308791	0.0021503	Protein CYR61
DNAJC6	-1.3380642	0.00222606	Putative tyrosine-protein phosphatase auxilin
CYP4V2	-1.0449548	0.00230652	Cytochrome P450 4V2
HSPA8	1.18548779	0.00231603	Heat shock cognate 71 kDa protein
LRRC3	-1.3247819	0.00232911	Leucine-rich repeat-containing protein 3
NR1H4	-1.1614291	0.00236055	Bile acid receptor
AVEN	1.28293428	0.00236805	Cell death regulator Aven
PRSS23	-1.0709074	0.00237571	Serine protease 23
C4orf33	-1.162707	0.00242278	UPF0462 protein C4orf33
LMCD1	-1.1903921	0.00244497	LIM and cysteine-rich domains protein 1
HERC3	-1.2154001	0.0024525	Probable E3 ubiquitin-protein ligase HERC3
PMEPA1	-1.1264706	0.00247581	Protein TMEPAI
HMMR	1.1040958	0.00249637	Hyaluronan mediated motility receptor
FNTB	1.0834249	0.00250348	Protein farnesyltransferase subunit beta
LGR5	1.60712256	0.00265951	Leucine-rich repeat-containing G-protein coupled receptor 5
AGXT2	-1.2504148	0.00271079	Alanine--glyoxylate aminotransferase 2, mitochondrial
CRISPLD2	1.19067634	0.00273147	Cysteine-rich secretory protein LCCL domain-containing 2
GUCA2B	-1.4404897	0.00279904	Guanylate cyclase activator 2B
DENND3	-1.0742664	0.00281607	DENN domain-containing protein 3
LRFN1	-1.4421328	0.00281607	Leucine-rich repeat and fibronectin type III domain-containing protein 1
RHPN1	-1.1675494	0.00292306	Rhophilin-1
TPD52L1	-1.2846688	0.00302735	Tumor protein D52 like 1
TIMM21	1.06179067	0.00304343	Mitochondrial import inner membrane translocase subunit Tim21
NBPF19	-1.1171545	0.00305557	Neuroblastoma breakpoint family, member 14
GABRA2	-1.0122379	0.00306116	Gamma-aminobutyric acid receptor subunit alpha-2
ZBTB46	-1.0211035	0.00310376	Zinc finger and BTB domain-containing protein 46
MYCN	-1.559391	0.00319067	N-myc proto-oncogene protein
FAM151A	-1.0648206	0.00322147	Protein FAM151A
HSPH1	1.14744783	0.00326548	Heat shock protein 105 kDa
VIPR1	-1.0446173	0.00339226	Vasoactive intestinal polypeptide receptor 1
ALPP	-2.3021591	0.00341611	Alkaline phosphatase, placental

ITGA2	1.60534129	0.00343095	Integrin alpha-2
CAPN13	-1.2090753	0.00347596	Calpain-13
HMGN5	-1.5352429	0.00361865	High mobility group nucleosome-binding domain-containing protein 5
EML6	-1.409264	0.00363159	Echinoderm microtubule-associated protein-like 6
APOBEC3C	-1.327768	0.00363507	DNA dC->dU-editing enzyme APOBEC-3C
PLEKHA4	-1.118064	0.00379277	Pleckstrin homology domain-containing family A member 4
NBPF3	-1.022401	0.00386208	Neuroblastoma breakpoint family member 3
TLR2	-2.1607712	0.00387953	Toll-like receptor 2
CENPI	1.01474331	0.00396064	Centromere protein I
CACNA1G	-1.5486996	0.00404589	Voltage-dependent T-type calcium channel subunit alpha-1G
SMIM10	-1.2332397	0.00404662	Small integral membrane protein 10
SPAG1	-1.0233711	0.00407544	Sperm-associated antigen 1
RAB31	-1.4638276	0.00411401	Ras-related protein Rab-31
ATG16L2	-1.2950761	0.00418324	Autophagy-related protein 16-2
IFT81	-1.0177418	0.00418491	Intraflagellar transport protein 81 homolog
TMEM63C	-1.2592989	0.0041872	Calcium permeable stress-gated cation channel 1
MUC13	1.42238237	0.0041951	Mucin-13
LAMB2	-1.0943457	0.00429233	Laminin subunit beta-2
ARL14	-1.891192	0.00431357	ADP-ribosylation factor-like protein 14
NMNAT1	1.37011429	0.00432713	Nicotinamide/nicotinic acid mononucleotide adenylyltransferase 1
TSTD1	-1.1994737	0.0045465	Thiosulfate:glutathione sulfurtransferase
GADD45B	-1.5776648	0.00457149	Growth arrest and DNA damage-inducible protein GADD45 beta
PER1	-1.0027184	0.00460711	Period circadian protein homolog 1
URAD	-1.6989683	0.00467073	Putative 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline decarboxylase
SEMA6C	-1.7194057	0.00467073	Semaphorin-6C
S100A11	1.2376758	0.00470222	Protein S100-A11
SFRP5	-1.3914931	0.00482421	Secreted frizzled-related protein 5
USH1G	-1.3952693	0.00493289	Usher syndrome type-1G protein
FBXL16	-1.4491827	0.00511337	F-box/LRR-repeat protein 16
FOXO4	-1.311291	0.00522266	Forkhead box protein O4
KLF2	1.97954492	0.00529825	Krueppel-like factor 2
KCNJ8	-2.1168505	0.00545924	ATP-sensitive inward rectifier potassium channel 8
FLNC	-1.3902371	0.0056018	Filamin-C
ANKRD37	1.84362532	0.00563566	Ankyrin repeat domain-containing protein 37
MORN2	-1.3091906	0.00565096	MORN repeat-containing protein 2
LOC389831	1.08008067	0.0056668	Homo sapiens uncharacterized LOC389831 (LOC389831), mRNA
SULT1A1	-1.0943077	0.00568896	Sulfotransferase 1A1
UGT1A6	1.14836606	0.00571749	UDP-glucuronosyltransferase 1-6
EFNB3	-1.7460366	0.00577008	Ephrin-B3
SOX8	-1.0890377	0.0062028	Transcription factor SOX-8

CAMKK1	-1.2288221	0.00621025	Calcium/calmodulin-dependent protein kinase kinase 1
SLCO1A2	-1.3072663	0.00651271	Solute carrier organic anion transporter family member 1A2
HMHA1	-1.3060071	0.00654058	Rho GTPase-activating protein 45
SH2D2A	1.35429922	0.00658804	SH2 domain-containing protein 2A
PUS10	1.02900548	0.00659534	Putative tRNA pseudouridine synthase Pus10
MMP1	2.08917462	0.0066665	Interstitial collagenase
ISPD	-1.4419897	0.00672688	D-ribitol-5-phosphate cytidylyltransferase
SKIDA1	-1.4509724	0.00686923	SKI/DACH domain containing 1
SOCS3	-1.0439034	0.00704982	Suppressor of cytokine signaling 3
CEP41	1.08665205	0.00705127	Centrosomal protein of 41 kDa
SLC17A4	-1.0944598	0.00711549	Probable small intestine urate exporter
SH3PXD2B	-1.1763575	0.00711682	SH3 and PX domain-containing protein 2B
GEM	-1.5387246	0.00711682	GTP-binding protein GEM
CX3CL1	-1.3287376	0.00719109	Fractalkine
FAIM	1.05877068	0.00736595	Fas apoptotic inhibitory molecule 1
SGMS2	-1.0770814	0.00739254	Phosphatidylcholine:ceramide cholinephosphotransferase 2
ELMO1	1.45530248	0.00799073	Engulfment and cell motility protein 1
ENPP7	-1.1145501	0.00802737	Ectonucleotide pyrophosphatase/phosphodiesterase family member 7
SPTBN4	-1.8488618	0.0080349	Spectrin beta chain, non-erythrocytic 4
ANXA9	-1.1938651	0.00806968	Annexin A9
BIRC7	-1.9833366	0.00847506	Baculoviral IAP repeat-containing protein 7
PTPRS	1.19549871	0.00857147	Receptor-type tyrosine-protein phosphatase S
FRMD5	-1.114392	0.00861652	FERM domain-containing protein 5
ARL4D	-1.3152778	0.00914128	ADP-ribosylation factor-like protein 4D
DOK3	1.05206205	0.00933211	Docking protein 3
SESN3	-1.0646596	0.00936201	Sestrin-3
NEDD9	-1.5351987	0.00958453	Enhancer of filamentation 1
AXL	-1.2761707	0.00962234	Tyrosine-protein kinase receptor UFO
BRCA2	1.22321443	0.00966768	Breast cancer type 2 susceptibility protein
ASPDH	-1.0560001	0.0097882	Putative L-aspartate dehydrogenase
NT5DC4	1.13432875	0.00990942	5'-nucleotidase domain containing 4
CD302	-1.0260898	0.00992262	CD302 antigen
WNT11	-1.5038404	0.00992506	Protein Wnt-11
LDHD	-1.1191983	0.01014833	Probable D-lactate dehydrogenase, mitochondrial
HTR1D	1.09386399	0.01016248	5-hydroxytryptamine receptor 1D
RSPH1	-1.3210367	0.01018892	Radial spoke head 1 homolog
MASP1	-1.0010484	0.01029326	Mannan binding lectin serine peptidase 1
C1orf112	1.2965438	0.01029437	Uncharacterized protein C1orf112
ALPK2	-1.2053535	0.01029437	Alpha-protein kinase 2
SCAMP5	-1.0000123	0.01068154	Secretory carrier-associated membrane protein 5

PER2	-1.0141647	0.01069012	Period circadian protein homolog 2
ACKR3	-1.6630832	0.01069012	Atypical chemokine receptor 3
KIFC2	-1.0699257	0.0107055	Kinesin-like protein KIFC2
UTP20	1.21771357	0.01073835	Small subunit processome component 20 homolog
SLC16A10	-1.1520868	0.01111368	Monocarboxylate transporter 10
TMCC1	1.07447053	0.01120098	Transmembrane and coiled-coil domain containing
FAM110C	-1.2015972	0.01149331	Protein FAM110C
TTK	1.10133168	0.01155791	Dual specificity protein kinase TTK
DDIT3	-1.000335	0.01209187	DNA damage-inducible transcript 3 protein
TXNDC5	1.02297493	0.01229982	Thioredoxin domain-containing protein 5
BBS12	-1.1785234	0.01251656	Bardet-Biedl syndrome 12 protein
HIST2H2BF	1.02315333	0.01265353	Histone H2B type 2-F
ZNF441	-1.2236168	0.01265353	Zinc finger protein 441
ZFYVE28	-1.1234692	0.01287937	Lateral signaling target protein 2 homolog
EMP1	-1.0580713	0.01299697	Epithelial membrane protein 1
CASZ1	-1.1933594	0.01300359	Zinc finger protein castor homolog 1
RASGEF1B	-1.1380157	0.01309616	Ras-GEF domain-containing family member 1B
CDNF	-1.3992188	0.01321801	Cerebral dopamine neurotrophic factor
DUSP9	-1.1176925	0.01327068	Dual specificity protein phosphatase 9
FBLN2	1.75945737	0.01462712	Fibulin-2
VSNL1	-1.258275	0.01466497	Visinin-like protein 1
C3orf70	1.6334826	0.01477013	UPF0524 protein C3orf70
SLC5A1	-1.0556782	0.01481578	Sodium/glucose cotransporter 1
CAPN8	1.12064698	0.01526247	Calpain-8
IRX5	-1.1123389	0.01541705	Iroquois-class homeodomain protein IRX-5
MAPKBP1	-1.0405517	0.01544713	Mitogen-activated protein kinase-binding protein 1
CXCL2	-1.5381948	0.01567207	C-X-C motif chemokine 2
CAV2	-1.1624831	0.01601541	Caveolin-2
KLK1	-1.2147095	0.01632026	Kallikrein-1
ERICH4	-1.3685956	0.01658698	Glutamate-rich protein 4
QTRT1	1.02404851	0.01674288	Queuine tRNA-ribosyltransferase catalytic subunit 1
GRASP	-1.1716593	0.01674865	General receptor for phosphoinositides 1-associated scaffold protein
FAM174A	-1.3017919	0.01687007	Membrane protein FAM174A
MAP4K2	-1.0685168	0.01688068	Mitogen-activated protein kinase kinase kinase kinase 2
B3GNT7	-1.1349224	0.01691965	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 7
FNDC4	-1.1876533	0.01750672	Fibronectin type III domain-containing protein 4
HEATR5A	1.07743848	0.01760927	HEAT repeat-containing protein 5A
UBTD1	-1.0763216	0.01768398	Ubiquitin domain-containing protein 1
DEF6	-1.1305057	0.01785563	Differentially expressed in FDCP 6 homolog
RENBP	-1.0800128	0.0179721	N-acylglucosamine 2-epimerase
NUAK2	-1.2030145	0.01839258	NUAK family SNF1-like kinase 2

HIST1H1B	1.05662315	0.01840755	Histone H1.5
SERPINB8	1.33322924	0.01847754	Serpin B8
ARHGAP28	-1.0176459	0.01847754	Rho GTPase-activating protein 28
FAM46B	-1.166762	0.01858598	Putative nucleotidyltransferase FAM46B
PRDM1	-1.2072452	0.01860571	PR domain zinc finger protein 1
HILPDA	1.52042485	0.01860729	Hypoxia-inducible lipid droplet-associated protein
SH2B2	-1.0620403	0.018977	SH2B adapter protein 2
N4BP2L1	-1.249248	0.01916947	NEDD4 binding protein 2 like 1
FAM122C	-1.3307053	0.01942702	Protein FAM122C
ITIH2	1.14752859	0.01945785	Inter-alpha-trypsin inhibitor heavy chain H2
FGD3	-1.0094944	0.02003107	FYVE, RhoGEF and PH domain-containing protein 3
EBF3	-1.0741285	0.02011701	Transcription factor COE3
PCSK4	-1.1311036	0.02039105	Proprotein convertase subtilisin/kexin type 4
EMP3	1.13488101	0.02095865	Epithelial membrane protein 3
NMB	-1.1902081	0.0216189	Neuromedin-B
PCNA	1.08116765	0.0220092	Proliferating cell nuclear antigen
FABP1	1.03951913	0.0220092	Fatty acid-binding protein, liver
AGPAT9	1.03299743	0.02205452	Glycerol-3-phosphate acyltransferase 3
HIST1H2AB	1.37068546	0.02218239	Histone H2A type 1-B/E
DHDH	-1.1228216	0.02234639	Dihydrodiol dehydrogenase
BRSK1	-1.2904139	0.02268829	Serine/threonine-protein kinase BRSK1
CD3G	-1.1259178	0.02291863	T-cell surface glycoprotein CD3 gamma chain
PLA2G2A	1.13266589	0.02314687	Phospholipase A2, membrane associated
PALM	-1.008309	0.02374947	Paralemmin-1
FRS3	-1.0786366	0.02388193	Fibroblast growth factor receptor substrate 3
NCMAP	-1.737772	0.02412003	Noncompact myelin-associated protein
MUC12	1.06246893	0.02415675	Mucin-12
MOK	-1.3802359	0.0244363	MAPK/MAK/MRK overlapping kinase
BTNL8	-1.2013058	0.02447409	Butyrophilin-like protein 8
GPR128	-1.1767034	0.02507528	Adhesion G-protein coupled receptor G7
IGDCC4	-1.1115645	0.02609711	Immunoglobulin superfamily DCC subclass member 4
PCP4L1	-1.243114	0.02620577	Purkinje cell protein 4 like 1
PDZK1	-1.0430802	0.02659304	Na(+)/H(+) exchange regulatory cofactor NHE-RF3
HIST1H2BO	1.03223279	0.02679398	Histone H2B type 1-O
GNAZ	-1.0864936	0.02696122	Guanine nucleotide-binding protein G(z) subunit alpha
NR1I3	-1.4889218	0.02705592	Nuclear receptor subfamily 1 group I member 3
MEF2C	-1.1841587	0.02749147	Myocyte-specific enhancer factor 2C
ARHGEF18	-1.043243	0.02792071	Rho guanine nucleotide exchange factor 18
DYSF	-1.7209554	0.02831562	Dysferlin
ZNF853	-1.0556558	0.02874354	Zinc finger protein 853
CHST9	-1.3355582	0.02881318	Carbohydrate sulfotransferase 9

WDR91	-1.2434382	0.02890883	WD repeat-containing protein 91
DGKQ	-1.1725723	0.02907335	Diacylglycerol kinase theta
BTN1A1	1.05499009	0.02996675	Butyrophilin subfamily 1 member A1
ADPRM	-1.0319086	0.03043963	Manganese-dependent ADP-ribose/CDP-alcohol diphosphatase
CDON	-1.040951	0.03106482	Cell adhesion molecule-related/down-regulated by oncogenes
NFE2	1.17007001	0.03230819	Transcription factor NF-E2 45 kDa subunit
FTCD	-1.3397131	0.03268553	Formimidoyltransferase-cyclodeaminase
HBEGF	-1.1873464	0.03281269	Proheparin-binding EGF-like growth factor
CCDC159	-1.0932446	0.03299288	Coiled-coil domain containing 159
CENPQ	1.07024477	0.03328991	Centromere protein Q
COL4A5	-1.1794158	0.0335361	Collagen alpha-5(IV) chain
HIST1H3H	1.19954476	0.03441985	Histone cluster 1 H3 family member h
HIVEP2	-1.1066683	0.03449877	Transcription factor HIVEP2
ACAD11	-1.3028211	0.03509592	Acyl-CoA dehydrogenase family member 11
RIMS4	-1.1957271	0.03536498	Regulating synaptic membrane exocytosis protein 4
NCAPG	1.06320737	0.035443	Condensin complex subunit 3
CYP2C19	-1.154126	0.03568885	Cytochrome P450 2C19
NR4A1	-1.1959557	0.03577197	Nuclear receptor subfamily 4 group A member 1
LIN9	1.06979665	0.0357984	Protein lin-9 homolog
NBPF10	-1.4676624	0.0357984	Neuroblastoma breakpoint family member 10
TPPP	-2.491627	0.0362121	Tubulin polymerization-promoting protein
HIST1H2BM	1.20489362	0.03644997	Histone H2B type 1-M
SCARA3	-1.1644954	0.0367292	Scavenger receptor class A member 3
ADAMTS6	-1.2199796	0.03709368	ADAM metallopeptidase with thrombospondin type 1 motif 6
SMIM1	-1.4753534	0.03748592	Small integral membrane protein 1
HIST1H3B	1.07744082	0.0389002	Histone H3.1
SLC41A1	-1.0124272	0.03923426	Solute carrier family 41 member 1
DUSP1	-1.1322501	0.03951051	Dual specificity protein phosphatase 1
AADAC	1.18422423	0.04073255	Arylacetamide deacetylase
MAFF	-1.1029135	0.04132858	Transcription factor MafF
STPG1	-1.011702	0.04163662	O(6)-methylguanine-induced apoptosis 2
PPP1R3G	1.15700075	0.04175256	Protein phosphatase 1 regulatory subunit 3G
NUGGC	-1.4014079	0.04204348	Nuclear GTPase SLIP-GC
CCDC87	-1.01173	0.04240046	Coiled-coil domain containing 87
BMF	-1.0284897	0.04250647	Bcl-2-modifying factor
MBLAC2	-1.1540651	0.04250647	Metallo-beta-lactamase domain containing 2
GINS3	1.14857533	0.04279325	GINS complex subunit 3
ARMC9	-1.0623427	0.04434124	LisH domain-containing protein ARMC9
ARG2	1.18246263	0.04646041	Arginase-2, mitochondrial
KIAA1462	-1.5160052	0.04658047	Junctional protein associated with coronary artery disease
ENTPD2	-1.1362275	0.04667444	Ectonucleoside triphosphate diphosphohydrolase 2

TSPAN4	-1.1315974	0.0479894	Tetraspanin-4
ASAH2	-1.2495249	0.04877554	Neutral ceramidase
GDF15	1.46253087	0.04947731	Growth differentiation factor 15

Appendix table 16. Significantly enriched human pathways between the uninfected controls and infected samples at 3 h post infection.

Ingenuity Canonical Pathways	-log(p-value)	Ratio	z-score
Salvage Pathways of Pyrimidine Ribonucleotides	4,65E00	2,14E-01	0,894
Sertoli Cell-Sertoli Cell Junction Signaling	4,25E00	1,68E-01	#NUM!
Germ Cell-Sertoli Cell Junction Signaling	4,11E00	1,7E-01	#NUM!
ATM Signaling	3,72E00	1,96E-01	-0,500
CD27 Signaling in Lymphocytes	3,71E00	2,45E-01	1,387
Protein Ubiquitination Pathway	3,61E00	1,43E-01	#NUM!
IL-17A Signaling in Fibroblasts	3,56E00	2,86E-01	#NUM!
Senescence Pathway	3,55E00	1,42E-01	1,000
AMPK Signaling	3,42E00	1,5E-01	0,426
Pyridoxal 5'-phosphate Salvage Pathway	3,25E00	2,12E-01	1,387
Hereditary Breast Cancer Signaling	3,19E00	1,64E-01	#NUM!
Antioxidant Action of Vitamin C	3,05E00	1,74E-01	-0,277
Epithelial Adherens Junction Signaling	3,05E00	1,58E-01	#NUM!
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	3,05E00	1,49E-01	1,732
NRF2-mediated Oxidative Stress Response	3,01E00	1,48E-01	-0,535
Sumoylation Pathway	2,94E00	1,75E-01	-0,535

Role of BRCA1 in DNA Damage Response	2,85E00	1,88E-01	0,707
Superpathway of Cholesterol Biosynthesis	2,84E00	2,76E-01	-2,828
B Cell Receptor Signaling	2,83E00	1,46E-01	0,577
NGF Signaling	2,81E00	1,67E-01	0,471
Molecular Mechanisms of Cancer	2,78E00	1,23E-01	#NUM!
Acute Phase Response Signaling	2,71E00	1,45E-01	-1,342
Superpathway of Geranylgeranyldiphosphate Biosynthesis I (via Mevalonate)	2,7E00	3,33E-01	-2,449
Role of PKR in Interferon Induction and Antiviral Response	2,68E00	1,62E-01	1,147
GNRH Signaling	2,6E00	1,45E-01	1,279
GADD45 Signaling	2,57E00	3,16E-01	#NUM!
ERK/MAPK Signaling	2,56E00	1,4E-01	0,600
PDGF Signaling	2,53E00	1,74E-01	0,258
Mevalonate Pathway I	2,47E00	3,57E-01	-2,236
EGF Signaling	2,46E00	0.2	0,905
RANK Signaling in Osteoclasts	2,43E00	1,7E-01	1,291
Adrenomedullin signaling pathway	2,43E00	1,37E-01	0,200
Endocannabinoid Developing Neuron Pathway	2,39E00	1,57E-01	-0,728
Estrogen Receptor Signaling	2,37E00	1,22E-01	-0,480
Cyclins and Cell Cycle Regulation	2,36E00	1,73E-01	-0,632
Xenobiotic Metabolism General Signaling Pathway	2,36E00	1,47E-01	0,894
Role of CHK Proteins in Cell Cycle Checkpoint Control	2,33E00	1,93E-01	1,000
CDK5 Signaling	2,31E00	1,57E-01	0,243
Insulin Secretion Signaling Pathway	2,24E00	1,28E-01	0,180
Cell Cycle: G1/S Checkpoint Regulation	2,23E00	1,79E-01	0,000
SAPK/JNK Signaling	2,2E00	1,57E-01	0,775

Melanocyte Development and Pigmentation Signaling	2,16E00	1,6E-01	-0,535
ErbB Signaling	2,16E00	1,6E-01	1,291
Chronic Myeloid Leukemia Signaling	2,15E00	1,55E-01	#NUM!
Thrombin Signaling	2,12E00	1,3E-01	1,279
Role of IL-17A in Arthritis	2,04E00	1,85E-01	#NUM!
FLT3 Signaling in Hematopoietic Progenitor Cells	2,01E00	1,63E-01	0,277
Thrombopoietin Signaling	2	1,75E-01	0,302
ERK5 Signaling	1,97E00	1,67E-01	-1,155
RAR Activation	1,96E00	1,29E-01	#NUM!
Prolactin Signaling	1,96E00	1,6E-01	0,905
4-1BB Signaling in T Lymphocytes	1,95E00	2,19E-01	-0,378
14-3-3-mediated Signaling	1,94E00	1,42E-01	1,155
Renin-Angiotensin Signaling	1,93E00	1,44E-01	0,500
3-phosphoinositide Degradation	1,93E00	1,35E-01	-0,218
Unfolded protein response	1,93E00	1,79E-01	-2,121
D-myo-inositol-5-phosphate Metabolism	1,9E00	1,34E-01	0,218
CD40 Signaling	1,9E00	1,69E-01	0,302
Kinetochore Metaphase Signaling Pathway	1,88E00	1,49E-01	0,535
IL-15 Signaling	1,84E00	1,6E-01	-1,155
Sirtuin Signaling Pathway	1,83E00	1,17E-01	-0,200
Inhibition of Angiogenesis by TSP1	1,81E00	2,06E-01	0,000
NER Pathway	1,81E00	1,46E-01	1,069
BMP signaling pathway	1,79E00	1,53E-01	-0,277
Glucocorticoid Receptor Signaling	1,77E00	1,13E-01	#NUM!
Remodeling of Epithelial Adherens Junctions	1,76E00	1,62E-01	#NUM!
UVC-Induced MAPK Signaling	1,75E00	1,76E-01	-1,000
Endocannabinoid Cancer Inhibition Pathway	1,74E00	1,33E-01	-0,229
DNA Methylation and Transcriptional Repression Signaling	1,74E00	0.2	#NUM!
Nucleotide Excision Repair Pathway	1,74E00	0.2	#NUM!
BAG2 Signaling Pathway	1,74E00	1,86E-01	1,414
Neuregulin Signaling	1,72E00	1,46E-01	0,535
IL-7 Signaling Pathway	1,71E00	1,54E-01	1,265

DNA Double-Strand Break Repair by Homologous Recombination	1,68E00	2,86E-01	#NUM!
Huntington's Disease Signaling	1,66E00	1,18E-01	0,577
Fc Epsilon RI Signaling	1,65E00	1,37E-01	1,000
Role of Tissue Factor in Cancer	1,65E00	1,37E-01	#NUM!
Acute Myeloid Leukemia Signaling	1,64E00	1,46E-01	-0,905
Paxillin Signaling	1,63E00	1,39E-01	0,832
IL-17 Signaling	1,63E00	1,5E-01	#NUM!
Small Cell Lung Cancer Signaling	1,63E00	1,55E-01	-0,333
Synaptogenesis Signaling Pathway	1,63E00	1,12E-01	0,522
Regulation of eIF4 and p70S6K Signaling	1,63E00	1,27E-01	-0,333
Sonic Hedgehog Signaling	1,62E00	2,07E-01	-0,447
VEGF Signaling	1,62E00	1,41E-01	0,277
Cell Cycle Regulation by BTG Family Proteins	1,61E00	1,89E-01	0,000
Protein Kinase A Signaling	1,61E00	1,08E-01	1,915
Endothelin-1 Signaling	1,61E00	1,22E-01	0,209
Pancreatic Adenocarcinoma Signaling	1,6E00	1,38E-01	1,155
Cardiac Hypertrophy Signaling	1,59E00	1,17E-01	2,353
Superpathway of Inositol Phosphate Compounds	1,59E00	1,21E-01	0,408
LPS-stimulated MAPK Signaling	1,56E00	1,46E-01	0,000
ILK Signaling	1,56E00	1,21E-01	0,000
Semaphorin Neuronal Repulsive Signaling Pathway	1,55E00	1,31E-01	0,243
IL-17A Signaling in Airway Cells	1,55E00	1,56E-01	0,000
HGF Signaling	1,54E00	1,35E-01	2,324
Heme Biosynthesis II	1,54E00	3,33E-01	#NUM!
Regulation of Cellular Mechanics by Calpain Protease	1,51E00	1,54E-01	0,000
Integrin Signaling	1,5E00	1,17E-01	2,294
April Mediated Signaling	1,5E00	1,79E-01	-0,378
D-myo-inositol (1,4,5,6)-Tetrakisphosphate Biosynthesis	1,49E00	1,27E-01	-0,471

D-myo-inositol (3,4,5,6)-tetrakisphosphate Biosynthesis	1,49E00	1,27E-01	-0,471
Role of NFAT in Cardiac Hypertrophy	1,48E00	1,17E-01	1,706
Extrinsic Prothrombin Activation Pathway	1,48E00	2,5E-01	#NUM!
FcγRIIB Signaling in B Lymphocytes	1,47E00	1,47E-01	0,000
Spliceosomal Cycle	1,47E00	1,67E-01	-2,121
HIF1α Signaling	1,45E00	1,17E-01	1,460
Erythropoietin Signaling	1,44E00	1,45E-01	#NUM!
GDNF Family Ligand-Receptor Interactions	1,44E00	1,45E-01	0,302
Neurotrophin/TRK Signaling	1,44E00	1,45E-01	0,302
FAK Signaling	1,44E00	1,37E-01	#NUM!
ErbB4 Signaling	1,43E00	1,49E-01	1,000
IL-6 Signaling	1,42E00	1,28E-01	-0,500
TGF-β Signaling	1,4E00	1,35E-01	-0,302
Antiproliferative Role of Somatostatin Receptor 2	1,4E00	1,43E-01	0,378
B Cell Activating Factor Signaling	1,39E00	1,71E-01	-0,378
PTEN Signaling	1,39E00	1,27E-01	0,500
Assembly of RNA Polymerase II Complex	1,38E00	1,6E-01	#NUM!
Amyloid Processing	1,38E00	1,6E-01	1,633
Telomerase Signaling	1,37E00	1,31E-01	-0,577
IL-17A Signaling in Gastric Cells	1,36E00	0.2	0,000
Ceramide Signaling	1,35E00	1,36E-01	-0,905
Pyrimidine Ribonucleotides Interconversion	1,34E00	1,67E-01	-0,378
p53 Signaling	1,34E00	1,33E-01	-0,378
IL-8 Signaling	1,34E00	1,15E-01	1,091
Role of JAK2 in Hormone-like Cytokine Signaling	1,31E00	1,76E-01	#NUM!
Reelin Signaling in Neurons	1,31E00	1,24E-01	1,500
White Adipose Tissue Browning Pathway	1,31E00	1,24E-01	-0,500
Induction of Apoptosis by HIV1	1,3E00	1,48E-01	-1,414
Role of MAPK Signaling in the Pathogenesis of Influenza	1,3E00	1,38E-01	#NUM!
Estrogen-mediated S-phase Entry	1,3E00	1,92E-01	-1,342
Glioma Signaling	1,28E00	1,27E-01	1,000
Tetrapyrrole Biosynthesis II	1,27E00	0.4	#NUM!
Galactose Degradation I (Leloir Pathway)	1,27E00	0.4	#NUM!
Cardiac β-adrenergic Signaling	1,27E00	1,21E-01	0,577
TWEAK Signaling	1,26E00	1,71E-01	-1,633
IL-1 Signaling	1,26E00	1,32E-01	-0,707
Type I Diabetes Mellitus Signaling	1,25E00	1,26E-01	-1,069

Lymphotoxin β Receptor Signaling	1,25E00	1,51E-01	-0,378
Pyrimidine Ribonucleotides De Novo Biosynthesis	1,25E00	1,59E-01	-0,378
Type II Diabetes Mellitus Signaling	1,24E00	1,2E-01	-0,277
IL-12 Signaling and Production in Macrophages	1,24E00	1,21E-01	#NUM!
G Beta Gamma Signaling	1,23E00	1,23E-01	1,069
Inhibition of ARE-Mediated mRNA Degradation Pathway	1,23E00	1,23E-01	-0,258
Non-Small Cell Lung Cancer Signaling	1,21E00	1,37E-01	0,707
iNOS Signaling	1,21E00	1,56E-01	-0,816
Mouse Embryonic Stem Cell Pluripotency	1,2E00	1,26E-01	-0,277
Assembly of RNA Polymerase I Complex	1,2E00	2,5E-01	#NUM!
Glioblastoma Multiforme Signaling	1,19E00	1,15E-01	1,500
PKCθ Signaling in T Lymphocytes	1,18E00	1,16E-01	1,069
Corticotropin Releasing Hormone Signaling	1,17E00	1,17E-01	-0,258
3-phosphoinositide Biosynthesis	1,17E00	1,14E-01	-0,229
FGF Signaling	1,17E00	1,31E-01	1,508
Regulation Of The Epithelial Mesenchymal Transition In Development Pathway	1,17E00	1,31E-01	0,000
Gap Junction Signaling	1,17E00	1,11E-01	#NUM!
STAT3 Pathway	1,16E00	1,19E-01	0,500
PFKFB4 Signaling Pathway	1,16E00	1,52E-01	0,378
HIPPO signaling	1,14E00	1,29E-01	0,816
Necroptosis Signaling Pathway	1,14E00	1,15E-01	-1,886
Xenobiotic Metabolism Signaling	1,14E00	1,05E-01	#NUM!
Zymosterol Biosynthesis	1,12E00	3,33E-01	#NUM!
Toll-like Receptor Signaling	1,12E00	1,32E-01	0,707
P2Y Purigenic Receptor Signaling Pathway	1,11E00	1,18E-01	0,000
Cholesterol Biosynthesis I	1,11E00	2,31E-01	#NUM!
NAD Phosphorylation and Dephosphorylation	1,11E00	2,31E-01	#NUM!
Cholesterol Biosynthesis II (via 24,25-dihydrolanosterol)	1,11E00	2,31E-01	#NUM!
Cholesterol Biosynthesis III (via Desmosterol)	1,11E00	2,31E-01	#NUM!
Methylthiopropionate Biosynthesis	1,1E00	1	#NUM!
Lipoate Salvage and Modification	1,1E00	1	#NUM!
Sulfite Oxidation IV	1,1E00	1	#NUM!

Hepatic Fibrosis Signaling Pathway	1,1E00	1,01E-01	0,667
tRNA Charging	1,07E00	1,54E-01	-2,449
Cancer Drug Resistance By Drug Efflux	1,07E00	1,38E-01	#NUM!
Agrin Interactions at Neuromuscular Junction	1,06E00	1,28E-01	-0,707
Insulin Receptor Signaling	1,05E00	1,14E-01	-1,069
Cholecystokinin/Gastrin-mediated Signaling	1,05E00	1,18E-01	1,069
Gα12/13 Signaling	1,04E00	1,15E-01	-0,775
Pyrimidine Deoxyribonucleotides De Novo Biosynthesis I	1,04E00	1,82E-01	0,000
Apoptosis Signaling	1,04E00	1,21E-01	0,000
Regulation of IL-2 Expression in Activated and Anergic T Lymphocytes	1,03E00	1,24E-01	#NUM!
Phenylalanine Degradation IV (Mammalian, via Side Chain)	1,03E00	2,14E-01	#NUM!
Colanic Acid Building Blocks Biosynthesis	1,03E00	2,14E-01	#NUM!
Dopamine-DARPP32 Feedback in cAMP Signaling	1,02E00	1,1E-01	0,243
Myc Mediated Apoptosis Signaling	1,01E00	1,4E-01	-0,378
TNFR1 Signaling	1,01E00	1,4E-01	-1,134
Semaphorin Signaling in Neurons	1	1,33E-01	#NUM!
JAK/Stat Signaling	1	1,25E-01	-1,265
Chemokine Signaling	1	1,25E-01	0,000
Natural Killer Cell Signaling	9,96E-01	1,07E-01	1,528
Thioredoxin Pathway	9,96E-01	2,86E-01	#NUM!
mTOR Signaling	9,63E-01	1,05E-01	-0,728
Regulation Of The Epithelial Mesenchymal Transition By Growth Factors Pathway	9,63E-01	1,06E-01	-0,688
Osteoarthritis Pathway	9,47E-01	1,04E-01	0,853
Melatonin Signaling	9,39E-01	1,25E-01	1,414
UVB-Induced MAPK Signaling	9,39E-01	1,35E-01	0,378
PPARα/RXRα Activation	9,28E-01	1,05E-01	1,500
IL-22 Signaling	9,28E-01	1,67E-01	0,000
Tumoricidal Function of Hepatic Natural Killer Cells	9,28E-01	1,67E-01	-1,000
Role of JAK1, JAK2 and TYK2 in Interferon Signaling	9,28E-01	1,67E-01	#NUM!
NF-κB Signaling	9,28E-01	1,06E-01	0,229
Phospholipase C Signaling	9,24E-01	1,01E-01	1,706
Oncostatin M Signaling	9,14E-01	1,4E-01	-0,816

Aldosterone Signaling in Epithelial Cells	9,14E-01	1,08E-01	2,236
tRNA Splicing	9,14E-01	1,4E-01	0,000
Ephrin Receptor Signaling	9,1E-01	1,06E-01	0,500
CCR5 Signaling in Macrophages	9,07E-01	1,17E-01	0,816
Opioid Signaling Pathway	9,07E-01	1,01E-01	-0,408
Regulation of the Epithelial-Mesenchymal Transition Pathway	8,96E-01	1,04E-01	#NUM!
Granzyme B Signaling	8,93E-01	1,88E-01	#NUM!
Mismatch Repair in Eukaryotes	8,93E-01	1,88E-01	#NUM!
Isoleucine Degradation I	8,93E-01	1,88E-01	#NUM!
Parkinson's Signaling	8,93E-01	1,88E-01	#NUM!
Hypoxia Signaling in the Cardiovascular System	8,86E-01	1,22E-01	#NUM!
PI3K Signaling in B Lymphocytes	8,79E-01	1,09E-01	-0,277
Role of JAK family kinases in IL-6-type Cytokine Signaling	8,79E-01	1,6E-01	#NUM!
Apelin Pancreas Signaling Pathway	8,79E-01	1,36E-01	0,816
Ovarian Cancer Signaling	8,63E-01	1,08E-01	-1,000
Endocannabinoid Neuronal Synapse Pathway	8,63E-01	1,09E-01	1,604
Relaxin Signaling	8,6E-01	1,07E-01	0,378
Coagulation System	8,54E-01	1,43E-01	-2,236
Virus Entry via Endocytic Pathways	8,54E-01	1,12E-01	#NUM!
p38 MAPK Signaling	8,48E-01	1,1E-01	-0,577
p70S6K Signaling	8,45E-01	1,09E-01	0,277
NAD Salvage Pathway II	8,36E-01	1,54E-01	0,000
Mitotic Roles of Polo-Like Kinase	8,24E-01	1,21E-01	-0,447
Asparagine Degradation I	8,18E-01	0.5	#NUM!
Lipoate Biosynthesis and Incorporation II	8,18E-01	0.5	#NUM!
All-trans-decaprenyl Diphosphate Biosynthesis	8,18E-01	0.5	#NUM!
Alanine Degradation III	8,18E-01	0.5	#NUM!
Alanine Biosynthesis II	8,18E-01	0.5	#NUM!
L-cysteine Degradation III	8,18E-01	0.5	#NUM!
Spermidine Biosynthesis I	8,18E-01	0.5	#NUM!
Putrescine Biosynthesis III	8,18E-01	0.5	#NUM!
UVA-Induced MAPK Signaling	8,18E-01	1,12E-01	0,378
Cell Cycle Control of Chromosomal Replication	8,15E-01	1,25E-01	-0,378
Interferon Signaling	8,15E-01	1,39E-01	0,447
PI3K/AKT Signaling	8,12E-01	1,03E-01	-1,807
Dopamine Receptor Signaling	8,1E-01	1,17E-01	#NUM!
Leucine Degradation I	8,07E-01	2,22E-01	#NUM!
Folate Transformations I	8,07E-01	2,22E-01	#NUM!
Aryl Hydrocarbon Receptor Signaling	7,9E-01	1,05E-01	0,302

IL-15 Production	7,9E-01	1,07E-01	0,832
Notch Signaling	7,8E-01	1,35E-01	0,000
1D-myo-inositol Hexakisphosphate Biosynthesis II (Mammalian)	7,77E-01	1,67E-01	#NUM!
D-myo-inositol (1,3,4)-trisphosphate Biosynthesis	7,77E-01	1,67E-01	#NUM!
Synaptic Long Term Depression	7,72E-01	1,01E-01	0,728
Thyroid Cancer Signaling	7,64E-01	1,14E-01	-1,000
EIF2 Signaling	7,59E-01	9,82E-02	1,604
RhoA Signaling	7,54E-01	1,06E-01	1,155
Adipogenesis pathway	7,54E-01	1,04E-01	#NUM!
Docosahexaenoic Acid (DHA) Signaling	7,47E-01	1,32E-01	#NUM!
Prostate Cancer Signaling	7,38E-01	1,1E-01	#NUM!
Calcium Transport I	7,33E-01	0.2	#NUM!
DNA damage-induced 14-3-3σ Signaling	7,28E-01	1,58E-01	#NUM!
Gαq Signaling	7,28E-01	1,01E-01	1,604
Androgen Signaling	7,21E-01	1,03E-01	-0,378
PEDF Signaling	6,99E-01	1,1E-01	-1,000
Cardiac Hypertrophy Signaling (Enhanced)	6,97E-01	9,03E-02	1,677
Actin Cytoskeleton Signaling	6,88E-01	9,63E-02	0,471
Colorectal Cancer Metastasis Signaling	6,86E-01	9,49E-02	0,218
Cardiomyocyte Differentiation via BMP Receptors	6,82E-01	1,5E-01	#NUM!
CREB Signaling in Neurons	6,82E-01	9,66E-02	0,577
CDP-diacylglycerol Biosynthesis I	6,82E-01	1,5E-01	#NUM!
Coronavirus Pathogenesis Pathway	6,8E-01	0.1	0,000
Ketogenesis	6,68E-01	1,82E-01	#NUM!
Leukocyte Extravasation Signaling	6,62E-01	9,64E-02	1,500
Tetrahydrobiopterin Biosynthesis I	6,6E-01	3,33E-01	#NUM!
Hypusine Biosynthesis	6,6E-01	3,33E-01	#NUM!
Geranylgeranyldiphosphate Biosynthesis	6,6E-01	3,33E-01	#NUM!
Thiosulfate Disproportionation III (Rhodanese)	6,6E-01	3,33E-01	#NUM!
Oxidized GTP and dGTP Detoxification	6,6E-01	3,33E-01	#NUM!
Tetrahydrobiopterin Biosynthesis II	6,6E-01	3,33E-01	#NUM!
Glutamate Degradation II	6,6E-01	3,33E-01	#NUM!
Tyrosine Biosynthesis IV	6,6E-01	3,33E-01	#NUM!
Aspartate Biosynthesis	6,6E-01	3,33E-01	#NUM!
Caveolar-mediated Endocytosis Signaling	6,58E-01	1,1E-01	#NUM!
VEGF Family Ligand-Receptor Interactions	6,58E-01	1,07E-01	1,000
G Protein Signaling Mediated by Tubby	6,42E-01	1,29E-01	#NUM!
α-Adrenergic Signaling	6,4E-01	1,04E-01	0,447
Endoplasmic Reticulum Stress Pathway	6,4E-01	1,43E-01	#NUM!
Activation of IRF by Cytosolic Pattern Recognition Receptors	6,38E-01	1,11E-01	0,378
IL-4 Signaling	6,38E-01	1,06E-01	#NUM!
Leptin Signaling in Obesity	6,36E-01	1,08E-01	#NUM!
Role of IL-17F in Allergic Inflammatory Airway Diseases	6,25E-01	1,19E-01	-0,447
Retinol Biosynthesis	6,25E-01	1,19E-01	-0,447
Bladder Cancer Signaling	6,23E-01	1,03E-01	-0,447
PAK Signaling	6,23E-01	1,03E-01	0,632

Cellular Effects of Sildenafil (Viagra)	6,22E-01	9,92E-02	#NUM!
Phospholipases	6,16E-01	1,09E-01	1,134
CXCR4 Signaling	6,02E-01	9,58E-02	0,535
Polyamine Regulation in Colon Cancer	0.6	1,36E-01	#NUM!
Phosphatidylglycerol Biosynthesis II (Non-plastidic)	0.6	1,36E-01	#NUM!
Transcriptional Regulatory Network in Embryonic Stem Cells	5,93E-01	1,11E-01	#NUM!
ErbB2-ErbB3 Signaling	5,93E-01	1,08E-01	-0,378
Tight Junction Signaling	5,88E-01	9,52E-02	#NUM!
Circadian Rhythm Signaling	5,78E-01	1,21E-01	#NUM!
Role of RIG1-like Receptors in Antiviral Innate Immunity	5,72E-01	1,14E-01	-0,447
Clathrin-mediated Endocytosis Signaling	5,69E-01	9,33E-02	#NUM!
CTLA4 Signaling in Cytotoxic T Lymphocytes	5,65E-01	1,01E-01	#NUM!
Superpathway of D-myo-inositol (1,4,5)-trisphosphate Metabolism	5,64E-01	1,3E-01	#NUM!
Assembly of RNA Polymerase III Complex	5,59E-01	1,54E-01	#NUM!
Serotonin Degradation	5,53E-01	1,04E-01	-0,378
Arsenate Detoxification I (Glutaredoxin)	5,51E-01	2,5E-01	#NUM!
Retinoate Biosynthesis II	5,51E-01	2,5E-01	#NUM!
Heme Biosynthesis from Uroporphyrinogen-III I	5,51E-01	2,5E-01	#NUM!
Spermine and Spermidine Degradation I	5,51E-01	2,5E-01	#NUM!
Phenylethylamine Degradation I	5,51E-01	2,5E-01	#NUM!
Glutathione Redox Reactions II	5,51E-01	2,5E-01	#NUM!
Melatonin Degradation II	5,51E-01	2,5E-01	#NUM!
L-cysteine Degradation I	5,51E-01	2,5E-01	#NUM!
Phenylalanine Degradation I (Aerobic)	5,51E-01	2,5E-01	#NUM!
Fatty Acid β-oxidation III (Unsaturated, Odd Number)	5,51E-01	2,5E-01	#NUM!
IL-3 Signaling	5,41E-01	1,01E-01	0,000
Signaling by Rho Family GTPases	5,3E-01	9,02E-02	2,524
CNTF Signaling	5,29E-01	1,05E-01	0,000
Vitamin-C Transport	5,29E-01	1,25E-01	#NUM!
FAT10 Cancer Signaling Pathway	5,24E-01	1,09E-01	-0,447
Noradrenaline and Adrenaline Degradation	5,23E-01	1,14E-01	-1,000
Factors Promoting Cardiogenesis in Vertebrates	5,21E-01	9,33E-02	1,604
Apelin Endothelial Signaling Pathway	5,19E-01	9,57E-02	-0,302
IL-10 Signaling	5,16E-01	1,01E-01	#NUM!
PPAR Signaling	5,1E-01	9,62E-02	1,897
IGF-1 Signaling	5,1E-01	9,62E-02	-0,333
T Cell Exhaustion Signaling Pathway	5,04E-01	9,14E-02	0,000
nNOS Signaling in Neurons	5,02E-01	1,06E-01	#NUM!
D-myo-inositol (1,4,5)-Trisphosphate Biosynthesis	4,98E-01	1,2E-01	#NUM!
T Cell Receptor Signaling	4,95E-01	9,52E-02	#NUM!
Maturity Onset Diabetes of Young (MODY) Signaling	4,88E-01	1,02E-01	#NUM!
Fcγ Receptor-mediated Phagocytosis in Macrophages and Monocytes	4,84E-01	9,57E-02	0,333
HMGB1 Signaling	4,84E-01	9,09E-02	0,000

Choline Biosynthesis III	4,74E-01	1,33E-01	#NUM!
Superpathway of Methionine Degradation	4,72E-01	1,08E-01	-1,000
Complement System	4,72E-01	1,08E-01	#NUM!
2-ketoglutarate Dehydrogenase Complex	4,71E-01	0.2	#NUM!
Tetrahydrofolate Salvage from 5,10-methenyltetrahydrofolate	4,71E-01	0.2	#NUM!
Serine Biosynthesis	4,71E-01	0.2	#NUM!
Lysine Degradation II	4,71E-01	0.2	#NUM!
Myo-inositol Biosynthesis	4,71E-01	0.2	#NUM!
Lysine Degradation V	4,71E-01	0.2	#NUM!
Tyrosine Degradation I	4,71E-01	0.2	#NUM!
Endometrial Cancer Signaling	4,7E-01	0.1	0,000
Apelin Liver Signaling Pathway	4,69E-01	1,15E-01	#NUM!
Systemic Lupus Erythematosus In B Cell Signaling Pathway	4,67E-01	8,73E-02	-0,408
Actin Nucleation by ARP-WASP Complex	4,63E-01	9,72E-02	-0,447
Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis	4,62E-01	8,65E-02	#NUM!
Cell Cycle: G2/M DNA Damage Checkpoint Regulation	4,6E-01	1,02E-01	1,000
CD28 Signaling in T Helper Cells	4,52E-01	9,17E-02	0,302
IL-2 Signaling	4,51E-01	9,84E-02	0,000
GPCR-Mediated Integration of Enteroendocrine Signaling Exemplified by an L Cell	4,46E-01	9,59E-02	-0,378
Melanoma Signaling	4,4E-01	0.1	-0,447
Chondroitin Sulfate Degradation (Metazoa)	4,37E-01	1,25E-01	#NUM!
Wnt/Ca+ pathway	4,34E-01	9,68E-02	0,816
Calcium Signaling	4,34E-01	8,74E-02	0,302
Estrogen-Dependent Breast Cancer Signaling	4,31E-01	9,46E-02	-1,134
Apelin Cardiomyocyte Signaling Pathway	4,15E-01	9,09E-02	1,667
NAD Biosynthesis from 2-amino-3-carboxymuconate Semialdehyde	4,08E-01	1,67E-01	#NUM!
Selenocysteine Biosynthesis II (Archaea and Eukaryotes)	4,08E-01	1,67E-01	#NUM!
NAD Biosynthesis III	4,08E-01	1,67E-01	#NUM!
Lactose Degradation III	4,08E-01	1,67E-01	#NUM!
GDP-mannose Biosynthesis	4,08E-01	1,67E-01	#NUM!
CCR3 Signaling in Eosinophils	4,05E-01	8,87E-02	0,707
RAN Signaling	4,03E-01	1,18E-01	#NUM!
Rac Signaling	4,03E-01	8,93E-02	-0,632
Dermatan Sulfate Degradation (Metazoa)	4,03E-01	1,18E-01	#NUM!
D-myo-inositol (1,4,5)-trisphosphate Degradation	4,03E-01	1,18E-01	#NUM!
Macropinocytosis Signaling	4,01E-01	9,21E-02	0,000
Role of p14/p19ARF in Tumor Suppression	3,9E-01	1,03E-01	#NUM!
Mechanisms of Viral Exit from Host Cells	3,86E-01	9,76E-02	#NUM!
nNOS Signaling in Skeletal Muscle Cells	3,86E-01	9,76E-02	#NUM!
FAT10 Signaling Pathway	3,74E-01	1,11E-01	#NUM!
MIF Regulation of Innate Immunity	3,68E-01	9,52E-02	0,000
TNFR2 Signaling	3,68E-01	0.1	#NUM!
Dopamine Degradation	3,68E-01	0.1	#NUM!

Death Receptor Signaling	3,63E-01	8,79E-02	-1,134
BEX2 Signaling Pathway	3,6E-01	8,86E-02	-1,134
Trehalose Degradation II (Trehalase)	3,58E-01	1,43E-01	#NUM!
Superpathway of Serine and Glycine Biosynthesis I	3,58E-01	1,43E-01	#NUM!
NAD Salvage Pathway III	3,58E-01	1,43E-01	#NUM!
Glycogen Biosynthesis II (from UDP-D-Glucose)	3,58E-01	1,43E-01	#NUM!
Glycoaminoglycan-protein Linkage Region Biosynthesis	3,58E-01	1,43E-01	#NUM!
Aspartate Degradation II	3,58E-01	1,43E-01	#NUM!
Synaptic Long Term Potentiation	3,51E-01	8,53E-02	0,302
Gustation Pathway	3,49E-01	8,44E-02	#NUM!
Renal Cell Carcinoma Signaling	3,47E-01	8,75E-02	0,447
Valine Degradation I	3,46E-01	1,05E-01	#NUM!
Methylglyoxal Degradation III	3,46E-01	1,05E-01	#NUM!
Apelin Muscle Signaling Pathway	3,46E-01	1,05E-01	#NUM!
Nur77 Signaling in T Lymphocytes	3,42E-01	8,82E-02	1,633
Role of JAK1 and JAK3 in γc Cytokine Signaling	3,29E-01	8,7E-02	#NUM!
Ethanol Degradation II	3,27E-01	9,38E-02	#NUM!
NF-κB Activation by Viruses	3,21E-01	8,54E-02	-1,134
Apelin Adipocyte Signaling Pathway	3,21E-01	8,54E-02	-0,378
The Visual Cycle	3,21E-01	0.1	#NUM!
Inflammasome pathway	3,21E-01	0.1	#NUM!
Histidine Degradation III	3,16E-01	1,25E-01	#NUM!
Superoxide Radicals Degradation	3,16E-01	1,25E-01	#NUM!
Salvage Pathways of Pyrimidine Deoxyribonucleotides	3,16E-01	1,25E-01	#NUM!
GABA Receptor Signaling	3,16E-01	8,42E-02	#NUM!
GM-CSF Signaling	3,16E-01	8,57E-02	0,000
Dendritic Cell Maturation	3,14E-01	8,2E-02	-0,535
IL-9 Signaling	3,08E-01	9,09E-02	#NUM!
LXR/RXR Activation	3,07E-01	8,26E-02	-0,707
eNOS Signaling	3,05E-01	8,18E-02	0,632
Role of Oct4 in Mammalian Embryonic Stem Cell Pluripotency	3,01E-01	8,7E-02	#NUM!
Sperm Motility	2,98E-01	8,07E-02	0,000
Putrescine Degradation III	2,98E-01	9,52E-02	#NUM!
HOTAIR Regulatory Pathway	2,97E-01	8,13E-02	0,277
Wnt/β-catenin Signaling	2,94E-01	8,09E-02	-0,302
Cytotoxic T Lymphocyte-mediated Apoptosis of Target Cells	2,91E-01	8,82E-02	#NUM!
MIF-mediated Glucocorticoid Regulation	2,91E-01	8,82E-02	#NUM!
Retinoate Biosynthesis I	2,91E-01	8,82E-02	#NUM!
Ephrin A Signaling	2,87E-01	8,51E-02	#NUM!
Triacylglycerol Degradation	2,87E-01	8,51E-02	-1,000
Sucrose Degradation V (Mammalian)	2,81E-01	1,11E-01	#NUM!
Phosphatidylethanolamine Biosynthesis II	2,81E-01	1,11E-01	#NUM!
Methionine Degradation I (to Homocysteine)	2,76E-01	9,09E-02	#NUM!
Nitric Oxide Signaling in the Cardiovascular System	2,75E-01	8,08E-02	-1,134
Gαi Signaling	2,72E-01	0.08	-0,707

Xenobiotic Metabolism CAR Signaling Pathway	2,7E-01	7,94E-02	0,775
Tec Kinase Signaling	2,66E-01	7,93E-02	1,265
FXR/RXR Activation	2,63E-01	7,94E-02	#NUM!
Angiopoietin Signaling	2,58E-01	0.08	-0,447
Apelin Cardiac Fibroblast Signaling Pathway	2,56E-01	8,7E-02	#NUM!
Ketolysis	2,5E-01	0.1	#NUM!
Dolichyl-diphosphooligosaccharide Biosynthesis	2,5E-01	0.1	#NUM!
Glycine Betaine Degradation	2,5E-01	0.1	#NUM!
Glutathione Redox Reactions I	2,38E-01	8,33E-02	#NUM!
TCA Cycle II (Eukaryotic)	2,38E-01	8,33E-02	#NUM!
Cysteine Biosynthesis III (mammalia)	2,38E-01	8,33E-02	#NUM!
γ-glutamyl Cycle	2,24E-01	9,09E-02	#NUM!
GDP-glucose Biosynthesis	2,24E-01	9,09E-02	#NUM!
Tryptophan Degradation X (Mammalian, via Tryptamine)	2,22E-01	0.08	#NUM!
Cleavage and Polyadenylation of Pre-mRNA	2,01E-01	8,33E-02	#NUM!
Glucose and Glucose-1-phosphate Degradation	2,01E-01	8,33E-02	#NUM!
NAD biosynthesis II (from tryptophan)	2,01E-01	8,33E-02	#NUM!
Guanosine Nucleotides Degradation III	2,01E-01	8,33E-02	#NUM!
Mineralocorticoid Biosynthesis	2,01E-01	8,33E-02	#NUM!

Appendix table 17. Significantly enriched human pathways between the uninfected controls and infected samples at 6 h post infection.

Ingenuity Canonical Pathways	**-log(p-value)**	**Ratio**	**z-score**
Role of Tissue Factor in Cancer	4,13E00	8,55E-02	#NUM!
Acute Phase Response Signaling	3,81E00	6,7E-02	-1,414
Extrinsic Prothrombin Activation Pathway	3,73E00	2,5E-01	#NUM!
Signalling by Rho Family GTPases	3,63E00	5,74E-02	-0,577
14-3-3-mediated Signaling	3,16E00	7,09E-02	0,816
Sumoylation Pathway	3,13E00	7,77E-02	0,000
Gαq Signaling	3,07E00	6,33E-02	0,333
Antioxidant Action of Vitamin C	2,97E00	7,34E-02	0,447
IL-17A Signaling in Gastric Cells	2,94E00	1,6E-01	#NUM!
NGF Signaling	2,84E00	7,02E-02	-1,134
IL-8 Signaling	2,82E00	5,5E-02	-0,333
PFKFB4 Signaling Pathway	2,77E00	1,09E-01	0,447
GP6 Signaling Pathway	2,73E00	6,72E-02	-0,707
IL-10 Signaling	2,72E00	8,7E-02	#NUM!
Thrombin Signaling	2,69E00	5,29E-02	0,333
Heme Degradation	2,68E00	0.5	#NUM!
Growth Hormone Signaling	2,66E00	8,45E-02	0,000
IL-6 Signaling	2,59E00	6,4E-02	0,000
Apoptosis Signaling	2,57E00	7,07E-02	-0,378
Phospholipase C Signaling	2,44E00	4,67E-02	1,667

Coagulation System	2,39E00	1,14E-01	-2,000
Hepatic Fibrosis Signaling Pathway	2,35E00	4,08E-02	-0,258
Germ Cell-Sertoli Cell Junction Signaling	2,28E00	5,26E-02	#NUM!
Semaphorin Signaling in Neurons	2,27E00	8,33E-02	#NUM!
GNRH Signaling	2,25E00	5,2E-02	0,333
Cardiac Hypertrophy Signaling	2,22E00	4,58E-02	0,000
April Mediated Signaling	2,22E00	1,03E-01	-1,000
Wnt/Ca+ pathway	2,21E00	8,06E-02	1,342
Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis	2,18E00	4,17E-02	#NUM!
Insulin Secretion Signaling Pathway	2,18E00	4,53E-02	1,508
Renin-Angiotensin Signaling	2,15E00	5,93E-02	-0,378
p38 MAPK Signaling	2,15E00	5,93E-02	0,378
B Cell Activating Factor Signaling	2,14E00	9,76E-02	-1,000
Cholecystokinin/Gastrin-mediated Signaling	2,13E00	5,88E-02	0,000
Opioid Signaling Pathway	2,13E00	4,45E-02	0,000
Death Receptor Signaling	2,13E00	6,59E-02	-0,447
Role of IL-17F in Allergic Inflammatory Airway Diseases	2,1E00	9,52E-02	-1,000
Sertoli Cell-Sertoli Cell Junction Signaling	2,07E00	4,86E-02	#NUM!
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	2,02E00	4,79E-02	0,000
Estrogen Receptor Signaling	2,01E00	3,96E-02	0,277
Sperm Motility	2	4,48E-02	0,816
ILK Signaling	2	4,74E-02	-0,707
P2Y Purigenic Receptor Signaling Pathway	1,99E00	5,51E-02	1,134
ERK/MAPK Signaling	1,95E00	4,66E-02	0,707
Cardiac Hypertrophy Signaling (Enhanced)	1,94E00	3,49E-02	1,213
Melatonin Signaling	1,94E00	6,94E-02	0,447
Neuropathic Pain Signaling In Dorsal Horn Neurons	1,91E00	5,94E-02	0,816
GPCR-Mediated Integration of Enteroendocrine Signaling Exemplified by an L Cell	1,91E00	6,85E-02	-0,447
Tec Kinase Signaling	1,9E00	4,88E-02	0,000
Erythropoietin Signaling	1,84E00	6,58E-02	#NUM!
GDNF Family Ligand-Receptor Interactions	1,84E00	6,58E-02	-1,000
Toll-like Receptor Signaling	1,84E00	6,58E-02	#NUM!
Calcium Transport I	1,84E00	0.2	#NUM!
PI3K Signaling in B Lymphocytes	1,8E00	5,07E-02	0,816
Role of BRCA1 in DNA Damage Response	1,76E00	6,25E-02	#NUM!
CD27 Signaling in Lymphocytes	1,75E00	7,55E-02	-1,000
Type II Diabetes Mellitus Signaling	1,74E00	4,93E-02	0,000
Prolactin Signaling	1,74E00	6,17E-02	2,000
TNFR2 Signaling	1,73E00	0.1	#NUM!
Role of IL-17A in Arthritis	1,73E00	7,41E-02	#NUM!
Protein Kinase A Signaling	1,71E00	3,51E-02	1,732
EGF Signaling	1,7E00	7,27E-02	0,000
Integrin Signaling	1,7E00	4,23E-02	0,707

Corticotropin Releasing Hormone Signaling	1,7E00	4,83E-02	0,816
Role of NFAT in Cardiac Hypertrophy	1,69E00	4,21E-02	1,000
Glucose and Glucose-1-phosphate Degradation	1,69E00	1,67E-01	#NUM!
Unfolded protein response	1,68E00	7,14E-02	0,000
Endocannabinoid Developing Neuron Pathway	1,66E00	5,22E-02	0,447
IL-9 Signaling	1,62E00	9,09E-02	#NUM!
Role of IL-17A in Psoriasis	1,62E00	1,54E-01	#NUM!
RANK Signaling in Osteoclasts	1,6E00	5,68E-02	-1,342
Inhibition of Angiogenesis by TSP1	1,59E00	8,82E-02	#NUM!
Role of JAK2 in Hormone-like Cytokine Signaling	1,59E00	8,82E-02	#NUM!
Regulation of IL-2 Expression in Activated and Anergic T Lymphocytes	1,58E00	5,62E-02	#NUM!
DNA Double-Strand Break Repair by Homologous Recombination	1,56E00	1,43E-01	#NUM!
TWEAK Signaling	1,56E00	8,57E-02	#NUM!
Induction of Apoptosis by HIV1	1,55E00	6,56E-02	#NUM!
FXR/RXR Activation	1,49E00	4,76E-02	#NUM!
Atherosclerosis Signaling	1,48E00	4,72E-02	#NUM!
Adrenomedullin 336ignalling pathway	1,47E00	4,06E-02	0,000
Molecular Mechanisms of Cancer	1,46E00	3,32E-02	#NUM!
α-Adrenergic Signaling	1,45E00	5,21E-02	#NUM!
Synaptic Long Term Potentiation	1,45E00	4,65E-02	1,633
Mismatch Repair in Eukaryotes	1,45E00	1,25E-01	#NUM!
Semaphorin Neuronal Repulsive Signaling Pathway	1,44E00	4,62E-02	0,816
HMGB1 Signaling	1,43E00	4,24E-02	-1,633
Palmitate Biosynthesis I (Animals)	1,43E00	0.5	#NUM!
Fatty Acid Biosynthesis Initiation II	1,43E00	0.5	#NUM!
p53 Signaling	1,42E00	5,1E-02	-1,000
UVA-Induced MAPK Signaling	1,42E00	5,1E-02	1,000
IL-12 Signaling and Production in Macrophages	1,41E00	4,55E-02	#NUM!
CXCR4 Signaling	1,41E00	4,19E-02	0,000
Remodeling of Epithelial Adherens Junctions	1,41E00	5,88E-02	#NUM!
HIF1α Signaling	1,39E00	3,9E-02	0,000
Axonal Guidance Signaling	1,38E00	3,1E-02	#NUM!
CREB Signaling in Neurons	1,37E00	3,86E-02	1,134
Intrinsic Prothrombin Activation Pathway	1,35E00	7,14E-02	#NUM!
Heparan Sulfate Biosynthesis (Late Stages)	1,35E00	5,63E-02	0,000
ERK5 Signaling	1,33E00	5,56E-02	0,000
tRNA Splicing	1,33E00	6,98E-02	#NUM!
Osteoarthritis Pathway	1,33E00	3,79E-02	-0,707
T Cell Receptor Signaling	1,31E00	4,76E-02	#NUM!
GADD45 Signaling	1,31E00	1,05E-01	#NUM!
Hereditary Breast Cancer Signaling	1,31E00	4,29E-02	#NUM!
Apelin Pancreas Signaling Pathway	1,3E00	6,82E-02	#NUM!
AMPK Signaling	1,3E00	3,74E-02	-0,816

Estrogen-Dependent Breast Cancer Signaling	1,29E00	5,41E-02	0,000
Leptin Signaling in Obesity	1,29E00	5,41E-02	#NUM!
Virus Entry via Endocytic Pathways	1,29E00	4,67E-02	#NUM!
Endocannabinoid Cancer Inhibition Pathway	1,27E00	4,2E-02	0,000
Xenobiotic Metabolism General Signaling Pathway	1,27E00	4,2E-02	-0,816
RhoGDI Signaling	1,26E00	3,89E-02	-0,447
Hypusine Biosynthesis	1,26E00	3,33E-01	#NUM!
Role of Oct4 in Mammalian Embryonic Stem Cell Pluripotency	1,26E00	6,52E-02	#NUM!
FAT10 Cancer Signaling Pathway	1,26E00	6,52E-02	#NUM!
Role of NFAT in Regulation of the Immune Response	1,25E00	3,87E-02	0,447
HGF Signaling	1,23E00	4,5E-02	-0,447
Heparan Sulfate Biosynthesis	1,22E00	5,13E-02	0,000
Rac Signaling	1,22E00	4,46E-02	-2,236
GPCR-Mediated Nutrient Sensing in Enteroendocrine Cells	1,22E00	4,46E-02	1,342
B Cell Receptor Signaling	1,21E00	3,78E-02	-0,378
BEX2 Signaling Pathway	1,21E00	5,06E-02	-1,000
Hepatic Fibrosis / Hepatic Stellate Cell Activation	1,2E00	3,76E-02	#NUM!
IL-17 Signaling	1,19E00	0.05	#NUM!
Relaxin Signaling	1,19E00	0.04	-1,000
Factors Promoting Cardiogenesis in Vertebrates	1,19E00	0.04	1,633
Endothelin-1 Signaling	1,18E00	3,72E-02	1,890
Myc Mediated Apoptosis Signaling	1,17E00	0.06	#NUM!
TNFR1 Signaling	1,17E00	0.06	#NUM!
fMLP Signaling in Neutrophils	1,16E00	4,31E-02	0,447
PPARα/RXRα Activation	1,16E00	3,68E-02	1,342
LPS-stimulated MAPK Signaling	1,16E00	4,88E-02	0,000
PEDF Signaling	1,16E00	4,88E-02	0,000
Sphingosine-1-phosphate Signaling	1,15E00	4,27E-02	0,000
Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses	1,15E00	3,9E-02	0,000
Arsenate Detoxification I (Glutaredoxin)	1,14E00	2,5E-01	#NUM!
IL-22 Signaling	1,13E00	8,33E-02	#NUM!
Tumoricidal Function of Hepatic Natural Killer Cells	1,13E00	8,33E-02	#NUM!
Vitamin-C Transport	1,13E00	8,33E-02	#NUM!
RAR Activation	1,13E00	3,61E-02	#NUM!
Lymphotoxin β Receptor Signaling	1,11E00	5,66E-02	#NUM!
IL-15 Production	1,1E00	4,13E-02	-1,342
Aldosterone Signaling in Epithelial Cells	1,1E00	3,8E-02	2,000
PDGF Signaling	1,1E00	4,65E-02	0,000
Natural Killer Cell Signaling	1,1E00	3,55E-02	0,378
Role of JAK family kinases in IL-6-type Cytokine Signaling	1,1E00	0.08	#NUM!
D-myo-inositol (1,4,5)-Trisphosphate Biosynthesis	1,1E00	0.08	#NUM!
Huntington's Disease Signaling	1,09E00	3,38E-02	0,000

Gap Junction Signaling	1,09E00	3,54E-02	#NUM!
Phagosome Formation	1,06E00	0.04	#NUM!
Dopamine-DARPP32 Feedback in cAMP Signaling	1,05E00	3,68E-02	1,633
Lysine Degradation V	1,04E00	0.2	#NUM!
Role of CHK Proteins in Cell Cycle Checkpoint Control	1,04E00	5,26E-02	#NUM!
Nicotine Degradation III	1,04E00	5,26E-02	#NUM!
IL-1 Signaling	1,03E00	4,4E-02	#NUM!
Calcium Signaling	1,02E00	3,4E-02	1,633
Tight Junction Signaling	1,01E00	3,57E-02	#NUM!
Gα12/13 Signaling	1	3,85E-02	-1,342
Maturity Onset Diabetes of Young (MODY) Signaling	1	5,08E-02	#NUM!
Fcγ Receptor-mediated Phagocytosis in Macrophages and Monocytes	9,91E-01	4,26E-02	0,000
CCR5 Signaling in Macrophages	9,91E-01	4,26E-02	#NUM!
Melanocyte Development and Pigmentation Signaling	9,91E-01	4,26E-02	0,000
ErbB Signaling	9,91E-01	4,26E-02	0,000
Sonic Hedgehog Signaling	9,87E-01	6,9E-02	#NUM!
Melatonin Degradation I	9,83E-01	0.05	#NUM!
Ceramide Biosynthesis	9,67E-01	1,67E-01	#NUM!
Adipogenesis pathway	9,63E-01	3,73E-02	#NUM!
ATM Signaling	9,55E-01	4,12E-02	-1,000
G Protein Signaling Mediated by Tubby	9,39E-01	6,45E-02	#NUM!
Thrombopoietin Signaling	9,36E-01	4,76E-02	#NUM!
VEGF Signaling	9,32E-01	4,04E-02	#NUM!
IL-17A Signaling in Airway Cells	9,21E-01	4,69E-02	#NUM!
Phospholipases	9,21E-01	4,69E-02	#NUM!
4-1BB Signaling in T Lymphocytes	9,17E-01	6,25E-02	#NUM!
Kinetochore Metaphase Signaling Pathway	9,1E-01	3,96E-02	-1,000
CD40 Signaling	9,07E-01	4,62E-02	#NUM!
Nicotine Degradation II	9,07E-01	4,62E-02	#NUM!
Superpathway of Melatonin Degradation	9,07E-01	4,62E-02	#NUM!
Trehalose Degradation II (Trehalase)	9,03E-01	1,43E-01	#NUM!
Thioredoxin Pathway	9,03E-01	1,43E-01	#NUM!
SAPK/JNK Signaling	8,96E-01	3,92E-02	-1,000
Circadian Rhythm Signaling	8,93E-01	6,06E-02	#NUM!
Cardiac β-adrenergic Signaling	8,93E-01	3,55E-02	#NUM!
Dendritic Cell Maturation	8,79E-01	3,28E-02	0,816
Aryl Hydrocarbon Receptor Signaling	8,76E-01	3,5E-02	1,000
IGF-1 Signaling	8,76E-01	3,85E-02	-1,000
Cytotoxic T Lymphocyte-mediated Apoptosis of Target Cells	8,73E-01	5,88E-02	#NUM!
MIF-mediated Glucocorticoid Regulation	8,73E-01	5,88E-02	#NUM!
Hepatic Cholestasis	8,63E-01	3,24E-02	#NUM!
DNA Methylation and Transcriptional Repression Signaling	8,54E-01	5,71E-02	#NUM!
IL-17A Signaling in Fibroblasts	8,54E-01	5,71E-02	#NUM!
Inositol Pyrophosphates Biosynthesis	8,51E-01	1,25E-01	#NUM!

Salvage Pathways of Pyrimidine Deoxyribonucleotides	8,51E-01	1,25E-01	#NUM!
Gαs Signaling	8,45E-01	3,74E-02	0,000
Systemic Lupus Erythematosus In B Cell Signaling Pathway	8,27E-01	2,91E-02	0,707
Small Cell Lung Cancer Signaling	8,24E-01	4,23E-02	#NUM!
Pancreatic Adenocarcinoma Signaling	8,24E-01	3,67E-02	-2,000
Notch Signaling	8,15E-01	5,41E-02	#NUM!
Actin Nucleation by ARP-WASP Complex	8,12E-01	4,17E-02	#NUM!
Type I Diabetes Mellitus Signaling	8,04E-01	3,6E-02	0,000
Phosphatidylethanolamine Biosynthesis II	8,04E-01	1,11E-01	#NUM!
Clathrin-mediated Endocytosis Signaling	8,01E-01	3,11E-02	#NUM!
Caveolar-mediated Endocytosis Signaling	7,99E-01	4,11E-02	#NUM!
Thyroid Hormone Metabolism II (via Conjugation and/or Degradation)	7,96E-01	5,26E-02	#NUM!
Gustation Pathway	7,8E-01	3,25E-02	#NUM!
TREM1 Signaling	7,75E-01	0.04	#NUM!
IL-15 Signaling	7,75E-01	0.04	#NUM!
PKCθ Signaling in T Lymphocytes	7,72E-01	3,23E-02	-0,447
Macropinocytosis Signaling	7,64E-01	3,95E-02	#NUM!
Apelin Endothelial Signaling Pathway	7,64E-01	3,48E-02	0,000
Dolichyl-diphosphooligosaccharide Biosynthesis	7,62E-01	0.1	#NUM!
Glycine Betaine Degradation	7,62E-01	0.1	#NUM!
Superpathway of Inositol Phosphate Compounds	7,59E-01	3,02E-02	0,000
D-myo-inositol-5-phosphate Metabolism	7,57E-01	3,18E-02	0,447
Xenobiotic Metabolism Signaling	7,54E-01	2,79E-02	#NUM!
Fc Epsilon RI Signaling	7,47E-01	3,42E-02	1,000
Mechanisms of Viral Exit from Host Cells	7,45E-01	4,88E-02	#NUM!
eNOS Signaling	7,42E-01	3,14E-02	0,000
Glucocorticoid Receptor Signaling	7,33E-01	2,68E-02	#NUM!
Sirtuin Signaling Pathway	7,33E-01	2,75E-02	-0,447
Thyroid Cancer Signaling	7,3E-01	3,8E-02	#NUM!
MIF Regulation of Innate Immunity	7,28E-01	4,76E-02	#NUM!
Triacylglycerol Biosynthesis	7,28E-01	4,76E-02	#NUM!
γ-glutamyl Cycle	7,24E-01	9,09E-02	#NUM!
GDP-glucose Biosynthesis	7,24E-01	9,09E-02	#NUM!
CD28 Signaling in T Helper Cells	7,19E-01	3,33E-02	0,000
FLT3 Signaling in Hematopoietic Progenitor Cells	7,19E-01	3,75E-02	#NUM!
Chemokine Signaling	7,19E-01	3,75E-02	#NUM!
LXR/RXR Activation	7,12E-01	3,31E-02	0,000
BAG2 Signaling Pathway	7,12E-01	4,65E-02	#NUM!
G Beta Gamma Signaling	7,03E-01	3,28E-02	1,000
Glioblastoma Multiforme Signaling	6,97E-01	3,03E-02	1,000
IL-23 Signaling Pathway	6,97E-01	4,55E-02	#NUM!
RhoA Signaling	6,95E-01	3,25E-02	1,000
mTOR Signaling	6,88E-01	2,86E-02	-1,342
CCR3 Signaling in Eosinophils	6,86E-01	3,23E-02	#NUM!

Regulation Of The Epithelial Mesenchymal Transition In Development Pathway	6,76E-01	3,57E-02	#NUM!
BMP 340ignalling pathway	6,68E-01	3,53E-02	#NUM!
UDP-N-acetyl-D-galactosamine Biosynthesis II	6,6E-01	7,69E-02	#NUM!
nNOS Signaling in Neurons	6,56E-01	4,26E-02	#NUM!
Triacylglycerol Degradation	6,56E-01	4,26E-02	#NUM!
Endocannabinoid Neuronal Synapse Pathway	6,54E-01	3,12E-02	2,000
Reelin Signaling in Neurons	6,46E-01	3,1E-02	0,000
Ceramide Signaling	6,38E-01	3,41E-02	#NUM!
Cellular Effects of Sildenafil (Viagra)	6,31E-01	3,05E-02	#NUM!
Mevalonate Pathway I	6,31E-01	7,14E-02	#NUM!
Acute Myeloid Leukemia Signaling	6,29E-01	3,37E-02	#NUM!
OX40 Signaling Pathway	6,2E-01	3,33E-02	#NUM!
Th17 Activation Pathway	6,11E-01	3,3E-02	#NUM!
Choline Biosynthesis III	6,06E-01	6,67E-02	#NUM!
UVC-Induced MAPK Signaling	6,04E-01	3,92E-02	#NUM!
NF-κB Signaling	6,04E-01	2,79E-02	1,000
Ephrin Receptor Signaling	5,99E-01	2,78E-02	1,342
Androgen Signaling	5,95E-01	2,94E-02	#NUM!
G-Protein Coupled Receptor Signaling	5,95E-01	2,57E-02	#NUM!
UVB-Induced MAPK Signaling	5,92E-01	3,85E-02	#NUM!
Protein Ubiquitination Pathway	5,92E-01	2,56E-02	#NUM!
Iron homeostasis 340ignalling pathway	5,88E-01	2,92E-02	#NUM!
Regulation of Actin-based Motility by Rho	5,85E-01	3,19E-02	#NUM!
Granzyme B Signaling	5,82E-01	6,25E-02	#NUM!
Phototransduction Pathway	5,8E-01	3,77E-02	#NUM!
Neuregulin Signaling	5,67E-01	3,12E-02	#NUM!
Insulin Receptor Signaling	5,67E-01	2,86E-02	1,000
Salvage Pathways of Pyrimidine Ribonucleotides	5,51E-01	3,06E-02	#NUM!
NRF2-mediated Oxidative Stress Response	5,47E-01	2,65E-02	-0,447
Cell Cycle Control of Chromosomal Replication	5,47E-01	3,57E-02	#NUM!
Xenobiotic Metabolism CAR Signaling Pathway	5,47E-01	2,65E-02	-0,447
Nitric Oxide Signaling in the Cardiovascular System	5,44E-01	3,03E-02	#NUM!
Apelin Cardiomyocyte Signaling Pathway	5,44E-01	3,03E-02	#NUM!
1D-myo-inositol Hexakisphosphate Biosynthesis II (Mammalian)	5,38E-01	5,56E-02	#NUM!
D-myo-inositol (1,3,4)-trisphosphate Biosynthesis	5,38E-01	5,56E-02	#NUM!
Superpathway of Geranylgeranyldiphosphate Biosynthesis I (via Mevalonate)	5,38E-01	5,56E-02	#NUM!
Xenobiotic Metabolism PXR Signaling Pathway	5,3E-01	2,6E-02	-1,342
DNA damage-induced 14-3-3σ Signaling	5,19E-01	5,26E-02	#NUM!
PPAR Signaling	5,06E-01	2,88E-02	#NUM!
Coronavirus Pathogenesis Pathway	5,04E-01	2,67E-02	-2,000
Leukocyte Extravasation Signaling	5,03E-01	2,54E-02	1,342

Cardiomyocyte Differentiation via BMP Receptors	0.5	0.05	#NUM!
CDP-diacylglycerol Biosynthesis I	0.5	0.05	#NUM!
Inflammasome pathway	0.5	0.05	#NUM!
Endoplasmic Reticulum Stress Pathway	4,83E-01	4,76E-02	#NUM!
Paxillin Signaling	4,79E-01	2,78E-02	#NUM!
Activation of IRF by Cytosolic Pattern Recognition Receptors	4,78E-01	3,17E-02	#NUM!
Neuroinflammation Signaling Pathway	4,78E-01	2,34E-02	-0,816
Polyamine Regulation in Colon Cancer	4,66E-01	4,55E-02	#NUM!
Colorectal Cancer Metastasis Signaling	4,66E-01	2,37E-02	-1,342
Phosphatidylglycerol Biosynthesis II (Non-plastidic)	4,66E-01	4,55E-02	#NUM!
PXR/RXR Activation	4,6E-01	3,08E-02	#NUM!
Netrin Signaling	4,6E-01	3,08E-02	#NUM!
Calcium-induced T Lymphocyte Apoptosis	4,51E-01	3,03E-02	#NUM!
Pyridoxal 5'-phosphate Salvage Pathway	4,51E-01	3,03E-02	#NUM!
Superpathway of D-myo-inositol (1,4,5)-trisphosphate Metabolism	4,51E-01	4,35E-02	#NUM!
HOTAIR Regulatory Pathway	4,5E-01	2,5E-02	0,000
ErbB4 Signaling	4,44E-01	2,99E-02	#NUM!
Serotonin Degradation	4,44E-01	2,99E-02	#NUM!
Role of JAK1, JAK2 and TYK2 in Interferon Signaling	4,37E-01	4,17E-02	#NUM!
Synaptogenesis Signaling Pathway	4,28E-01	2,24E-02	0,378
Role of JAK1 and JAK3 in γc Cytokine Signaling	4,27E-01	2,9E-02	#NUM!
Bupropion Degradation	4,23E-01	0.04	#NUM!
Role of PKR in Interferon Induction and Antiviral Response	4,21E-01	2,56E-02	#NUM!
Cdc42 Signaling	4,15E-01	2,4E-02	-2,000
Glycolysis I	4,09E-01	3,85E-02	#NUM!
Apelin Liver Signaling Pathway	4,09E-01	3,85E-02	#NUM!
Th1 Pathway	3,99E-01	2,48E-02	#NUM!
Non-Small Cell Lung Cancer Signaling	3,97E-01	2,74E-02	#NUM!
Th1 and Th2 Activation Pathway	3,96E-01	2,34E-02	#NUM!
Inhibition of ARE-Mediated mRNA Degradation Pathway	3,94E-01	2,46E-02	#NUM!
Glioma Invasiveness Signaling	3,9E-01	2,7E-02	#NUM!
FcγRIIB Signaling in B Lymphocytes	3,83E-01	2,67E-02	#NUM!
Neurotrophin/TRK Signaling	3,77E-01	2,63E-02	#NUM!
Superpathway of Cholesterol Biosynthesis	3,73E-01	3,45E-02	#NUM!
Antiproliferative Role of Somatostatin Receptor 2	3,7E-01	2,6E-02	#NUM!
cAMP-mediated signaling	3,68E-01	2,19E-02	2,000
VDR/RXR Activation	3,64E-01	2,56E-02	#NUM!
Breast Cancer Regulation by Stathmin1	3,61E-01	2,03E-02	#NUM!
p70S6K Signaling	3,58E-01	2,33E-02	#NUM!
White Adipose Tissue Browning Pathway	3,58E-01	2,33E-02	#NUM!
IL-3 Signaling	3,57E-01	2,53E-02	#NUM!
Acetone Degradation I (to Methylglyoxal)	3,52E-01	3,23E-02	#NUM!
Role of MAPK Signaling in the Pathogenesis of Influenza	3,51E-01	2,5E-02	#NUM!

JAK/Stat Signaling	3,51E-01	2,5E-02	#NUM!
NF-κB Activation by Viruses	3,38E-01	2,44E-02	#NUM!
STAT3 Pathway	3,3E-01	2,22E-02	#NUM!
TR/RXR Activation	3,27E-01	2,38E-02	#NUM!
HER-2 Signaling in Breast Cancer	3,27E-01	2,38E-02	#NUM!
VEGF Family Ligand-Receptor Interactions	3,27E-01	2,38E-02	#NUM!
FGF Signaling	3,27E-01	2,38E-02	#NUM!
Th2 Pathway	3,25E-01	2,21E-02	#NUM!
Regulation Of The Epithelial Mesenchymal Transition By Growth Factors Pathway	3,25E-01	2,13E-02	0,000
Synaptic Long Term Depression	3,21E-01	2,12E-02	2,000
Xenobiotic Metabolism AHR Signaling Pathway	3,21E-01	2,35E-02	#NUM!
Nucleotide Excision Repair Pathway	3,13E-01	2,86E-02	#NUM!
Regulation of the Epithelial-Mesenchymal Transition Pathway	3,1E-01	2,08E-02	#NUM!
Crosstalk between Dendritic Cells and Natural Killer Cells	0.3	2,25E-02	#NUM!
D-myo-inositol (1,4,5,6)-Tetrakisphosphate Biosynthesis	2,99E-01	2,11E-02	#NUM!
D-myo-inositol (3,4,5,6)-tetrakisphosphate Biosynthesis	2,99E-01	2,11E-02	#NUM!
Cell Cycle Regulation by BTG Family Proteins	2,97E-01	2,7E-02	#NUM!
Prostate Cancer Signaling	2,9E-01	2,2E-02	#NUM!
Docosahexaenoic Acid (DHA) Signaling	2,88E-01	2,63E-02	#NUM!
FAK Signaling	2,71E-01	2,11E-02	#NUM!
GABA Receptor Signaling	2,71E-01	2,11E-02	#NUM!
nNOS Signaling in Skeletal Muscle Cells	2,66E-01	2,44E-02	#NUM!
Estrogen Biosynthesis	2,66E-01	2,44E-02	#NUM!
Bladder Cancer Signaling	2,61E-01	2,06E-02	#NUM!
PAK Signaling	2,61E-01	2,06E-02	#NUM!
Epithelial Adherens Junction Signaling	2,61E-01	1,97E-02	#NUM!
Retinol Biosynthesis	2,6E-01	2,38E-02	#NUM!
Pyrimidine Ribonucleotides Interconversion	2,6E-01	2,38E-02	#NUM!
Oncostatin M Signaling	2,53E-01	2,33E-02	#NUM!
Role of Hypercytokinemia/hyperchemokinemia in the Pathogenesis of Influenza	2,53E-01	2,33E-02	#NUM!
3-phosphoinositide Degradation	2,48E-01	1,92E-02	#NUM!
Role of RIG1-like Receptors in Antiviral Innate Immunity	2,46E-01	2,27E-02	#NUM!
Pyrimidine Ribonucleotides De Novo Biosynthesis	2,46E-01	2,27E-02	#NUM!
Necroptosis Signaling Pathway	2,44E-01	1,91E-02	#NUM!
iNOS Signaling	2,4E-01	2,22E-02	#NUM!
NER Pathway	2,37E-01	1,94E-02	#NUM!
Dermatan Sulfate Biosynthesis (Late Stages)	2,34E-01	2,17E-02	#NUM!
Ephrin A Signaling	2,28E-01	2,13E-02	#NUM!
PD-1, PD-L1 cancer immunotherapy pathway	2,25E-01	1,89E-02	#NUM!
Spliceosomal Cycle	2,23E-01	2,08E-02	#NUM!

Chondroitin Sulfate Biosynthesis (Late Stages)	2,23E-01	2,08E-02	#NUM!
Stearate Biosynthesis I (Animals)	2,23E-01	2,08E-02	#NUM!
Hematopoiesis from Pluripotent Stem Cells	2,18E-01	2,04E-02	#NUM!
Cell Cycle: G2/M DNA Damage Checkpoint Regulation	2,18E-01	2,04E-02	#NUM!
Amyloid Processing	2,12E-01	0.02	#NUM!

Appendix table 18. Significantly enriched human pathways between the uninfected controls and infected samples at 12 h post infection.

Ingenuity Canonical Pathways	-log(p-value)	Ratio	z-score
FXR/RXR Activation	5,93E00	1,27E-01	#NUM!
Protein Kinase A Signaling	4,94E00	7,27E-02	1,706
Antioxidant Action of Vitamin C	3,98E00	1,1E-01	1,134
PKCθ Signaling in T Lymphocytes	3,61E00	9,03E-02	0,302
LXR/RXR Activation	3,55E00	9,92E-02	0,632
TWEAK Signaling	3,26E00	1,71E-01	0,000
Cardiac Hypertrophy Signaling (Enhanced)	3,08E00	5,75E-02	1,347
Extrinsic Prothrombin Activation Pathway	2,96E00	2,5E-01	#NUM!
Opioid Signaling Pathway	2,88E00	6,88E-02	1,000
B Cell Receptor Signaling	2,85E00	7,57E-02	0,000
Corticotropin Releasing Hormone Signaling	2,84E00	8,28E-02	0,905
Death Receptor Signaling	2,79E00	9,89E-02	1,000
Endocannabinoid Neuronal Synapse Pathway	2,79E00	8,59E-02	0,632
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	2,78E00	7,45E-02	-0,832
White Adipose Tissue Browning Pathway	2,76E00	8,53E-02	3,317
Relaxin Signaling	2,72E00	0.08	1,890
GADD45 Signaling	2,67E00	2,11E-01	#NUM!
Wnt/Ca+ pathway	2,61E00	1,13E-01	1,890
Apoptosis Signaling	2,54E00	9,09E-02	1,000
PI3K Signaling in B Lymphocytes	2,53E00	7,97E-02	0,302
Sirtuin Signaling Pathway	2,5E00	6,19E-02	0,832
PEDF Signaling	2,5E00	9,76E-02	0,378
Cell Cycle: G2/M DNA Damage Checkpoint Regulation	2,48E00	1,22E-01	-1,633
Coagulation System	2,44E00	1,43E-01	-1,342
γ-glutamyl Cycle	2,43E00	2,73E-01	#NUM!
Type II Diabetes Mellitus Signaling	2,43E00	7,75E-02	-2,121
Calcium Signaling	2,42E00	6,8E-02	2,714
Phospholipase C Signaling	2,31E00	6,23E-02	2,309
CD27 Signaling in Lymphocytes	2,31E00	1,13E-01	-2,236

Regulation of IL-2 Expression in Activated and Anergic T Lymphocytes	2,28E00	8,99E-02	#NUM!
ERK5 Signaling	2,24E00	9,72E-02	-0,378
Gα12/13 Signaling	2,24E00	7,69E-02	-1,000
April Mediated Signaling	2,23E00	1,28E-01	-0,447
IL-12 Signaling and Production in Macrophages	2,19E00	7,58E-02	#NUM!
Signaling by Rho Family GTPases	2,15E00	6,15E-02	0,832
Glycolysis I	2,15E00	1,54E-01	2,000
B Cell Activating Factor Signaling	2,14E00	1,22E-01	-1,000
Role of IL-17F in Allergic Inflammatory Airway Diseases	2,09E00	1,19E-01	-0,447
Gαq Signaling	2,09E00	6,96E-02	0,333
Semaphorin Signaling in Neurons	2,05E00	0.1	#NUM!
Induction of Apoptosis by HIV1	2,01E00	9,84E-02	0,000
Hepatic Cholestasis	2	6,49E-02	#NUM!
Thrombin Signaling	1,99E00	6,25E-02	0,333
iNOS Signaling	1,97E00	1,11E-01	-2,236
Kinetochore Metaphase Signaling Pathway	1,96E00	7,92E-02	2,121
Osteoarthritis Pathway	1,94E00	6,16E-02	-0,277
TNFR2 Signaling	1,93E00	1,33E-01	-2,000
Netrin Signaling	1,88E00	9,23E-02	2,449
4-1BB Signaling in T Lymphocytes	1,83E00	1,25E-01	#NUM!
Insulin Secretion Signaling Pathway	1,82E00	5,76E-02	3,207
GNRH Signaling	1,81E00	6,36E-02	1,414
Myc Mediated Apoptosis Signaling	1,78E00	0.1	2,236
TNFR1 Signaling	1,78E00	0.1	-1,342
IL-10 Signaling	1,76E00	8,7E-02	#NUM!
Gustation Pathway	1,75E00	6,49E-02	#NUM!
Acute Phase Response Signaling	1,71E00	6,15E-02	-0,333
IL-1 Signaling	1,71E00	7,69E-02	-2,000
IL-17A Signaling in Fibroblasts	1,7E00	1,14E-01	#NUM!
Ephrin Receptor Signaling	1,7E00	6,11E-02	1,667
NGF Signaling	1,67E00	7,02E-02	-0,707
Hypoxia Signaling in the Cardiovascular System	1,63E00	8,11E-02	1,000
Sphingosine-1-phosphate Signaling	1,61E00	6,84E-02	0,378
Cardiac β-adrenergic Signaling	1,58E00	6,38E-02	1,134
Synaptogenesis Signaling Pathway	1,58E00	5,13E-02	2,840
AMPK Signaling	1,55E00	5,61E-02	1,000
PPARα/RXRα Activation	1,55E00	5,79E-02	2,121
Cardiac Hypertrophy Signaling	1,53E00	5,42E-02	2,309
Hepatic Fibrosis Signaling Pathway	1,53E00	4,89E-02	0,471
Heparan Sulfate Biosynthesis	1,53E00	7,69E-02	0,816
Sulfite Oxidation IV	1,52E00	1	#NUM!
Maturity Onset Diabetes of Young (MODY) Signaling	1,5E00	8,47E-02	#NUM!
nNOS Signaling in Skeletal Muscle Cells	1,47E00	9,76E-02	#NUM!
PTEN Signaling	1,44E00	6,35E-02	0,000
Intrinsic Prothrombin Activation Pathway	1,44E00	9,52E-02	-1,000

Atherosclerosis Signaling	1,43E00	6,3E-02	#NUM!
P2Y Purigenic Receptor Signaling Pathway	1,43E00	6,3E-02	2,121
LPS/IL-1 Mediated Inhibition of RXR Function	1,42E00	5,36E-02	0,000
T Cell Receptor Signaling	1,41E00	6,67E-02	#NUM!
tRNA Splicing	1,41E00	9,3E-02	1,000
TR/RXR Activation	1,4E00	7,14E-02	#NUM!
Regulation Of The Epithelial Mesenchymal Transition In Development Pathway	1,4E00	7,14E-02	0,000
Ketogenesis	1,39E00	1,82E-01	#NUM!
Role of RIG1-like Receptors in Antiviral Innate Immunity	1,38E00	9,09E-02	#NUM!
Role of NFAT in Regulation of the Immune Response	1,33E00	5,52E-02	0,707
Molecular Mechanisms of Cancer	1,32E00	4,6E-02	#NUM!
RANK Signaling in Osteoclasts	1,31E00	6,82E-02	-1,342
Dendritic Cell Maturation	1,31E00	5,46E-02	0,333
iCOS-iCOSL Signaling in T Helper Cells	1,31E00	6,31E-02	-0,447
Type I Diabetes Mellitus Signaling	1,31E00	6,31E-02	-1,633
GPCR-Mediated Nutrient Sensing in Enteroendocrine Cells	1,29E00	6,25E-02	2,646
Bile Acid Biosynthesis, Neutral Pathway	1,25E00	1,54E-01	#NUM!
Endocannabinoid Developing Neuron Pathway	1,24E00	6,09E-02	1,134
Role of NFAT in Cardiac Hypertrophy	1,24E00	5,14E-02	2,530
Axonal Guidance Signaling	1,23E00	4,34E-02	#NUM!
Alanine Degradation III	1,23E00	0.5	#NUM!
Alanine Biosynthesis II	1,23E00	0.5	#NUM!
Palmitate Biosynthesis I (Animals)	1,23E00	0.5	#NUM!
Fatty Acid Biosynthesis Initiation II	1,23E00	0.5	#NUM!
Glycine Degradation (Creatine Biosynthesis)	1,23E00	0.5	#NUM!
Small Cell Lung Cancer Signaling	1,21E00	7,04E-02	-2,000
Heparan Sulfate Biosynthesis (Late Stages)	1,21E00	7,04E-02	1,342
Role of Tissue Factor in Cancer	1,21E00	5,98E-02	#NUM!
Regulation of Actin-based Motility by Rho	1,2E00	6,38E-02	1,633
CCR5 Signaling in Macrophages	1,2E00	6,38E-02	#NUM!
Melanocyte Development and Pigmentation Signaling	1,2E00	6,38E-02	1,633
Mevalonate Pathway I	1,19E00	1,43E-01	#NUM!
Ephrin B Signaling	1,19E00	6,94E-02	1,000
ERK/MAPK Signaling	1,18E00	5,18E-02	1,667
GABA Receptor Signaling	1,18E00	6,32E-02	#NUM!
G-Protein Coupled Receptor Signaling	1,17E00	4,78E-02	#NUM!
Neuroinflammation Signaling Pathway	1,17E00	4,68E-02	-0,832
CD28 Signaling in T Helper Cells	1,16E00	5,83E-02	-1,000
Estrogen-Dependent Breast Cancer Signaling	1,15E00	6,76E-02	1,000
ATM Signaling	1,15E00	6,19E-02	-1,000
Superpathway of Citrulline Metabolism	1,14E00	1,33E-01	#NUM!
Adenosine Nucleotides Degradation II	1,14E00	1,33E-01	#NUM!
Lymphotoxin β Receptor Signaling	1,13E00	7,55E-02	#NUM!
Phototransduction Pathway	1,13E00	7,55E-02	#NUM!

Circadian Rhythm Signaling	1,12E00	9,09E-02	#NUM!
RhoA Signaling	1,11E00	5,69E-02	0,816
IL-8 Signaling	1,1E00	0.05	0,333
Cytotoxic T Lymphocyte-mediated Apoptosis of Target Cells	1,09E00	8,82E-02	#NUM!
Chondroitin Sulfate Degradation (Metazoa)	1,09E00	1,25E-01	#NUM!
Isoleucine Degradation I	1,09E00	1,25E-01	#NUM!
IL-6 Signaling	1,09E00	5,6E-02	-1,134
Cell Cycle Control of Chromosomal Replication	1,06E00	7,14E-02	1,000
Chondroitin Sulfate Biosynthesis	1,06E00	7,14E-02	1,000
Hypusine Biosynthesis	1,06E00	3,33E-01	#NUM!
Dermatan Sulfate Degradation (Metazoa)	1,04E00	1,18E-01	#NUM!
Synaptic Long Term Potentiation	1,03E00	5,43E-02	1,134
Reelin Signaling in Neurons	1,03E00	5,43E-02	0,378
CREB Signaling in Neurons	1,03E00	4,83E-02	2,449
Semaphorin Neuronal Repulsive Signaling Pathway	1,02E00	5,38E-02	-0,378
Cellular Effects of Sildenafil (Viagra)	1	5,34E-02	#NUM!
LPS-stimulated MAPK Signaling	1	6,1E-02	-1,342
NF-κB Activation by Viruses	1	6,1E-02	-1,342
D-myo-inositol-5-phosphate Metabolism	1	5,1E-02	1,890
Dermatan Sulfate Biosynthesis	1	6,78E-02	1,000
Purine Nucleotides Degradation II (Aerobic)	1	1,11E-01	#NUM!
Superpathway of Geranylgeranyldiphosphate Biosynthesis I (via Mevalonate)	1	1,11E-01	#NUM!
Gαs Signaling	9,91E-01	5,61E-02	2,449
Thyroid Hormone Metabolism II (via Conjugation and/or Degradation)	9,83E-01	7,89E-02	#NUM!
Melatonin Degradation I	9,79E-01	6,67E-02	1,000
Valine Degradation I	9,59E-01	1,05E-01	#NUM!
DNA damage-induced 14-3-3σ Signaling	9,59E-01	1,05E-01	#NUM!
Heme Degradation	9,39E-01	2,5E-01	#NUM!
Heme Biosynthesis from Uroporphyrinogen-III I	9,39E-01	2,5E-01	#NUM!
Proline Biosynthesis I	9,39E-01	2,5E-01	#NUM!
Fatty Acid β-oxidation III (Unsaturated, Odd Number)	9,39E-01	2,5E-01	#NUM!
Synaptic Long Term Depression	9,36E-01	4,76E-02	2,121
Dopamine-DARPP32 Feedback in cAMP Signaling	9,32E-01	4,91E-02	1,414
Activation of IRF by Cytosolic Pattern Recognition Receptors	9,21E-01	6,35E-02	-1,000
Rac Signaling	9,21E-01	5,36E-02	0,000
Actin Cytoskeleton Signaling	9,17E-01	4,59E-02	0,333
IL-17A Signaling in Airway Cells	9,03E-01	6,25E-02	0,000
Phospholipases	9,03E-01	6,25E-02	1,000
CD40 Signaling	8,86E-01	6,15E-02	0,000
Cdc42 Signaling	8,86E-01	4,79E-02	0,378
Superpathway of Melatonin Degradation	8,86E-01	6,15E-02	1,000
OX40 Signaling Pathway	8,76E-01	5,56E-02	0,000
Sperm Motility	8,73E-01	4,48E-02	0,000

Neuroprotective Role of THOP1 in Alzheimer's Disease	8,7E-01	5,17E-02	#NUM!
Serotonin Receptor Signaling	8,63E-01	6,98E-02	#NUM!
Endocannabinoid Cancer Inhibition Pathway	8,6E-01	4,9E-02	-1,134
Adrenomedullin signaling pathway	8,57E-01	4,57E-02	1,890
Pyrimidine Deoxyribonucleotides De Novo Biosynthesis I	8,54E-01	9,09E-02	#NUM!
Protein Citrullination	8,51E-01	0.2	#NUM!
Creatine-phosphate Biosynthesis	8,51E-01	0.2	#NUM!
α-tocopherol Degradation	8,51E-01	0.2	#NUM!
Lysine Degradation V	8,51E-01	0.2	#NUM!
Citrulline-Nitric Oxide Cycle	8,51E-01	0.2	#NUM!
Folate Polyglutamylation	8,51E-01	0.2	#NUM!
Th1 and Th2 Activation Pathway	8,48E-01	4,68E-02	#NUM!
Superpathway of Inositol Phosphate Compounds	8,39E-01	4,52E-02	1,414
Nur77 Signaling in T Lymphocytes	8,33E-01	5,88E-02	2,000
Apelin Cardiac Fibroblast Signaling Pathway	8,24E-01	8,7E-02	#NUM!
IL-15 Production	8,07E-01	4,96E-02	0,000
PI3K/AKT Signaling	8,07E-01	4,57E-02	-1,890
Dermatan Sulfate Biosynthesis (Late Stages)	7,99E-01	6,52E-02	#NUM!
FAT10 Cancer Signaling Pathway	7,99E-01	6,52E-02	#NUM!
PFKFB4 Signaling Pathway	7,99E-01	6,52E-02	#NUM!
G Beta Gamma Signaling	7,96E-01	4,92E-02	1,633
α-Adrenergic Signaling	7,93E-01	5,21E-02	#NUM!
Tumoricidal Function of Hepatic Natural Killer Cells	7,93E-01	8,33E-02	#NUM!
Role of JAK1, JAK2 and TYK2 in Interferon Signaling	7,93E-01	8,33E-02	#NUM!
Vitamin-C Transport	7,93E-01	8,33E-02	#NUM!
Factors Promoting Cardiogenesis in Vertebrates	7,85E-01	4,67E-02	2,646
Growth Hormone Signaling	7,85E-01	5,63E-02	#NUM!
Triacylglycerol Degradation	7,8E-01	6,38E-02	#NUM!
Arginine Biosynthesis IV	7,77E-01	1,67E-01	#NUM!
Pyruvate Fermentation to Lactate	7,77E-01	1,67E-01	#NUM!
Urea Cycle	7,77E-01	1,67E-01	#NUM!
Adenine and Adenosine Salvage III	7,77E-01	1,67E-01	#NUM!
Rapoport-Luebering Glycolytic Shunt	7,77E-01	1,67E-01	#NUM!
GDP-mannose Biosynthesis	7,77E-01	1,67E-01	#NUM!
Actin Nucleation by ARP-WASP Complex	7,72E-01	5,56E-02	#NUM!
p53 Signaling	7,67E-01	5,1E-02	#NUM!
D-myo-inositol (1,4,5)-Trisphosphate Biosynthesis	7,67E-01	0.08	#NUM!
Salvage Pathways of Pyrimidine Ribonucleotides	7,67E-01	5,1E-02	2,236
UVA-Induced MAPK Signaling	7,67E-01	5,1E-02	#NUM!
Chondroitin Sulfate Biosynthesis (Late Stages)	7,62E-01	6,25E-02	#NUM!
RhoGDI Signaling	7,59E-01	4,44E-02	-1,633
GPCR-Mediated Integration of Enteroendocrine Signaling Exemplified by an L Cell	7,57E-01	5,48E-02	0,000

Apelin Cardiomyocyte Signaling Pathway	7,54E-01	5,05E-02	1,342
Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses	7,45E-01	4,55E-02	#NUM!
Leptin Signaling in Obesity	7,42E-01	5,41E-02	#NUM!
Hematopoiesis from Pluripotent Stem Cells	7,42E-01	6,12E-02	#NUM!
NAD Salvage Pathway II	7,4E-01	7,69E-02	#NUM!
Gluconeogenesis I	7,4E-01	7,69E-02	#NUM!
TREM1 Signaling	7,28E-01	5,33E-02	0,000
FcγRIIB Signaling in B Lymphocytes	7,28E-01	5,33E-02	#NUM!
IL-15 Signaling	7,28E-01	5,33E-02	-1,000
Angiopoietin Signaling	7,28E-01	5,33E-02	1,000
Estrogen Receptor Signaling	7,26E-01	3,96E-02	3,051
Integrin Signaling	7,19E-01	4,23E-02	1,667
Trehalose Degradation II (Trehalase)	7,17E-01	1,43E-01	#NUM!
Thioredoxin Pathway	7,17E-01	1,43E-01	#NUM!
Purine Ribonucleosides Degradation to Ribose-1-phosphate	7,17E-01	1,43E-01	#NUM!
Glycoaminoglycan-protein Linkage Region Biosynthesis	7,17E-01	1,43E-01	#NUM!
Macropinocytosis Signaling	7,14E-01	5,26E-02	#NUM!
Hepatic Fibrosis / Hepatic Stellate Cell Activation	7,08E-01	4,3E-02	#NUM!
Chronic Myeloid Leukemia Signaling	7,08E-01	4,85E-02	#NUM!
PPAR Signaling	6,95E-01	4,81E-02	2,236
Systemic Lupus Erythematosus In B Cell Signaling Pathway	6,9E-01	0.04	-0,302
Xenobiotic Metabolism CAR Signaling Pathway	6,84E-01	4,23E-02	2,121
Glucocorticoid Receptor Signaling	6,78E-01	3,87E-02	#NUM!
ILK Signaling	6,76E-01	4,21E-02	0,378
Superpathway of Cholesterol Biosynthesis	6,68E-01	6,9E-02	#NUM!
Sonic Hedgehog Signaling	6,68E-01	6,9E-02	#NUM!
Role of Osteoblasts, Osteoclasts and Chondrocytes in Rheumatoid Arthritis	6,66E-01	4,09E-02	#NUM!
Salvage Pathways of Pyrimidine Deoxyribonucleotides	6,66E-01	1,25E-01	#NUM!
Role of BRCA1 in DNA Damage Response	6,62E-01	0.05	0,000
Virus Entry via Endocytic Pathways	6,62E-01	4,67E-02	#NUM!
Role of IL-17A in Arthritis	6,6E-01	5,56E-02	#NUM!
Xenobiotic Metabolism PXR Signaling Pathway	6,6E-01	4,17E-02	2,121
Tec Kinase Signaling	6,54E-01	4,27E-02	1,342
HMGB1 Signaling	6,46E-01	4,24E-02	-1,342
Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis	6,42E-01	3,85E-02	#NUM!
CXCR4 Signaling	6,29E-01	4,19E-02	0,000
G Protein Signaling Mediated by Tubby	6,25E-01	6,45E-02	#NUM!
Natural Killer Cell Signaling	6,22E-01	4,06E-02	0,707
Sucrose Degradation V (Mammalian)	6,22E-01	1,11E-01	#NUM!
Phosphatidylethanolamine Biosynthesis II	6,22E-01	1,11E-01	#NUM!
Citrulline Biosynthesis	6,22E-01	1,11E-01	#NUM!
Leucine Degradation I	6,22E-01	1,11E-01	#NUM!
Heme Biosynthesis II	6,22E-01	1,11E-01	#NUM!
Hereditary Breast Cancer Signaling	6,14E-01	4,29E-02	#NUM!

Nicotine Degradation III	6,14E-01	5,26E-02	#NUM!
Glutamate Receptor Signaling	6,14E-01	5,26E-02	#NUM!
cAMP-mediated signaling	6,09E-01	3,95E-02	2,121
Xenobiotic Metabolism AHR Signaling Pathway	0.6	4,71E-02	1,000
Germ Cell-Sertoli Cell Junction Signaling	5,97E-01	4,09E-02	#NUM!
D-myo-inositol (1,4,5,6)-Tetrakisphosphate Biosynthesis	5,97E-01	4,23E-02	1,342
D-myo-inositol (3,4,5,6)-tetrakisphosphate Biosynthesis	5,97E-01	4,23E-02	1,342
Aryl Hydrocarbon Receptor Signaling	5,88E-01	4,2E-02	#NUM!
Xenobiotic Metabolism General Signaling Pathway	5,88E-01	4,2E-02	0,447
Glycine Betaine Degradation	5,82E-01	0.1	#NUM!
Apelin Endothelial Signaling Pathway	5,82E-01	4,35E-02	-0,447
Retinoic acid Mediated Apoptosis Signaling	5,72E-01	0.05	#NUM!
MIF-mediated Glucocorticoid Regulation	5,67E-01	5,88E-02	#NUM!
HIF1α Signaling	5,64E-01	3,9E-02	2,121
Role of PKR in Interferon Induction and Antiviral Response	5,62E-01	4,27E-02	-1,342
Autophagy	5,59E-01	4,92E-02	#NUM!
Renin-Angiotensin Signaling	5,54E-01	4,24E-02	#NUM!
p38 MAPK Signaling	5,54E-01	4,24E-02	1,342
GDP-glucose Biosynthesis	5,47E-01	9,09E-02	#NUM!
Prostate Cancer Signaling	5,38E-01	4,4E-02	#NUM!
Interferon Signaling	5,33E-01	5,56E-02	#NUM!
Th1 Pathway	5,27E-01	4,13E-02	1,000
Epithelial Adherens Junction Signaling	5,16E-01	3,95E-02	#NUM!
Glucose and Glucose-1-phosphate Degradation	5,14E-01	8,33E-02	#NUM!
Glycogen Degradation II	5,14E-01	8,33E-02	#NUM!
Guanosine Nucleotides Degradation III	5,14E-01	8,33E-02	#NUM!
Nicotine Degradation II	5,1E-01	4,62E-02	#NUM!
Senescence Pathway	5,07E-01	3,64E-02	-1,667
Docosahexaenoic Acid (DHA) Signaling	0.5	5,26E-02	#NUM!
Mitotic Roles of Polo-Like Kinase	4,99E-01	4,55E-02	#NUM!
Eicosanoid Signaling	4,99E-01	4,55E-02	#NUM!
Sertoli Cell-Sertoli Cell Junction Signaling	4,96E-01	3,78E-02	#NUM!
Phagosome Formation	4,95E-01	0.04	#NUM!
Serotonin Degradation	4,88E-01	4,48E-02	#NUM!
3-phosphoinositide Degradation	4,87E-01	3,85E-02	1,342
Role of IL-17A in Psoriasis	4,85E-01	7,69E-02	#NUM!
Urate Biosynthesis/Inosine 5'-phosphate Degradation	4,85E-01	7,69E-02	#NUM!
NAD Phosphorylation and Dephosphorylation	4,85E-01	7,69E-02	#NUM!
UDP-N-acetyl-D-galactosamine Biosynthesis II	4,85E-01	7,69E-02	#NUM!
Antigen Presentation Pathway	4,85E-01	5,13E-02	#NUM!
Necroptosis Signaling Pathway	4,8E-01	3,82E-02	-0,816
Endothelin-1 Signaling	4,78E-01	3,72E-02	-0,378
Remodeling of Epithelial Adherens Junctions	4,78E-01	4,41E-02	#NUM!
eNOS Signaling	4,66E-01	3,77E-02	2,000
p70S6K Signaling	4,63E-01	3,88E-02	1,000

VEGF Signaling	4,62E-01	4,04E-02	#NUM!
Colorectal Cancer Metastasis Signaling	4,6E-01	3,56E-02	0,707
DNA Double-Strand Break Repair by Homologous Recombination	4,6E-01	7,14E-02	#NUM!
Leukotriene Biosynthesis	4,6E-01	7,14E-02	#NUM!
Glycogen Degradation III	4,6E-01	7,14E-02	#NUM!
Colanic Acid Building Blocks Biosynthesis	4,6E-01	7,14E-02	#NUM!
Estrogen Biosynthesis	4,57E-01	4,88E-02	#NUM!
Regulation of the Epithelial-Mesenchymal Transition Pathway	4,52E-01	3,65E-02	#NUM!
Clathrin-mediated Endocytosis Signaling	4,46E-01	3,63E-02	#NUM!
Neuropathic Pain Signaling In Dorsal Horn Neurons	4,46E-01	3,96E-02	1,000
Xenobiotic Metabolism Signaling	4,46E-01	3,48E-02	#NUM!
MIF Regulation of Innate Immunity	4,44E-01	4,76E-02	#NUM!
Retinol Biosynthesis	4,44E-01	4,76E-02	#NUM!
Pyrimidine Ribonucleotides Interconversion	4,44E-01	4,76E-02	#NUM!
SAPK/JNK Signaling	4,38E-01	3,92E-02	1,000
Sumoylation Pathway	4,29E-01	3,88E-02	#NUM!
T Helper Cell Differentiation	4,27E-01	4,11E-02	#NUM!
Glioblastoma Multiforme Signaling	4,27E-01	3,64E-02	2,236
Adipogenesis pathway	4,27E-01	3,73E-02	#NUM!
3-phosphoinositide Biosynthesis	4,21E-01	3,61E-02	1,342
Pyrimidine Ribonucleotides De Novo Biosynthesis	4,19E-01	4,55E-02	#NUM!
Apelin Pancreas Signaling Pathway	4,19E-01	4,55E-02	#NUM!
Gap Junction Signaling	4,18E-01	3,54E-02	#NUM!
Glutaryl-CoA Degradation	4,15E-01	6,25E-02	#NUM!
Androgen Signaling	4,13E-01	3,68E-02	1,000
Th2 Pathway	4,13E-01	3,68E-02	#NUM!
Iron homeostasis signaling pathway	4,07E-01	3,65E-02	#NUM!
PD-1, PD-L1 cancer immunotherapy pathway	4,06E-01	3,77E-02	-1,000
Erythropoietin Signaling	3,99E-01	3,95E-02	#NUM!
Role of Oct4 in Mammalian Embryonic Stem Cell Pluripotency	3,95E-01	4,35E-02	#NUM!
CDK5 Signaling	3,91E-01	3,7E-02	0,000
Antiproliferative Role of Somatostatin Receptor 2	3,9E-01	3,9E-02	#NUM!
Dopamine Receptor Signaling	3,9E-01	3,9E-02	#NUM!
Ephrin A Signaling	3,84E-01	4,26E-02	#NUM!
Agrin Interactions at Neuromuscular Junction	3,82E-01	3,85E-02	#NUM!
1D-myo-inositol Hexakisphosphate Biosynthesis II (Mammalian)	3,75E-01	5,56E-02	#NUM!
D-myo-inositol (1,3,4)-trisphosphate Biosynthesis	3,75E-01	5,56E-02	#NUM!
FLT3 Signaling in Hematopoietic Progenitor Cells	3,66E-01	3,75E-02	#NUM!
Protein Ubiquitination Pathway	3,64E-01	3,3E-02	#NUM!
Granzyme A Signaling	3,58E-01	5,26E-02	#NUM!
Cyclins and Cell Cycle Regulation	3,58E-01	3,7E-02	#NUM!
Apelin Muscle Signaling Pathway	3,58E-01	5,26E-02	#NUM!
mTOR Signaling	3,56E-01	3,33E-02	#NUM!

Primary Immunodeficiency Signaling	3,53E-01	0.04	#NUM!
Apelin Adipocyte Signaling Pathway	3,5E-01	3,66E-02	#NUM!
NF-κB Signaling	3,49E-01	3,35E-02	0,447
Cardiomyocyte Differentiation via BMP Receptors	3,41E-01	0.05	#NUM!
Inflammasome pathway	3,41E-01	0.05	#NUM!
fMLP Signaling in Neutrophils	3,37E-01	3,45E-02	1,000
FGF Signaling	3,35E-01	3,57E-02	#NUM!
Coronavirus Pathogenesis Pathway	3,3E-01	3,33E-02	-1,342
HIPPO signaling	3,28E-01	3,53E-02	#NUM!
IL-4 Signaling	3,28E-01	3,53E-02	#NUM!
BMP signaling pathway	3,28E-01	3,53E-02	#NUM!
Endoplasmic Reticulum Stress Pathway	3,26E-01	4,76E-02	#NUM!
Airway Pathology in Chronic Obstructive Pulmonary Disease	3,25E-01	3,39E-02	#NUM!
GP6 Signaling Pathway	3,2E-01	3,36E-02	-1,000
Polyamine Regulation in Colon Cancer	3,12E-01	4,55E-02	#NUM!
Regulation Of The Epithelial Mesenchymal Transition By Growth Factors Pathway	3,05E-01	3,19E-02	-0,816
Superpathway of D-myo-inositol (1,4,5)-trisphosphate Metabolism	2,98E-01	4,35E-02	#NUM!
Tryptophan Degradation III (Eukaryotic)	2,98E-01	4,35E-02	#NUM!
Unfolded protein response	2,98E-01	3,57E-02	#NUM!
Role of CHK Proteins in Cell Cycle Checkpoint Control	2,9E-01	3,51E-02	#NUM!
14-3-3-mediated Signaling	2,76E-01	3,15E-02	#NUM!
IL-17A Signaling in Gastric Cells	2,73E-01	0.04	#NUM!
Bupropion Degradation	2,73E-01	0.04	#NUM!
Leukocyte Extravasation Signalling	2,67E-01	3,05E-02	1,342
FAK Signalling	2,64E-01	3,16E-02	#NUM!
Apelin Liver Signalling Pathway	2,62E-01	3,85E-02	#NUM!
Neuregulin Signaling	2,58E-01	3,12E-02	#NUM!
TGF-β Signaling	2,58E-01	3,12E-02	#NUM!
Amyotrophic Lateral Sclerosis Signaling	2,53E-01	3,09E-02	#NUM!
PAK Signaling	2,53E-01	3,09E-02	#NUM!
Role of PI3K/AKT Signaling in the Pathogenesis of Influenza	2,4E-01	3,12E-02	#NUM!
PXR/RXR Activation	2,34E-01	3,08E-02	#NUM!
Regulation of Cellular Mechanics by Calpain Protease	2,34E-01	3,08E-02	#NUM!
Calcium-induced T Lymphocyte Apoptosis	2,28E-01	3,03E-02	#NUM!
Pyridoxal 5'-phosphate Salvage Pathway	2,28E-01	3,03E-02	#NUM!
Dopamine Degradation	2,23E-01	3,33E-02	#NUM!
Acetone Degradation I (to Methylglyoxal)	2,14E-01	3,23E-02	#NUM!
Glutathione-mediated Detoxification	2,06E-01	3,12E-02	#NUM!
Fatty Acid β-oxidation I	2,06E-01	3,12E-02	#NUM!
IL-9 Signalling	1,98E-01	3,03E-02	#NUM!

Appendix table 19. Significantly enriched human pathways between the uninfected controls and infected samples at 24 h post infection.

Ingenuity Canonical Pathways	-log(p-value)	Ratio	z-score
Bile Acid Biosynthesis, Neutral Pathway	6,05E00	3,08E-01	2,000
γ-glutamyl Cycle	4,46E00	2,73E-01	#NUM!
FXR/RXR Activation	3,91E00	4,76E-02	#NUM!
α-tocopherol Degradation	3,44E00	0.4	#NUM!
LXR/RXR Activation	3,06E00	4,13E-02	2,236
Atherosclerosis Signaling	2,97E00	3,94E-02	#NUM!
Methylglyoxal Degradation III	2,23E00	1,05E-01	#NUM!
p38 MAPK Signaling	2,23E00	3,39E-02	1,000
Glutamine Biosynthesis I	2,22E00	1	#NUM!
ERK/MAPK Signaling	2,19E00	2,59E-02	0,447
Calcium-induced T Lymphocyte Apoptosis	2,12E00	4,55E-02	#NUM!
Eicosanoid Signaling	2,12E00	4,55E-02	#NUM!
ERK5 Signaling	2,02E00	4,17E-02	#NUM!
Neuroinflammation Signaling Pathway	2	2,01E-02	1,000
Corticotropin Releasing Hormone Signaling	1,92E00	2,76E-02	#NUM!
β-alanine Degradation I	1,92E00	0.5	#NUM!
TR/RXR Activation	1,84E00	3,57E-02	#NUM!
Huntington's Disease Signaling	1,83E00	2,11E-02	0,000
Crosstalk between Dendritic Cells and Natural Killer Cells	1,77E00	3,37E-02	#NUM!
OX40 Signaling Pathway	1,76E00	3,33E-02	#NUM!
Cytotoxic T Lymphocyte-mediated Apoptosis of Target Cells	1,74E00	5,88E-02	#NUM!
4-aminobutyrate Degradation I	1,74E00	3,33E-01	#NUM!
GABA Receptor Signaling	1,69E00	3,16E-02	#NUM!
Phospholipase C Signaling	1,69E00	1,95E-02	#NUM!
GNRH Signaling	1,67E00	2,31E-02	#NUM!
Salvage Pathways of Pyrimidine Ribonucleotides	1,66E00	3,06E-02	#NUM!
Acute Phase Response Signaling	1,62E00	2,23E-02	#NUM!
Estrogen Biosynthesis	1,59E00	4,88E-02	#NUM!
Role of IL-17F in Allergic Inflammatory Airway Diseases	1,57E00	4,76E-02	#NUM!
Synaptic Long Term Depression	1,55E00	2,12E-02	#NUM!
Protein Citrullination	1,52E00	0.2	#NUM!
Lysine Degradation V	1,52E00	0.2	#NUM!
Glutamate Degradation III (via 4-aminobutyrate)	1,52E00	0.2	#NUM!
Clathrin-mediated Endocytosis Signaling	1,52E00	2,07E-02	#NUM!
iCOS-iCOSL Signaling in T Helper Cells	1,52E00	2,7E-02	#NUM!
Neuroprotective Role of THOP1 in Alzheimer's Disease	1,47E00	2,59E-02	#NUM!
Role of Tissue Factor in Cancer	1,46E00	2,56E-02	#NUM!
Renin-Angiotensin Signaling	1,45E00	2,54E-02	#NUM!
Glycerol Degradation I	1,44E00	1,67E-01	#NUM!
CD28 Signaling in T Helper Cells	1,43E00	2,5E-02	#NUM!
AMPK Signaling	1,38E00	1,87E-02	0,000
Synaptogenesis Signaling Pathway	1,38E00	1,6E-02	1,342

IL-12 Signaling and Production in Macrophages	1,33E00	2,27E-02	#NUM!
Nicotine Degradation III	1,33E00	3,51E-02	#NUM!
Sperm Motility	1,33E00	1,79E-02	#NUM!
Salvage Pathways of Pyrimidine Deoxyribonucleotides	1,32E00	1,25E-01	#NUM!
Maturity Onset Diabetes of Young (MODY) Signaling	1,3E00	3,39E-02	#NUM!
Th2 Pathway	1,3E00	2,21E-02	#NUM!
Iron homeostasis signaling pathway	1,29E00	2,19E-02	#NUM!
Melatonin Degradation I	1,29E00	3,33E-02	#NUM!
Systemic Lupus Erythematosus In T Cell Signaling Pathway	1,27E00	1,5E-02	1,000
Sucrose Degradation V (Mammalian)	1,27E00	1,11E-01	#NUM!
Glycine Betaine Degradation	1,23E00	0.1	#NUM!
Nicotine Degradation II	1,23E00	3,08E-02	#NUM!
Superpathway of Melatonin Degradation	1,23E00	3,08E-02	#NUM!
Insulin Secretion Signaling Pathway	1,21E00	1,65E-02	#NUM!
Factors Promoting Cardiogenesis in Vertebrates	1,2E00	0.02	#NUM!
Opioid Signaling Pathway	1,19E00	1,62E-02	1,000
Ketogenesis	1,19E00	9,09E-02	#NUM!
Gustation Pathway	1,17E00	1,95E-02	#NUM!
PKCθ Signaling in T Lymphocytes	1,16E00	1,94E-02	#NUM!
Guanosine Nucleotides Degradation III	1,15E00	8,33E-02	#NUM!
Estrogen-Dependent Breast Cancer Signaling	1,13E00	2,7E-02	#NUM!
Axonal Guidance Signaling	1,12E00	1,24E-02	#NUM!
Urate Biosynthesis/Inosine 5'-phosphate Degradation	1,12E00	7,69E-02	#NUM!
Cardiac Hypertrophy Signaling (Enhanced)	1,11E00	1,23E-02	0,816
Leukotriene Biosynthesis	1,09E00	7,14E-02	#NUM!
Mevalonate Pathway I	1,09E00	7,14E-02	#NUM!
FLT3 Signaling in Hematopoietic Progenitor Cells	1,07E00	2,5E-02	#NUM!
Th1 and Th2 Activation Pathway	1,06E00	1,75E-02	#NUM!
Adenosine Nucleotides Degradation II	1,06E00	6,67E-02	#NUM!
FGF Signaling	1,03E00	2,38E-02	#NUM!
Androgen Biosynthesis	1,03E00	6,25E-02	#NUM!
Role of NFAT in Regulation of the Immune Response	1,01E00	1,66E-02	#NUM!
Dendritic Cell Maturation	1	1,64E-02	#NUM!
CTLA4 Signaling in Cytotoxic T Lymphocytes	9,91E-01	2,25E-02	#NUM!
B Cell Receptor Signaling	9,87E-01	1,62E-02	#NUM!
Purine Nucleotides Degradation II (Aerobic)	9,83E-01	5,56E-02	#NUM!
Superpathway of Geranylgeranyldiphosphate Biosynthesis I (via Mevalonate)	9,83E-01	5,56E-02	#NUM!
Death Receptor Signaling	9,75E-01	2,2E-02	#NUM!
Endothelin-1 Signaling	9,71E-01	1,6E-02	#NUM!
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	9,71E-01	1,6E-02	#NUM!
ILK Signaling	9,63E-01	1,58E-02	#NUM!
Valine Degradation I	9,63E-01	5,26E-02	#NUM!
CCR5 Signaling in Macrophages	9,51E-01	2,13E-02	#NUM!

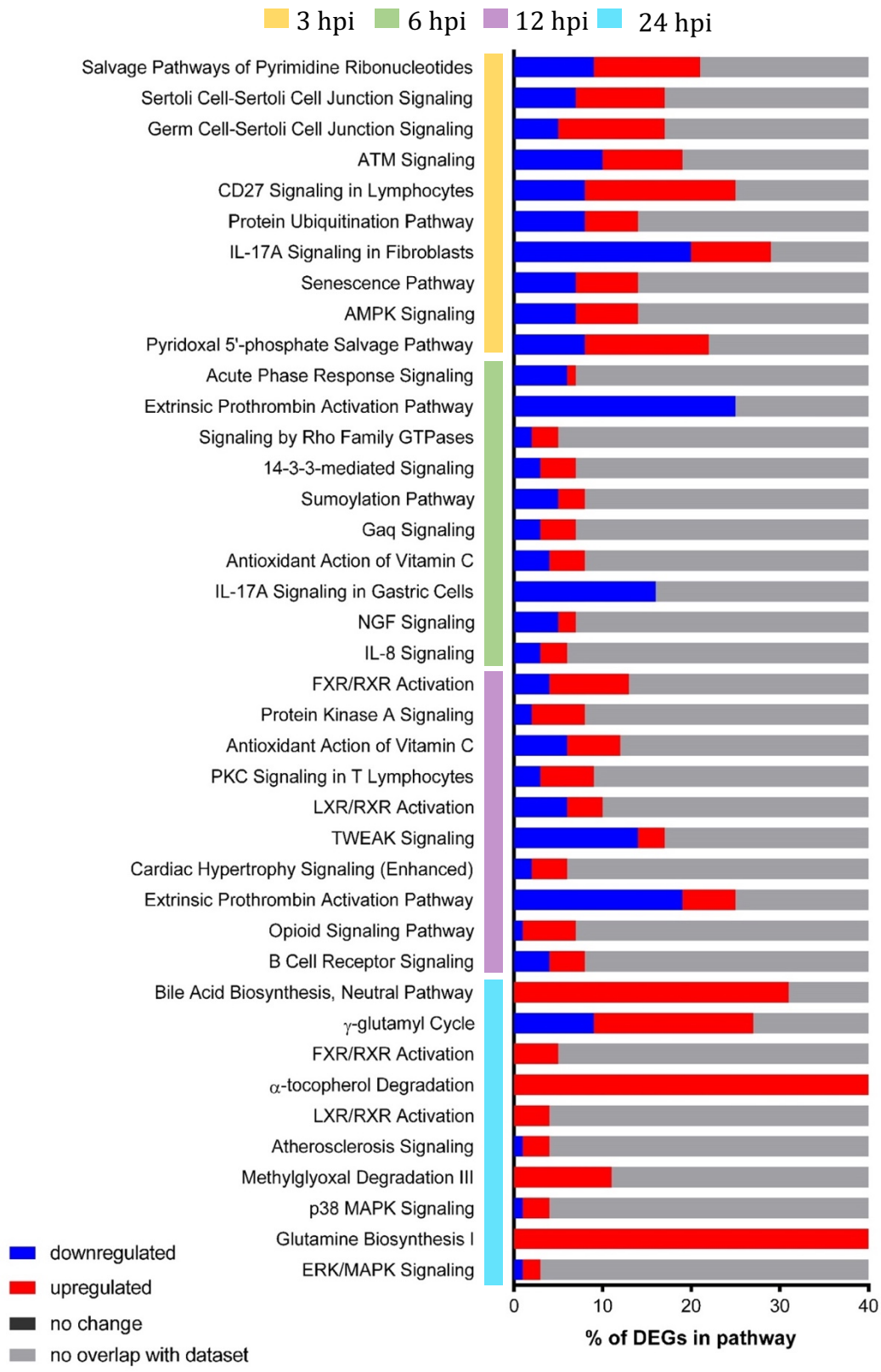
Melanocyte Development and Pigmentation Signaling	9,51E-01	2,13E-02	#NUM!
FAK Signaling	9,43E-01	2,11E-02	#NUM!
The Visual Cycle	9,39E-01	0.05	#NUM!
Amyotrophic Lateral Sclerosis Signaling	9,28E-01	2,06E-02	#NUM!
p53 Signaling	9,24E-01	2,04E-02	#NUM!
UVA-Induced MAPK Signaling	9,24E-01	2,04E-02	#NUM!
Apoptosis Signaling	9,17E-01	2,02E-02	#NUM!
Pyrimidine Deoxyribonucleotides De Novo Biosynthesis I	9,03E-01	4,55E-02	#NUM!
Calcium Signaling	8,86E-01	1,46E-02	#NUM!
CREB Signaling in Neurons	8,79E-01	1,45E-02	#NUM!
PD-1, PD-L1 cancer immunotherapy pathway	8,66E-01	1,89E-02	#NUM!
Osteoarthritis Pathway	8,63E-01	1,42E-02	#NUM!
Bupropion Degradation	8,51E-01	0.04	#NUM!
Type I Diabetes Mellitus Signaling	8,36E-01	1,8E-02	#NUM!
NAD Salvage Pathway II	8,36E-01	3,85E-02	#NUM!
Apelin Liver Signaling Pathway	8,36E-01	3,85E-02	#NUM!
Actin Cytoskeleton Signaling	8,33E-01	1,38E-02	#NUM!
Glucocorticoid Receptor Signaling	8,3E-01	1,19E-02	#NUM!
GPCR-Mediated Nutrient Sensing in Enteroendocrine Cells	8,3E-01	1,79E-02	#NUM!
NGF Signaling	8,18E-01	1,75E-02	#NUM!
Superpathway of Cholesterol Biosynthesis	7,9E-01	3,45E-02	#NUM!
Th1 Pathway	7,77E-01	1,65E-02	#NUM!
Acetone Degradation I (to Methylglyoxal)	7,64E-01	3,23E-02	#NUM!
CCR3 Signaling in Eosinophils	7,59E-01	1,61E-02	#NUM!
IL-6 Signaling	7,54E-01	1,6E-02	#NUM!
Phagosome Formation	7,54E-01	1,6E-02	#NUM!
Glutathione-mediated Detoxification	7,52E-01	3,12E-02	#NUM!
PTEN Signaling	7,5E-01	1,59E-02	#NUM!
P2Y Purigenic Receptor Signaling Pathway	7,45E-01	1,57E-02	#NUM!
Circadian Rhythm Signaling	7,4E-01	3,03E-02	#NUM!
Endocannabinoid Neuronal Synapse Pathway	7,4E-01	1,56E-02	#NUM!
Synaptic Long Term Potentiation	7,33E-01	1,55E-02	#NUM!
Reelin Signaling in Neurons	7,33E-01	1,55E-02	#NUM!
White Adipose Tissue Browning Pathway	7,33E-01	1,55E-02	#NUM!
Gα12/13 Signaling	7,28E-01	1,54E-02	#NUM!
MIF-mediated Glucocorticoid Regulation	7,28E-01	2,94E-02	#NUM!
Role of JAK2 in Hormone-like Cytokine Signaling	7,28E-01	2,94E-02	#NUM!
Retinoate Biosynthesis I	7,28E-01	2,94E-02	#NUM!
TWEAK Signaling	7,17E-01	2,86E-02	#NUM!
B Cell Development	7,06E-01	2,78E-02	#NUM!
Interferon Signaling	7,06E-01	2,78E-02	#NUM!
Complement System	6,95E-01	2,7E-02	#NUM!
Docosahexaenoic Acid (DHA) Signaling	6,84E-01	2,63E-02	#NUM!
Thyroid Hormone Metabolism II (via Conjugation and/or Degradation)	6,84E-01	2,63E-02	#NUM!
Antiproliferative Role of TOB in T Cell Signaling	6,74E-01	2,56E-02	#NUM!
Antigen Presentation Pathway	6,74E-01	2,56E-02	#NUM!

D-myo-inositol (1,4,5,6)-Tetrakisphosphate Biosynthesis	6,72E-01	1,41E-02	#NUM!
D-myo-inositol (3,4,5,6)-tetrakisphosphate Biosynthesis	6,72E-01	1,41E-02	#NUM!
Molecular Mechanisms of Cancer	6,7E-01	1,02E-02	#NUM!
Endocannabinoid Cancer Inhibition Pathway	6,68E-01	1,4E-02	#NUM!
Xenobiotic Metabolism General Signaling Pathway	6,68E-01	1,4E-02	#NUM!
nNOS Signaling in Skeletal Muscle Cells	6,56E-01	2,44E-02	#NUM!
MIF Regulation of Innate Immunity	6,46E-01	2,38E-02	#NUM!
Retinol Biosynthesis	6,46E-01	2,38E-02	#NUM!
Pyrimidine Ribonucleotides Interconversion	6,46E-01	2,38E-02	#NUM!
G-Protein Coupled Receptor Signaling	6,4E-01	1,1E-02	#NUM!
Serotonin Receptor Signaling	6,36E-01	2,33E-02	#NUM!
BAG2 Signaling Pathway	6,36E-01	2,33E-02	#NUM!
Coronavirus Pathogenesis Pathway	6,36E-01	1,33E-02	#NUM!
Phagosome Maturation	6,33E-01	1,32E-02	#NUM!
Pyrimidine Ribonucleotides De Novo Biosynthesis	6,29E-01	2,27E-02	#NUM!
Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses	6,2E-01	1,3E-02	#NUM!
Role of Oct4 in Mammalian Embryonic Stem Cell Pluripotency	6,13E-01	2,17E-02	#NUM!
FAT10 Cancer Signaling Pathway	6,13E-01	2,17E-02	#NUM!
PFKFB4 Signaling Pathway	6,13E-01	2,17E-02	#NUM!
3-phosphoinositide Degradation	6,11E-01	1,28E-02	#NUM!
D-myo-inositol-5-phosphate Metabolism	6,07E-01	1,27E-02	#NUM!
Necroptosis Signaling Pathway	6,07E-01	1,27E-02	#NUM!
Aldosterone Signaling in Epithelial Cells	6,04E-01	1,27E-02	#NUM!
Ephrin A Signaling	6,04E-01	2,13E-02	#NUM!
Gαq Signaling	6,04E-01	1,27E-02	#NUM!
eNOS Signaling	0.6	1,26E-02	#NUM!
Graft-versus-Host Disease Signaling	5,95E-01	2,08E-02	#NUM!
Autoimmune Thyroid Disease Signaling	5,88E-01	2,04E-02	#NUM!
Hematopoiesis from Pluripotent Stem Cells	5,88E-01	2,04E-02	#NUM!
Cell Cycle: G2/M DNA Damage Checkpoint Regulation	5,88E-01	2,04E-02	#NUM!
Dopamine-DARPP32 Feedback in cAMP Signaling	5,85E-01	1,23E-02	#NUM!
Sirtuin Signaling Pathway	5,85E-01	1,03E-02	#NUM!
Myc Mediated Apoptosis Signaling	5,8E-01	0.02	#NUM!
3-phosphoinositide Biosynthesis	5,73E-01	1,2E-02	#NUM!
CXCR4 Signaling	5,7E-01	1,2E-02	#NUM!
Cdc42 Signaling	5,7E-01	1,2E-02	#NUM!
CD27 Signaling in Lymphocytes	5,59E-01	1,89E-02	#NUM!
Lymphotoxin β Receptor Signaling	5,59E-01	1,89E-02	#NUM!
EGF Signaling	5,45E-01	1,82E-02	#NUM!
T Cell Exhaustion Signaling Pathway	5,42E-01	1,14E-02	#NUM!
Cell Cycle Control of Chromosomal Replication	5,39E-01	1,79E-02	#NUM!
Breast Cancer Regulation by Stathmin1	5,38E-01	8,46E-03	#NUM!
CNTF Signaling	5,32E-01	1,75E-02	#NUM!
Role of CHK Proteins in Cell Cycle Checkpoint Control	5,32E-01	1,75E-02	#NUM!

Glutamate Receptor Signaling	5,32E-01	1,75E-02	#NUM!
RhoGDI Signaling	5,26E-01	1,11E-02	#NUM!
Ephrin Receptor Signaling	5,26E-01	1,11E-02	#NUM!
Induction of Apoptosis by HIV1	5,07E-01	1,64E-02	#NUM!
Autophagy	5,07E-01	1,64E-02	#NUM!
Hepatic Fibrosis / Hepatic Stellate Cell Activation	5,06E-01	1,08E-02	#NUM!
Wnt/Ca+ pathway	5,02E-01	1,61E-02	#NUM!
Phospholipases	4,91E-01	1,56E-02	#NUM!
PXR/RXR Activation	4,85E-01	1,54E-02	#NUM!
Netrin Signaling	4,85E-01	1,54E-02	#NUM!
Regulation of Cellular Mechanics by Calpain Protease	4,85E-01	1,54E-02	#NUM!
Mitotic Roles of Polo-Like Kinase	4,8E-01	1,52E-02	#NUM!
Pyridoxal 5'-phosphate Salvage Pathway	4,8E-01	1,52E-02	#NUM!
Serotonin Degradation	4,75E-01	1,49E-02	#NUM!
Adrenomedullin signaling pathway	4,74E-01	1,02E-02	#NUM!
Nur77 Signaling in T Lymphocytes	4,7E-01	1,47E-02	#NUM!
Gap Junction Signaling	4,7E-01	1,01E-02	#NUM!
Superpathway of Inositol Phosphate Compounds	4,67E-01	1,01E-02	#NUM!
GM-CSF Signaling	4,6E-01	1,43E-02	#NUM!
Growth Hormone Signaling	4,55E-01	1,41E-02	#NUM!
Small Cell Lung Cancer Signaling	4,55E-01	1,41E-02	#NUM!
Basal Cell Carcinoma Signaling	4,5E-01	1,39E-02	#NUM!
Caveolar-mediated Endocytosis Signaling	4,45E-01	1,37E-02	#NUM!
T Helper Cell Differentiation	4,45E-01	1,37E-02	#NUM!
Non-Small Cell Lung Cancer Signaling	4,45E-01	1,37E-02	#NUM!
GPCR-Mediated Integration of Enteroendocrine Signaling Exemplified by an L Cell	4,45E-01	1,37E-02	#NUM!
Thrombin Signaling	4,42E-01	9,62E-03	#NUM!
Hypoxia Signaling in the Cardiovascular System	4,4E-01	1,35E-02	#NUM!
FcγRIIB Signaling in B Lymphocytes	4,35E-01	1,33E-02	#NUM!
IL-15 Signaling	4,35E-01	1,33E-02	#NUM!
GDNF Family Ligand-Receptor Interactions	4,31E-01	1,32E-02	#NUM!
Neurotrophin/TRK Signaling	4,31E-01	1,32E-02	#NUM!
Integrin Signaling	4,31E-01	9,39E-03	#NUM!
Role of NFAT in Cardiac Hypertrophy	4,27E-01	9,35E-03	#NUM!
Antiproliferative Role of Somatostatin Receptor 2	4,26E-01	1,3E-02	#NUM!
BEX2 Signaling Pathway	4,18E-01	1,27E-02	#NUM!
Role of BRCA1 in DNA Damage Response	4,13E-01	1,25E-02	#NUM!
Role of MAPK Signaling in the Pathogenesis of Influenza	4,13E-01	1,25E-02	#NUM!
JAK/Stat Signaling	4,13E-01	1,25E-02	#NUM!
Hepatic Fibrosis Signaling Pathway	4,12E-01	8,15E-03	#NUM!
PEDF Signaling	4,05E-01	1,22E-02	#NUM!
LPS/IL-1 Mediated Inhibition of RXR Function	4,03E-01	8,93E-03	#NUM!
VEGF Family Ligand-Receptor Interactions	3,97E-01	1,19E-02	#NUM!
cAMP-mediated signaling	3,94E-01	8,77E-03	#NUM!
IL-4 Signaling	3,93E-01	1,18E-02	#NUM!
BMP signaling pathway	3,93E-01	1,18E-02	#NUM!

Xenobiotic Metabolism AHR Signaling Pathway	3,93E-01	1,18E-02	#NUM!
Systemic Lupus Erythematosus Signaling	3,91E-01	8,73E-03	#NUM!
Allograft Rejection Signaling	3,89E-01	1,16E-02	#NUM!
Regulation of IL-2 Expression in Activated and Anergic T Lymphocytes	3,78E-01	1,12E-02	#NUM!
Altered T Cell and B Cell Signaling in Rheumatoid Arthritis	3,75E-01	1,11E-02	#NUM!
Prostate Cancer Signaling	3,71E-01	1,1E-02	#NUM!
Cardiac Hypertrophy Signaling	3,68E-01	8,33E-03	#NUM!
Fcγ Receptor-mediated Phagocytosis in Macrophages and Monocytes	3,61E-01	1,06E-02	#NUM!
Protein Kinase A Signaling	3,6E-01	7,52E-03	#NUM!
Signaling by Rho Family GTPases	3,6E-01	8,2E-03	#NUM!
α-Adrenergic Signaling	3,54E-01	1,04E-02	#NUM!
Communication between Innate and Adaptive Immune Cells	3,54E-01	1,04E-02	#NUM!
ATM Signaling	3,5E-01	1,03E-02	#NUM!
Nitric Oxide Signaling in the Cardiovascular System	3,44E-01	1,01E-02	#NUM!
VEGF Signaling	3,44E-01	1,01E-02	#NUM!
Apelin Cardiomyocyte Signaling Pathway	3,44E-01	1,01E-02	#NUM!
Neuropathic Pain Signaling In Dorsal Horn Neurons	3,37E-01	9,9E-03	#NUM!
Kinetochore Metaphase Signaling Pathway	3,37E-01	9,9E-03	#NUM!
Chronic Myeloid Leukemia Signaling	3,31E-01	9,71E-03	#NUM!
T Cell Receptor Signaling	3,25E-01	9,52E-03	#NUM!
Virus Entry via Endocytic Pathways	3,2E-01	9,35E-03	#NUM!
Gαs Signaling	3,2E-01	9,35E-03	#NUM!
CDK5 Signaling	3,16E-01	9,26E-03	#NUM!
Paxillin Signaling	3,16E-01	9,26E-03	#NUM!
Pancreatic Adenocarcinoma Signaling	3,13E-01	9,17E-03	#NUM!
Antioxidant Action of Vitamin C	3,13E-01	9,17E-03	#NUM!
Rac Signaling	3,05E-01	8,93E-03	#NUM!
Protein Ubiquitination Pathway	3,05E-01	7,33E-03	#NUM!
Senescence Pathway	3,02E-01	7,27E-03	#NUM!
Endocannabinoid Developing Neuron Pathway	2,98E-01	8,7E-03	#NUM!
fMLP Signaling in Neutrophils	2,95E-01	8,62E-03	#NUM!
Fc Epsilon RI Signaling	2,92E-01	8,55E-03	#NUM!
Sphingosine-1-phosphate Signaling	2,92E-01	8,55E-03	#NUM!
Cholecystokinin/Gastrin-mediated Signaling	2,87E-01	8,4E-03	#NUM!
Role of NANOG in Mammalian Embryonic Stem Cell Pluripotency	2,87E-01	8,4E-03	#NUM!
GP6 Signaling Pathway	2,87E-01	8,4E-03	#NUM!
IL-15 Production	2,82E-01	8,26E-03	#NUM!
G Beta Gamma Signaling	2,8E-01	8,2E-03	#NUM!
RhoA Signaling	2,77E-01	8,13E-03	#NUM!
Gαi Signaling	2,72E-01	0.008	#NUM!
p70S6K Signaling	2,64E-01	7,75E-03	#NUM!
Semaphorin Neuronal Repulsive Signaling Pathway	2,61E-01	7,69E-03	#NUM!
Cellular Effects of Sildenafil (Viagra)	2,59E-01	7,63E-03	#NUM!
Human Embryonic Stem Cell Pluripotency	2,51E-01	7,41E-03	#NUM!
STAT3 Pathway	2,51E-01	7,41E-03	#NUM!

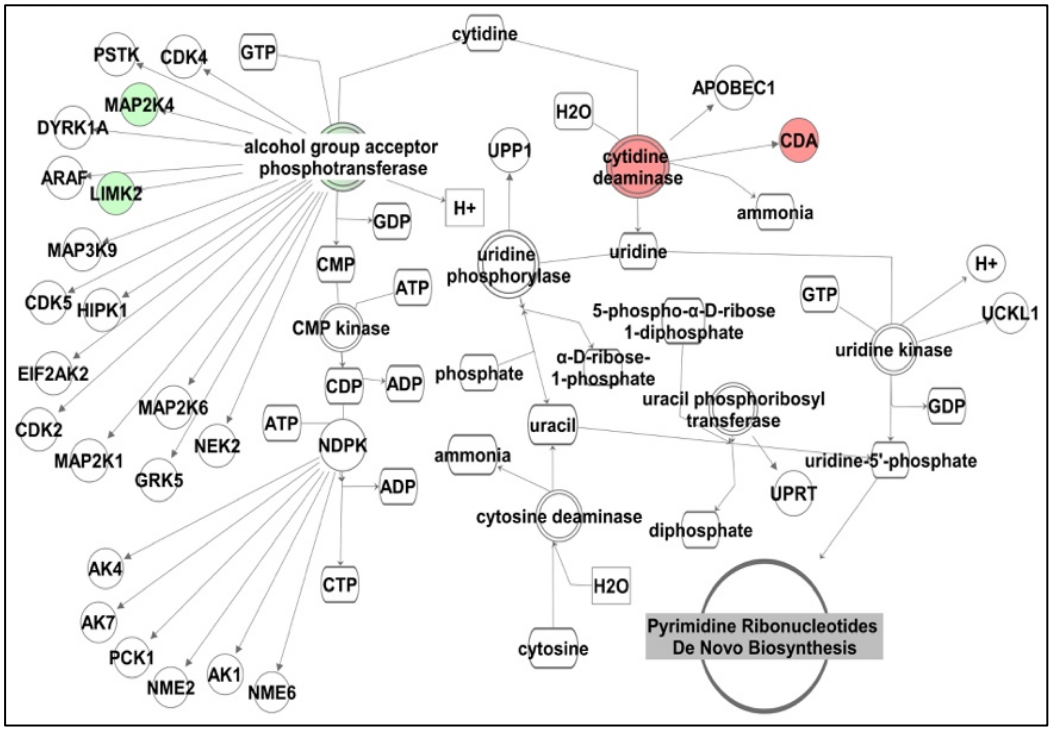
Androgen Signaling	2,49E-01	7,35E-03	#NUM!
Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis	2,46E-01	6,41E-03	#NUM!
PI3K Signaling in B Lymphocytes	2,45E-01	7,25E-03	#NUM!
Type II Diabetes Mellitus Signaling	2,37E-01	7,04E-03	#NUM!
Aryl Hydrocarbon Receptor Signaling	2,35E-01	6,99E-03	#NUM!
Estrogen Receptor Signaling	2,26E-01	6,1E-03	#NUM!
Relaxin Signaling	2,22E-01	6,67E-03	#NUM!
Epithelial Adherens Junction Signaling	2,18E-01	6,58E-03	#NUM!
HOTAIR Regulatory Pathway	2,05E-01	6,25E-03	#NUM!
Natural Killer Cell Signaling	0	5,08E-03	#NUM!
NRF2-mediated Oxidative Stress Response	0	5,29E-03	#NUM!
PPARα/RXRα Activation	0	5,26E-03	#NUM!
Hepatic Cholestasis	0	5,41E-03	#NUM!
Mitochondrial Dysfunction	0	5,85E-03	#NUM!
RAR Activation	0	5,15E-03	#NUM!
IL-8 Signaling	0	0.005	#NUM!
Colorectal Cancer Metastasis Signaling	0	3,95E-03	#NUM!
mTOR Signaling	0	4,76E-03	#NUM!
Role of Osteoblasts, Osteoclasts and Chondrocytes in Rheumatoid Arthritis	0	4,55E-03	#NUM!
Glioblastoma Multiforme Signaling	0	6,06E-03	#NUM!
Sertoli Cell-Sertoli Cell Junction Signaling	0	5,41E-03	#NUM!
Agranulocyte Adhesion and Diapedesis	0	5,18E-03	#NUM!
Regulation of the Epithelial-Mesenchymal Transition Pathway	0	5,21E-03	#NUM!
PI3K/AKT Signaling	0	5,71E-03	#NUM!
Xenobiotic Metabolism Signaling	0	3,48E-03	#NUM!
Wnt/β-catenin Signaling	0	5,78E-03	#NUM!
NF-κB Signaling	0	5,59E-03	#NUM!
Systemic Lupus Erythematosus In B Cell Signaling Pathway	0	3,64E-03	#NUM!
Xenobiotic Metabolism CAR Signaling Pathway	0	5,29E-03	#NUM!
Xenobiotic Metabolism PXR Signaling Pathway	0	5,21E-03	#NUM!
Regulation of The Epithelial Mesenchymal Transition By Growth Factors Pathway	0	5,32E-03	#NUM!



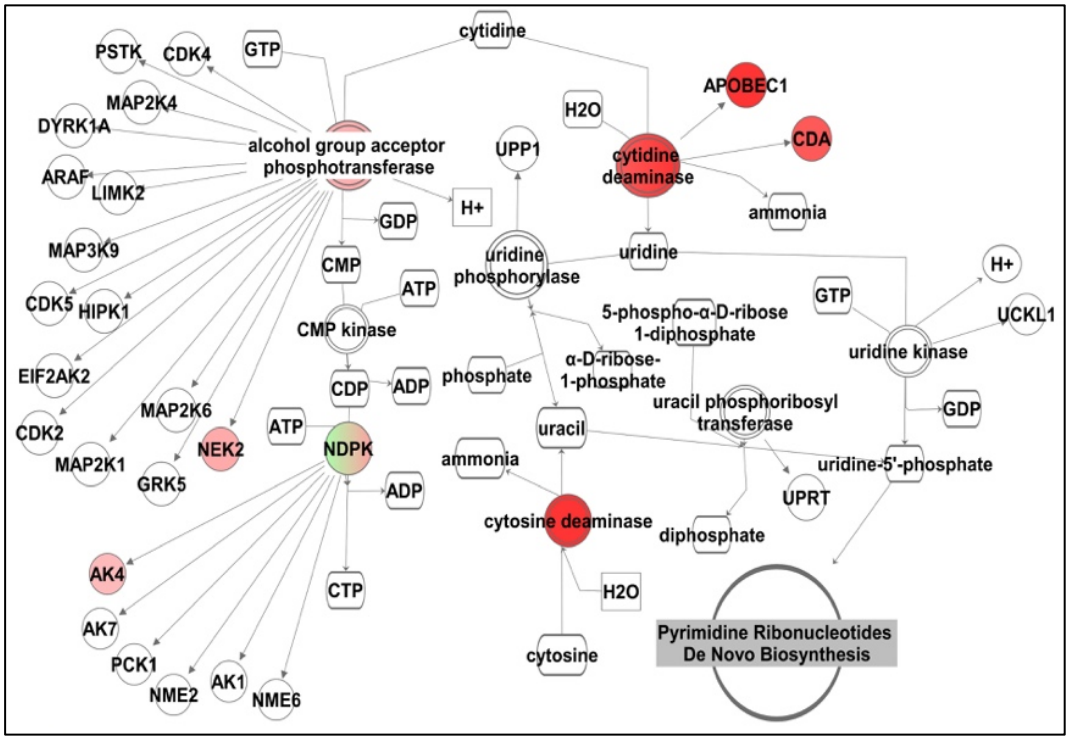
Appendix figure 1. The top 10 most significantly modulated pathways at each timepoint.

The percentage of upregulated (red) and downregulated (blue) genes in each pathway is indicated. Figure produced by Ludmila Rodrigues Pinto Ferreria.

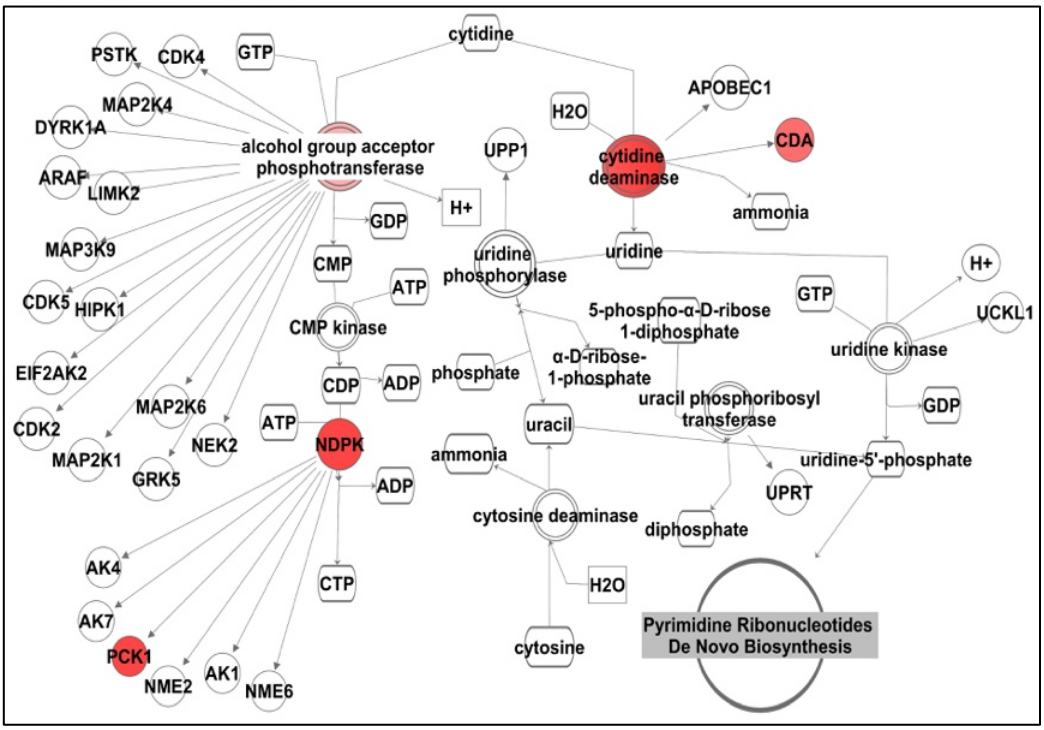
A.


cytidine
PSTK
CDK4
GTP
MAP2K4
DYRK1A
alcohol group acceptor phosphotransferase
ARAF
LIMK2
GDP
H+
MAP3K9
CMP
ATP
CDK5
HIPK1
CMP kinase
EIF2AK2
CDP
ADP
phosphate
MAP2K6
CDK2
ATP
NDPK
NEK2
MAP2K1
GRK5
ADP
AK4
CTP
AK7
PCK1
NME2
AK1
NME6
UPP1
H2O
cytidine deaminase
APOBEC1
CDA
ammonia
uridine
uridine phosphorylase
5-phospho-α-D-ribose 1-diphosphate
α-D-ribose-1-phosphate
uracil
uracil phosphoribosyl transferase
ammonia
cytosine deaminase
diphosphate
UPRT
uridine-5'-phosphate
H2O
cytosine
GTP
H+
UCKL1
uridine kinase
GDP
Pyrimidine Ribonucleotides De Novo Biosynthesis

B.


cytidine
PSTK
CDK4
GTP
MAP2K4
DYRK1A
alcohol group acceptor phosphotransferase
ARAF
LIMK2
GDP
H+
MAP3K9
CMP
ATP
CDK5
HIPK1
CMP kinase
EIF2AK2
CDP
ADP
phosphate
MAP2K6
CDK2
ATP
NDPK
NEK2
MAP2K1
GRK5
ADP
AK4
CTP
AK7
PCK1
NME2
AK1
NME6
UPP1
H2O
cytidine deaminase
APOBEC1
CDA
ammonia
uridine
uridine phosphorylase
5-phospho-α-D-ribose 1-diphosphate
α-D-ribose-1-phosphate
uracil
uracil phosphoribosyl transferase
ammonia
cytosine deaminase
diphosphate
UPRT
uridine-5'-phosphate
H2O
cytosine
GTP
H+
UCKL1
uridine kinase
GDP
Pyrimidine Ribonucleotides De Novo Biosynthesis

C.



Appendix figure 2. The differential expression of genes involved in the salvage pathways of pyrimidine ribonucleotides at 6 h (A), 12 h (B), and 24 h (C) post infection.

Upregulated genes are indicated by red colouring and downregulated genes are indicated by green colouring. Figure produced by Ludmila Rodrigues Pinto Ferreria with the Qiagen IPA software.

Bibliography

Abhyankar, M. M. *et al.* (2020) 'Immune Profiling To Predict Outcome of *Clostridioides difficile* Infection', *mBio*, 11(3), pp. 1–10.

Abt, M. C. *et al.* (2015) 'Innate immune defenses mediated by two ILC subsets are critical for protection against acute *Clostridium difficile* infection', *Cell Host and Microbe*, 18(1), pp. 27–37. doi: 10.1016/j.chom.2015.06.011.

Abt, M. C., McKenney, P. T. and Pamer, E. G. (2016) '*Clostridium difficile* colitis: pathogenesis and host defence', *Nature Reviews Microbiology*, 14(10), pp. 609–620. doi: 10.1038/s41395-018-0061-4.

Adiconis, X. *et al.* (2013) 'Comparative analysis of RNA sequencing methods for degraded or low-input samples', *Nature Methods*, 10(7), pp. 623–629. doi: 10.1038/NMETH.2483.

Aktories, K. and Wegner, A. (1992) 'Mechanisms of the cytopathic action of actin-ADP-ribosylating toxins', *Molecular Microbiology*, 6(20), pp. 2905–2908. doi: 10.1111/j.1365-2958.1992.tb01749.x.

Alcantara, C. *et al.* (2001) 'Role of Inducible Cyclooxygenase and Prostaglandins in *Clostridium difficile* Toxin A–Induced Secretion and Inflammation in an Animal Model', *The Journal of Infectious Diseases*, 184(5), pp. 648–652. doi: 10.1086/322799.

Anderson, D. J. *et al.* (2017) 'Enhanced terminal room disinfection and acquisition and infection caused by multidrug-resistant organisms and *Clostridium difficile* (the Benefits of Enhanced Terminal Room Disinfection study): a cluster-randomised, multicentre, crossover study', *The Lancet*,

389(10071), pp. 805–814. doi: 10.1016/S0140-6736(16)31588-4.

Ando, Y. *et al.* (2007) 'Inactivation of Rho GTPases with *Clostridium difficile* Toxin B Impairs Centrosomal Activation of Aurora-A in G2/M Transition of HeLa Cells', *Molecular Biology of the Cell*, 18(October), pp. 3752–3763. doi: 10.1091/mbc.E07.

Anjuwon-Foster, B. R. and Tamayo, R. (2017) *A genetic switch controls the production of flagella and toxins in* Clostridium difficile, *PLoS Genetics*. doi: 10.1371/journal.pgen.1006701.

Anonye, B. O. *et al.* (2019) 'Probing *Clostridium difficile* Infection in Complex Human Gut Cellular Models', *Frontiers in Microbiology*, 10, pp. 1–15. doi: 10.3389/fmicb.2019.00879.

Aprianto, R. *et al.* (2016) 'Time-resolved dual RNA-seq reveals extensive rewiring of lung epithelial and pneumococcal transcriptomes during early infection', *Genome Biology*, 17(1), pp. 1–16. doi: 10.1186/s13059-016-1054-5.

Arancibia, S. A. *et al.* (2007) 'Toll-like receptors are key participants in innate immune responses', *Biological Research*, 40(2), pp. 97–112. doi: 10.4067/S0716-97602007000200001.

Arato, V. *et al.* (2019) 'Dual role of the colonization factor CD2831 in *Clostridium difficile* pathogenesis', *Scientific Reports*, 9(1), p. 5554. doi: 10.1038/s41598-019-42000-8.

Arnon, S. S. *et al.* (1984) 'Rapid death of infant rhesus monkeys injected with *Clostridium difficile* toxins A and B: Physiologic and pathologic basis', *Journal of Pediatrics*, 104(1), pp. 34–40.

Aronsson, B. *et al.* (1985) 'Serum antibody response to *Clostridium difficile* toxins in patients with *Clostridium difficile* diarrhoea', *Infection*, 13(3), pp. 97–

101. doi: 10.1007/BF01642866.

Arroyo, L. G., Weese, J. S. and Staempfli, H. R. (2004) 'Experimental *Clostridium difficile* Enterocolitis in Foals', *Journal of Veterinary Internal Medicine*, 18(1), pp. 734–738. doi: 10.11501/3078355.

Avraham, R. *et al.* (2016) 'Pathogen Cell-to-cell Variability Drives Heterogeneity In Host Immune Responses', *HHS Access*, 162(6), pp. 1309–1321. doi: 10.1016/j.cell.2015.08.027.Pathogen.

Baban, S. T. *et al.* (2013) 'The Role of Flagella in *Clostridium difficile* Pathogenesis: Comparison between a Non-Epidemic and an Epidemic Strain', *PLoS ONE*, 8(9), p. e73026. doi: 10.1371/journal.pone.0073026.

Baddal, B. *et al.* (2015) 'Dual RNA-seq of Nontypeable *Haemophilus influenzae* and Host Cell Transcriptomes Reveals Novel Insights into Host-Pathogen Cross Talk', *mBio*, 6(6), pp. e01765-15. doi: 10.1128/mBio.01765-15.Editor.

Baines, S. D. *et al.* (2006) 'Tigecycline does not induce proliferation or cytotoxin production by epidemic *Clostridium difficile* strains in a human gut model', *Journal of Antimicrobial Chemotherapy*, 58(5), pp. 1062–1065. doi: 10.1093/jac/dkl364.

Baines, S. D. *et al.* (2008) 'Comparison of oritavancin versus vancomycin as treatments for clindamycin-induced *Clostridium difficile* PCR ribotype 027 infection in a human gut model', *Journal of Antimicrobial Chemotherapy*, 62(5), pp. 1078–1085. doi: 10.1093/jac/dkn358.

Baines, S. D. *et al.* (2009) 'Activity of vancomycin against epidemic *Clostridium difficile* strains in a human gut model', *Journal of Antimicrobial Chemotherapy*, 63(3), pp. 520–525. doi: 10.1093/jac/dkn502.

Baines, S. D., Freeman, J. and Wilcox, M. H. (2005) 'Effects of

piperacillin/tazobactam on *Clostridium difficile* growth and toxin production in a human gut model', *Journal of Antimicrobial Chemotherapy*, 55(6), pp. 974–982. doi: 10.1093/jac/dki120.

Barketi-Klai, A. *et al.* (2011) 'Role of fibronectin-binding protein A in *Clostridium difficile* intestinal colonization', *Journal of Medical Microbiology*, 60(8), pp. 1155–1161. doi: 10.1099/jmm.0.029553-0.

Barth, H. *et al.* (2004) 'Binary Bacterial Toxins: Biochemistry, Biology, and Applications of Common Clostridium and Bacillus Proteins', *Microbiology and Molecular Biology Reviews*, 68(3), pp. 373–402. doi: 10.1128/MMBR.68.3.373.

Bartlett, J. G. *et al.* (1978) 'Antibiotic-Associated Pseudomembranous Colitis Due to Toxin-Producing Clostridia', *New England Journal of Medicine*, 298, pp. 531–534.

Battaglioli, E. J. *et al.* (2018) '*Clostridioides difficile* uses amino acids associated with gut microbial dysbiosis in a subset of patients with diarrhea', *Science Translational Medicine*, 10(464), pp. 1–12. doi: 10.1126/scitranslmed.aam7019.

Baud'huin, M. *et al.* (2013) 'Osteoprotegerin : Multiple partners for multiple functions', *Cytokine & Growth Factor Reviews*, 24, pp. 1–9.

Beinke, S. and Ley, S. C. (2004) 'Functions of NF-κB1 and NF-κB2 in immune cell biology', *Biochemical Journal*, 382(2), pp. 393–409. doi: 10.1042/BJ20040544.

Best, E. L., Freeman, J. and Wilcox, M. H. (2012) 'Models for the study of *Clostridium difficile* infection', *Gut Microbes*, 3(2), pp. 145–167. doi: 10.4161/gmic.19526.

Blankenship-Paris, T. L. *et al.* (1995) '*Clostridium difficile* Infection in Hamsters Fed an Atherogenic Diet', *Veterinary Pathology*, 32(3), pp. 269–273. doi: 10.1177/030098589503200308.

Bordeleau, E. *et al.* (2015) 'Cyclic Di-GMP Riboswitch-Regulated Type IV Pili Contribute to Aggregation of *Clostridium difficile*', *Journal of Bacteriology*, 197(5), pp. 819–832. doi: 10.1128/JB.02340-14.

Borriello, S. P. *et al.* (1987) '*Clostridium difficile*: A spectrum of virulence and analysis of putative virulence determinants in the hamster model of antibiotic-associated colitis', *Journal of Medical Microbiology*, 24(1), pp. 53–64. doi: 10.1099/00222615-24-1-53.

Borriello, S. P. and Barclay, F. E. (1986) 'An *in-vitro* model of colonisation resistance to *Clostridium difficile* infection', *Journal of Medical Microbiology*, 21(4), pp. 299–309. doi: 10.1099/00222615-21-4-299.

Bouillaut, L., Self, W. T. and Sonenshein, A. L. (2013) 'Proline-Dependent Regulation of *Clostridium difficile* Stickland Metabolism', *Journal of Bacteriology*, 195(4), pp. 844–854. doi: 10.1128/JB.01492-12.

Bruxelle, J. F. *et al.* (2016) 'Immunogenic properties of the surface layer precursor of *Clostridium difficile* and vaccination assays in animal models', *Anaerobe*, 37, pp. 78–84. doi: 10.1016/j.anaerobe.2015.10.010.

Buonomo, E. L. *et al.* (2013) 'Role of interleukin 23 signaling in *Clostridium difficile* colitis', *Journal of Infectious Diseases*, 208(6), pp. 917–920. doi: 10.1093/infdis/jit277.

Butel, M. J. *et al.* (1998) 'Clostridial pathogenicity in experimental necrotising enterocolitis in gnotobiotic quails and protective role of bifidobacteria', *Journal of Medical Microbiology*, 47(5), pp. 391–399. doi: 10.1099/00222615-47-5-391.

Byrd, C. A. *et al.* (1999) 'Heat shock protein 90 mediates macrophage activation by Taxol and bacterial lipopolysaccharide', *Proceedings of the National Academy of Sciences of the United States of America*, 96(10), pp. 5645–5650. doi:

10.1073/pnas.96.10.5645.

Cafardi, V. *et al.* (2013) 'Identification of a novel zinc metalloprotease through a global analysis of *Clostridium difficile* extracellular proteins', *PLoS ONE*, 8(11), pp. 1–14. doi: 10.1371/journal.pone.0081306.

Calabi, E. *et al.* (2002) 'Binding of *Clostridium difficile* Surface Layer Proteins to Gastrointestinal Tissues', *International Journal of Cardiology*, 70(10), pp. 5770–5778. doi: 10.1128/IAI.70.10.5770.

Carter, G. P., Rood, J. I. and Lyras, D. (2012) 'The role of toxin A and toxin B in the virulence of *Clostridium difficile*', *Trends in Microbiology*, 20(1), pp. 21–29. doi: 10.1016/j.tim.2011.11.003.

Castagliuolo, I. *et al.* (1997) 'IL-11 inhibits enterotoxicity *Clostridium difficile* toxin A enterotoxicity in rat ileum', *American Physiological Society*, 273, pp. G333-41.

Cerquetti, M. *et al.* (2002) 'Binding of *Clostridium difficile* to Caco-2 epithelial cell line and to extracellular matrix proteins', *FEMS Immunology and Medical Microbiology*, 32(3), pp. 211–218. doi: 10.1016/S0928-8244(01)00301-7.

Chae, S. *et al.* (2006) 'Epithelial Cell IκB-Kinase β Has an Important Protective Role in *Clostridium difficile* Toxin A-Induced Mucosal Injury', *The Journal of Immunology*, 177(2), pp. 1214–1220. doi: 10.4049/jimmunol.177.2.1214.

Chang, T. W. *et al.* (1978) 'Clindamycin-induced enterocolitis in hamsters as a model of pseudomembranous colitis in patients', *Infection and Immunity*, 20(2), pp. 526–529. doi: 10.1128/iai.20.2.526-529.1978.

Chen, P. *et al.* (2018) 'Structural basis for recognition of frizzled proteins by *Clostridium difficile* toxin B', *Science*, 360(6389), pp. 664–669. doi: 10.1126/science.aar1999.

Chen, S. *et al.* (2015) 'The role of Rho GTPases in toxicity of *Clostridium difficile* toxins', *Toxins*, 7(12), pp. 5254–5267. doi: 10.3390/toxins7124874.

Chen, X. *et al.* (2006) '*Saccharomyces boulardii* inhibits ERK1/2 mitogen-activated protein kinase activation both in vitro and in vivo and protects against *Clostridium difficile* toxin A-induced enteritis', *Journal of Biological Chemistry*, 281(34), pp. 24449–24454. doi: 10.1074/jbc.M605200200.

Chen, X. *et al.* (2008) 'A Mouse Model of *Clostridium difficile*–Associated Disease', *Gastroenterology*, 135(6), pp. 1984–1992. doi: 10.1053/j.gastro.2008.09.002.

Chen, Y. S. *et al.* (2020) 'IL-17-producing γδ T cells protect against *Clostridium difficile* infection', *Journal of Clinical Investigation*, 130(5), pp. 2377–2390. doi: 10.1172/JCI127242.

Cheng, K. *et al.* (2013) 'Rab25 Small GTPase Mediates Secretion of Tumor Necrosis Factor Receptor Superfamily Member 11b (osteoprotegerin) Protecting Cancer Cells from Effects of TRAIL', *Journal of Genetic Syndromes & Gene Therapy*, 4. doi: 10.4172/2157-7412.1000153.Rab25.

Cho, J. S. *et al.* (2010) 'IL-17 is essential for host defense against cutaneous *Staphylococcus aureus* infection in mice', *Journal of Clinical Investigation*, 120(5), pp. 1762–1773. doi: 10.1172/JCI40891.

Collins, J. *et al.* (2015) 'Humanized microbiota mice as a model of recurrent *Clostridium difficile* disease', *Microbiome*, 3, p. 35. doi: 10.1186/s40168-015-0097-2.

Collins, J. *et al.* (2018) 'Dietary trehalose enhances virulence of epidemic *Clostridium difficile*', *HHS Access*, 553(7688), pp. 291–294. doi: 10.1038/nature25178.Dietary.

Corver, J. *et al.* (2017) 'Covalent attachment and Pro-Pro endopeptidase (PPEP-1)-mediated release of *Clostridium difficile* cell surface proteins involved in adhesion', *Molecular Microbiology*, 105(5), pp. 663–673. doi: 10.1111/mmi.13736.

Coutinho-Silva, R. *et al.* (1999) 'P2Z/P2X 7 receptor-dependent apoptosis of dendritic cells', *American Journal of Physiology*, 276(5), pp. 1139–1147.

Coutinho-Silva, R. *et al.* (2005) 'P2X and P2Y purinergic receptors on human intestinal epithelial carcinoma cells: Effects of extracellular nucleotides on apoptosis and cell proliferation', *American Journal of Physiology - Gastrointestinal and Liver Physiology*, 288(5 51-5), pp. 1024–1035. doi: 10.1152/ajpgi.00211.2004.

Cowardin, C. A. *et al.* (2017) 'The binary toxin CDT enhances *Clostridium difficile* virulence by suppressing protective colonic eosinophilia', *Nat Microbiol*, 1(8), pp. 1–21. doi: 10.1038/nmicrobiol.2016.108.The.

Crowther, G. S. *et al.* (2014) 'Development and Validation of a Chemostat Gut Model To Study Both Planktonic and Biofilm Modes of Growth of *Clostridium difficile* and Human Microbiota', *PLoS ONE*, 9(2). doi: 10.1371/journal.pone.0088396.

Czepiel, J. *et al.* (2014) 'The role of local and systemic cytokines in patients infected with *Clostridium difficile*', *Journal of Physiology and Pharmacology*, 65(5), pp. 695–703.

D'Auria, K. M. *et al.* (2012) 'Systems analysis of the transcriptional response of human ileocecal epithelial cells to *Clostridium difficile* toxins and effects on cell cycle control', *BMC Systems Biology*, 6(2). doi: 10.1186/1752-0509-6-2.

D'Auria, K. M. *et al.* (2013) 'In Vivo Physiological and Transcriptional Profiling

Reveals Host Responses to *Clostridium difficile* Toxin A and Toxin B', *Infection and Immunity*, 81(10), pp. 3814–3824. doi: 10.1128/iai.00869-13.

Dai, Y. *et al.* (2018) 'Overexpression of MUC13, a Poor Prognostic Predictor, Promotes Cell Growth by Activating Wnt Signaling in Hepatocellular Carcinoma', *American Journal of Pathology*, 188(2), pp. 378–391. doi: 10.1016/j.ajpath.2017.10.016.

Dapa, T. *et al.* (2013) 'Multiple factors modulate biofilm formation by the anaerobic pathogen *Clostridium difficile*', *Journal of Bacteriology*, 195(3), pp. 545–555. doi: 10.1128/JB.01980-12.

Dawson, L. F. *et al.* (2012) 'Characterisation of *Clostridium difficile* Biofilm Formation, a Role for Spo0A', *PLoS ONE*, 7(12). doi: 10.1371/journal.pone.0050527.

Dawson, L. F. *et al.* (2021) 'Extracellular DNA, cell surface proteins and c-di-GMP promote biofilm formation in *Clostridioides difficile*', *Scientific Reports*, 11(1), pp. 1–21. doi: 10.1038/s41598-020-78437-5.

Deakin, L. J. *et al.* (2012) '*Clostridium difficile* spo0A gene is a persistence and transmission factor', *Infection and immunity*, pp. 147–12.

Desbaillets, I. *et al.* (1997) 'Upregulation of interleukin 8 by oxygen-deprived cells in glioblastoma suggests a role in leukocyte activation, chemotaxis, and angiogenesis', *Journal of Experimental Medicine*, 186(8), pp. 1201–1212. doi: 10.1084/jem.186.8.1201.

Dial, S. *et al.* (2008) 'Patterns of antibiotic use and risk of hospital admission because of *Clostridium difficile* infection', *Canadian Medical Association*, 179(8), pp. 767–772.

Dineen, S. S. *et al.* (2007) 'Repression of *Clostridium difficile* toxin gene

expression by CodY', *Molecular Microbiology*, 66(1), pp. 206–219. doi: 10.1111/j.1365-2958.2007.05906.x.

Dineen, S. S., McBride, S. M. and Sonenshein, A. L. (2010) 'Integration of metabolism and virulence by *Clostridium difficile* CodY', *Journal of Bacteriology*, 192(20), pp. 5350–5362. doi: 10.1128/JB.00341-10.

Dingle, T. C., Mulvey, G. L. and Armstrong, G. D. (2011) 'Mutagenic Analysis of the *Clostridium difficile* Flagellar Proteins, FliC and FliD, and Their Contribution to Virulence in Hamsters', *Infection and Immunity*, 79(10), pp. 4061–4067. doi: 10.1128/IAI.05305-11.

Donelli, G. *et al.* (2012) 'Biofilm-growing intestinal anaerobic bacteria', *Immunology and Medical Microbiology*, 65, pp. 318–325. doi: 10.1111/j.1574-695X.2012.00962.x.

Drudy, D. *et al.* (2004) 'Human antibody response to surface layer proteins in *Clostridium difficile* infection', *FEMS Immunology and Medical Microbiology*, 41(3), pp. 237–242. doi: 10.1016/j.femsim.2004.03.007.

Eckstein, B. C. *et al.* (2007) 'Reduction of *Clostridium difficile* and vancomycin-resistant Enterococcus contamination of environmental surfaces after an intervention to improve cleaning methods', *BMC Infectious Diseases*, 7, pp. 1–6. doi: 10.1186/1471-2334-7-61.

Edwards, A. N., Nawrocki, K. L. and McBride, S. M. (2014) 'Conserved oligopeptide permeases modulate sporulation initiation in *Clostridium difficile*', *Infection and Immunity*, 82(10), pp. 4276–4291. doi: 10.1128/IAI.02323-14.

Egerer, M. *et al.* (2007) 'Auto-catalytic cleavage of *Clostridium difficile* toxins A and B depends on cysteine protease activity', *Journal of Biological Chemistry*, 282(35), pp. 25314–25321. doi: 10.1074/jbc.M703062200.

Eglow, R. *et al.* (1992) 'Diminished *Clostridium difficile* toxin A sensitivity in newborn rabbit ileum is associated with decreased toxin a receptor', *Journal of Clinical Investigation*, 90(3), pp. 822–829. doi: 10.1172/JCI115957.

von Eichel-Streiber, C. and Sauerborn, M. (1990) '*Clostridium difficile* toxin A carries a C-terminal repetitive structure homologous to the carbohydrate binding region of streptococcal glycosyltransferases', *Gene*, 96(1), pp. 107–113. doi: 10.1016/0378-1119(90)90348-U.

Emery, J. G. *et al.* (1998) 'Osteoprotegerin is a receptor for the cytotoxic ligand TRAIL', *Journal of Biological Chemistry*, 273(23), pp. 14363–14367. doi: 10.1074/jbc.273.23.14363.

Engevik, M. *et al.* (2020) '*Fusobacterium nucleatum* adheres to *Clostridioides difficile* via the RadD adhesin to enhance biofilm formation in intestinal mucus', *Gastroenterology*, p. 104743. Available at: https://doi.org/10.1016/j.phrs.2020.104743.

Engevik, M. A. *et al.* (2015) 'Human *Clostridium difficile* infection: altered mucus production and composition', *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 308(6), pp. 510–524. doi: 10.1152/ajpgi.00091.2014.

Eze, P. *et al.* (2017) 'Risk factors for *Clostridium difficile* infections – an overview of the evidence base and challenges in data synthesis', *Journal of Global Health*, 7(1), pp. 1–9. doi: 10.7189/jogh.07.010417.

Farrow, M. A. *et al.* (2013) '*Clostridium difficile* toxin B-induced necrosis is mediated by the host epithelial cell NADPH oxidase complex', *PNAS*, 110(46), pp. 18674–18679. doi: 10.1073/pnas.1313658110.

El Feghaly, R. E. *et al.* (2013) 'Markers of intestinal inflammation, not bacterial burden, correlate with clinical outcomes in *Clostridium difficile* infection',

Clinical Infectious Diseases, 56(12), pp. 1713–1721. doi: 10.1093/cid/cit147.

Fekety, R. G. *et al.* (1979) 'Clindamycin-induced colitis', *The American Journal of Clinical Nutrician*, 32(1), pp. 244–250.

Fiorentini, C. *et al.* (1998) '*Clostridium difficile* toxin B induces apoptosis in intestinal cultured cells', *Infection and Immunity*, 66(6), pp. 2660–2665. doi: 10.1128/iai.66.6.2660-2665.1998.

Fletcher, J. R. *et al.* (2018) 'Shifts in the Gut Metabolome and *Clostridium difficile* Transcriptome throughout Colonization and Infection in a Mouse Model', *mSphere*, 3(2), pp. 1–18. doi: 10.1128/msphere.00089-18.

Fletcher, J. R. *et al.* (2021) '*Clostridioides difficile* exploits toxin-mediated inflammation to alter the host nutritional landscape and exclude competitors from the gut microbiota', *Nature Communications*, 12(1), p. 462.

van der Flier, L. G. and Clevers, H. (2009) 'Stem Cells, Self-Renewal, and Differentiation in the Intestinal Epithelium', *Annual Review of Physiology*, 71(1), pp. 241–260. doi: 10.1146/annurev.physiol.010908.163145.

Foulke-Abel, J. *et al.* (2020) 'Phosphodiesterase 5 (PDE5) restricts intracellular cGMP accumulation during enterotoxigenic *Escherichia coli* infection', *Gut Microbes*, 12(1). doi: 10.1080/19490976.2020.1752125.

Frädrich, C., Beer, L. A. and Gerhard, R. (2016) 'Reactive oxygen species as additional determinants for cytotoxicity of *Clostridium difficile* toxins A and B', *Toxins*, 8(1), pp. 1–12. doi: 10.3390/toxins8010025.

Francione, L. *et al.* (2009) '*Legionella pneumophila* multiplication is enhanced by chronic AMPK signalling in mitochondrially diseased Dictyostelium cells', *DMM Disease Models and Mechanisms*, 2(9–10), pp. 479–489. doi: 10.1242/dmm.003319.

Freeman, J. *et al.* (2005) 'Comparison of the efficacy of ramoplanin and vancomycin in both *in vitro* and *in vivo* models of clindamycin-induced *Clostridium difficile* infection', *Journal of Antimicrobial Chemotherapy*, 56(4), pp. 717–725. doi: 10.1093/jac/dki321.

Freeman, J. *et al.* (2007) 'Effect of metronidazole on growth and toxin production by epidemic *Clostridium difficile* PCR ribotypes 001 and 027 in a human gut model', *Journal of Antimicrobial Chemotherapy*, 60(1), pp. 83–91. doi: 10.1093/jac/dkm113.

Freeman, J., O'Neill, F. J. and Wilcox, M. H. (2003) 'Effects of cefotaxime and desacetylcefotaxime upon *Clostridium difficile* proliferation and toxin production in a triple-stage chemostat model of the human gut', *Journal of Antimicrobial Chemotherapy*, 52(1), pp. 96–102. doi: 10.1093/jac/dkg267.

Freeman, J. and Wilcox, M. H. (1999) 'Antibiotics and *Clostridium difficile*', *Microbes and Infection*, 1, pp. 377–384.

Frisbee, A. L. *et al.* (2019) 'IL-33 drives group 2 innate lymphoid cell-mediated protection during *Clostridium difficile* infection', *Nature Communications*, 10(2712), pp. 1–13. doi: 10.1038/s41467-019-10733-9.

Frost, L. R., Cheng, J. K. J. and Unnikrishnan, M. (2021) '*Clostridioides difficile* biofilms: A mechanism of persistence in the gut?', *PLoS Pathogens*, 17(3), p. e1009348.

G. Martirosian *et al.* (2004) 'Isolation of Non-Toxigenic Strains of *Clostridium difficile* from Cases of Diarrhea Among Patients Hospitalized in Hematology/Oncology Ward', *Polish Journal of Microbiology*, 53(3), pp. 197–200. Available at: www.cemera.pl.

Gadaleta, R. M. *et al.* (2011) 'Farnesoid X receptor activation inhibits

inflammation and preserves the intestinal barrier in inflammatory bowel disease', *Gut*, 60(4), pp. 463–472. doi: 10.1136/gut.2010.212159.

Garavaglia, M., Rossi, E. and Landini, P. (2012) 'The pyrimidine nucleotide biosynthetic pathway modulates production of biofilm determinants in Escherichia coli', *PLoS ONE*, 7(2), pp. 1–10. doi: 10.1371/journal.pone.0031252.

Gengenbacher, M. and Kaufmann, S. H. E. (2012) '*Mycobacterium tuberculosis*: Success through dormancy', *FEMS Microbiology Reviews*, 36(3), pp. 514–532. doi: 10.1038/jid.2014.371.

Genth, H. *et al.* (2006) 'Cellular stability of Rho-GTPases glucosylated by *Clostridium difficile* toxin B', *FEBS Letters*, 580(14), pp. 3565–3569. doi: 10.1016/j.febslet.2006.04.100.

Gerding, D. N. *et al.* (2015) 'Administration of spores of nontoxigenic *Clostridium difficile* strain M3 for prevention of recurrent *C. difficile* infection: A Randomized clinical trial', *JAMA - Journal of the American Medical Association*, 313(17), pp. 1719–1727. doi: 10.1001/jama.2015.3725.

Gerhard, R. *et al.* (2008) 'Glucosylation of Rho GTPases by *Clostridium difficile* toxin A triggers apoptosis in intestinal epithelial cells', *Journal of Medical Microbiology*, 57(6), pp. 765–770. doi: 10.1099/jmm.0.47769-0.

Gerondakis, S. *et al.* (1999) 'Genetic approaches in mice to understand Rel/NF-κB and IκB function: Transgenics and knockouts', *Oncogene*, 18(49), pp. 6888–6895. doi: 10.1038/sj.onc.1203236.

Ghose, C. *et al.* (2016) 'Immunogenicity and protective efficacy of recombinant *Clostridium difficile* flagellar protein FliC', *Emerging Microbes & Infections*, 5(1), pp. 1–10. doi: 10.1038/emi.2016.8.

Giel, J. L. *et al.* (2010) 'Metabolism of bile salts in mice influences spore

germination in *Clostridium difficile*', *PLoS ONE*, 5(1). doi: 10.1371/journal.pone.0008740.

Göke, M., Kanai, M. and Podolsky, D. K. (1998) 'Intestinal fibroblasts regulate intestinal epithelial cell proliferation via hepatocyte growth factor', *The American Physiological Society*, 1, pp. 809–818.

Gopal-Srivastava, R. and Hylemon, P. B. (1988) 'Purification and characterization of bile salt hydrolase from Clostridium perfringens.', *Journal of Lipid Research*, 29(8), pp. 1079–1085.

Gorbach, S. L. and Thadepalli, H. (1975) 'Isolation of Clostridium in Human Infections: Evaluation of 114 Cases', *Journal of Infectious Diseases*, 131(Supplement), pp. S81–S85. doi: 10.1093/infdis/131.supplement.s81.

Gravisse, J. *et al.* (2003) '*Clostridium difficile* brain empyema after prolonged intestinal carriage', *Journal of Clinical Microbiology*, 41(1), pp. 509–511. doi: 10.1128/JCM.41.1.509-511.2003.

Greco, A. *et al.* (2006) 'Carbohydrate recognition by *Clostridium difficile* toxin A', *Nature Structural and Molecular Biology*, 13(5), pp. 460–461. doi: 10.1038/nsmb1084.

Hall, Ivan, C. and O'Toole, E. (1935) 'Intestinal flora of new-born infants with a description of a new pathogenic anaerobe Bacillus difficilis', *American Journal of Diseases of Children*, 49(2), pp. 390–402.

Hamm, E. E., Voth, D. E. and Ballard, J. D. (2006) 'Identification of *Clostridium difficile* toxin B cardiotoxicity using a zebrafish embryo model of intoxication', *Proceedings of the National Academy of Sciences of the United States of America*, 103(38), pp. 14176–14181. doi: 10.1073/pnas.0604725103.

Harris, T. J. C. and Tepass, U. (2010) 'Adherens junctions: From molecules to

morphogenesis', *Nature Reviews Molecular Cell Biology*, 11(7), pp. 502–514. doi: 10.1038/nrm2927.

Hasegawa, M. *et al.* (2011) 'Nucleotide-Binding Oligomerization Domain 1 Mediates Recognition of *Clostridium difficile* and Induces Neutrophil Recruitment and Protection against the Pathogen', *The Journal of Immunology*, 186(8), pp. 4872–4880. doi: 10.4049/jimmunol.1003761.

Hasegawa, M. *et al.* (2012) 'Protective Role of Commensals against *Clostridium difficile* Infection via an IL-1β–Mediated Positive-Feedback Loop', *The Journal of Immunology*, 189(6), pp. 3085–3091. doi: 10.4049/jimmunol.1200821.

He, D. *et al.* (2002) '*Clostridium difficile* Toxin A Triggers Human Colonocyte IL-8 Release via Mitochondrial Oxygen Radical Generation', *Gastroenterology*, 122(4), pp. 1048–1057. doi: 10.1053/gast.2002.32386.

He, R. *et al.* (2017) 'Glucosyltransferase Activity of *Clostridium difficile* Toxin B Triggers Autophagy-mediated Cell Growth Arrest', *Scientific Reports*, 7(1), pp. 1–14. doi: 10.1038/s41598-017-11336-4.

Hennequin, C. *et al.* (2003) 'Identification and characterization of a fibronectin-binding protein from *Clostridium difficile*', *Microbiology*, 149(10), pp. 2779–2787. doi: 10.1099/mic.0.26145-0.

Hennequin, C., Collignon, A. and Karjalainen, T. (2001) 'Analysis of expression of GroEL (Hsp60) of *Clostridium difficile* in response to stress', *Microbial Pathogenesis*, 31(5), pp. 255–260. doi: 10.1006/mpat.2001.0468.

Hensbergen, P. J. *et al.* (2014) 'A Novel Secreted Metalloprotease (CD2830) from *Clostridium difficile* Cleaves Specific Proline Sequences in LPXTG Cell Surface Proteins', *Molecular & Cellular Proteomics*, 13(5), pp. 1231–1244. doi: 10.1074/mcp.M113.034728.

Hensbergen, P. J. *et al.* (2015) '*Clostridium difficile* secreted Pro-Pro endopeptidase PPEP-1 (ZMP1/CD2830) modulates adhesion through cleavage of the collagen binding protein CD2831', *FEBS Letters*, 589(24), pp. 3952–3958. doi: 10.1016/j.febslet.2015.10.027.

Higgins, D. and Dworkin, J. (2012) 'Recent progress in *Bacillus subtilis* sporulation', *FEMS Microbiology Reviews*, 36(1), pp. 131–148. doi: 10.1111/j.1574-6976.2011.00310.x.

Higuchi, Y. *et al.* (2015) 'Gastrointestinal fibroblasts have specialized, diverse transcriptional phenotypes: A comprehensive gene expression analysis of human fibroblasts', *PLoS ONE*, 10(6), pp. 1–19. doi: 10.1371/journal.pone.0129241.

Hinchy, E. C. *et al.* (2018) 'Mitochondria-derived ROS activate AMP-activated protein kinase (AMPK) indirectly', *Journal of Biological Chemistry*, 293(44), pp. 17208–17217. doi: 10.1074/jbc.RA118.002579.

Hippenstiel, S. *et al.* (2002) 'Rho protein inactivation induced apoptosis of cultured human endothelial cells', *American Journal of Physiology - Lung Cellular and Molecular Physiology*, 283(4 27-4), pp. 830–838. doi: 10.1152/ajplung.00467.2001.

Ho, T. D. and Ellermeier, C. D. (2015) 'Ferric uptake regulator fur control of putative iron acquisition systems in *Clostridium difficile*', *Journal of Bacteriology*, 197(18), pp. 2930–2940. doi: 10.1128/JB.00098-15.

Hoffmann, E. *et al.* (2002) 'Multiple control of interleukin-8 gene expression.', *Journal of Leukocyte Biology*, 72(5), pp. 847–55. doi: 10.1189/jlb.1106655.

Holdeman, L. V., Good, I. J. and Moore, W. E. C. (1976) 'Human fecal flora: variation in bacterial composition within individuals and a possible effect of

emotional stress', *Applied and Environmental Microbiology*, 31(3), pp. 359–375. doi: 10.1128/aem.31.3.359-375.1976.

Houston, S. *et al.* (2012) 'Activation and proteolytic activity of the *Treponema pallidum* metalloprotease, pallilysin', *PLoS Pathogens*, 8(7), p. 17. doi: 10.1371/journal.ppat.1002822.

In, J. *et al.* (2016) 'Enterohemorrhagic *Escherichia coli* Reduces Mucus and Intermicrovillar Bridges in Human Stem Cell-Derived Colonoids', *Cmgh*, 2(1), pp. 48-62.e3. doi: 10.1016/j.jcmgh.2015.10.001.

Inami, A., Kiyono, H. and Kurashima, Y. (2018) 'ATP as a Pathophysiologic Mediator of Bacteria-Host Crosstalk in the Gastrointestinal Tract', *International Journal of Molecular Sciences*, 19(8), pp. 7–9. doi: 10.3390/ijms19082371.

Ishida, Y. *et al.* (2004) 'Essential Involvement of IFN-γ in *Clostridium difficile* Toxin A-Induced Enteritis', *The Journal of Immunology*, 172(5), pp. 3018–3025. doi: 10.4049/jimmunol.172.5.3018.

Ishigame, H. *et al.* (2009) 'Differential Roles of Interleukin-17A and -17F in Host Defense against Mucoepithelial Bacterial Infection and Allergic Responses', *Immunity*, 30(1), pp. 108–119. doi: 10.1016/j.immuni.2008.11.009.

Islam, J. *et al.* (2014) 'The role of the humoral immune response to <i>Clostridium difficile<i> toxins A and B in susceptibility to Clostridium difficile Infection: a case-control study', *Anaerobe*, 27, pp. 82–86. doi: 10.1038/jid.2014.371.

Jackson, S. *et al.* (2006) 'Analysis of Proline Reduction in the Nosocomial Pathogen *Clostridium difficile*', *Journal of Bacteriology*, 188(24), pp. 8487–8495. doi: 10.1128/JB.01370-06.

Jafari, N. V *et al.* (2013) '*Clostridium difficile* Modulates Host Innate Immunity

via Toxin-Independent and Dependent Mechanism (s)', *PloS one*, 8(7), pp. 1–10. doi: 10.1371/journal.pone.0069846.

Jafari, N. V *et al.* (2016) '*Clostridium difficile* -mediated effects on human intestinal epithelia : Modelling host-pathogen interactions in a vertical diffusion chamber', *Anaerobe*, 37, pp. 96–102.

Janoir, C. *et al.* (2007) 'Cwp84, a surface-associated protein of *Clostridium difficile*, is a cysteine protease with degrading activity on extracellular matrix proteins', *Journal of bacteriology*, 189(20), pp. 7174–7180.

Janoir, C. *et al.* (2013) 'Adaptive strategies and pathogenesis of *Clostridium difficile* from *In vivo* transcriptomics', *Infection and Immunity*, 81(10), pp. 3757–3769. doi: 10.1128/IAI.00515-13.

Janvilisri, T., Scaria, J. and Chang, Y.-F. (2015) 'Transcriptional profiling of *Clostridium difficile* and Caco-2 cells during infection', *NIH Publis Access*, 85(0 1), pp. 1–27. doi: 10.1016/j.neuroimage.2013.08.045.The.

Jarchum, I. *et al.* (2011) 'Toll-like receptor 5 stimulation protects mice from acute *Clostridium difficile* colitis', *Infection and Immunity*, 79(4), pp. 1498–1503. doi: 10.1128/IAI.01196-10.

Jarchum, I. *et al.* (2012) 'Critical role for myd88-Mediated Neutrophil recruitment during *Clostridium difficile* colitis', *Infection and Immunity*, 80(9), pp. 2989–2996. doi: 10.1128/IAI.00448-12.

Jefferson, K. K., Smith, M. F. and David, A. (1999) 'Roles of Intracellular Calcium and NF-κB in the *Clostridium difficile* Toxin A-Induced Up-Regulation and Secretion of IL-8 from Human Monocytes', *Journal of Immunology*, 163, pp. 5183–5191.

Jiang, Z. D. *et al.* (2007) 'Association of Interleukin-8 Polymorphism and

Immunoglobulin G Anti-Toxin A in Patients With *Clostridium difficile*-Associated Diarrhea', *Clinical Gastroenterology and Hepatology*, 5(8), pp. 964–968. doi: 10.1016/j.cgh.2007.04.018.

Johal, S. S. *et al.* (2004) 'Colonic IgA producing cells and macrophages are reduced in recurrent and non-recurrent *Clostridiun difficile* associated diarrhoea', *Journal of Clinical Pathology*, 57(9), pp. 973–979. doi: 10.1136/jcp.2003.015875.

Johnson, S., Gerding, D. N. and Janoff, E. N. (1992) 'Systemic and Mucosal Antibody Responses to Toxin A in Patients Infected with *Clostridium difficile*', *Journal of Infectious Diseases*, 166(6), pp. 1287–1294. doi: 10.1093/infdis/166.6.1287.

Kajiwara, C. *et al.* (2018) 'Metformin Mediates Protection against *Legionella* Pneumonia through Activation of AMPK and Mitochondrial Reactive Oxygen Species', *The Journal of Immunology*, 200(2), pp. 623–631. doi: 10.4049/jimmunol.1700474.

Karban, A. S. *et al.* (2004) 'Functional annotation of a novel NFKB1 promoter polymorphism that increases risk for ulcerative colitis', *Human Molecular Genetics*, 13(1), pp. 35–45. doi: 10.1093/hmg/ddh008.

Karjalainen, T. *et al.* (2001) 'Molecular and Genomic Analysis of Genes Encoding Surface- Anchored Proteins from *Clostridium difficile*', *Infection and Immunity*, 69(5), pp. 3442–3446. doi: 10.1128/IAI.69.5.3442.

Karlsson, S., Burman, L. G. and Akerlund, T. (1999) 'Suppression of toxin production in *Clostridium difficile* VPI 10463 by amino acids', *Microbiology*, 145, pp. 1683–1693.

Katchar, K. *et al.* (2007) 'Association Between IgG2 and IgG3 Subclass

Responses to Toxin A and Recurrent *Clostridium difficile*-Associated Disease', *Clinical Gastroenterology and Hepatology*, 5(6), pp. 707–713. doi: 10.1016/j.cgh.2007.02.025.

Keighley, M. R. B. *et al.* (1978) 'Randomised controlled trial of vancomycin for pseudomembranous colitis and postoperative diarrhoea', *British Medical Journal*, 2(6153), pp. 1667–1669. doi: 10.1136/bmj.2.6153.1667.

Kelly, C. P. *et al.* (1992) 'Human Colonic Aspirates Containing Immunoglobulin A Antibody to *Clostridium dificile* Toxin A Inhibit Toxin A-Receptor Binding', *Gastroenterology*, 102, pp. 35–40.

Kelly, C. P. and Kyne, L. (2011) 'The host immune response to *Clostridium difficile*', *Journal of Medical Microbiology*, 60(8), pp. 1070–1079. doi: 10.1099/jmm.0.030015-0.

Kelly, C. R. *et al.* (2015) 'Update on FMT 2015: Indications, Methodologies, Mechanisms and Outlook', *Gastroenterology*, 149(1), pp. 223–237. doi: 10.1097/CCM.0b013e31823da96d.Hydrogen.

Kim, H. *et al.* (2005) '*Clostridium difficile* toxin A regulates inducible cyclooxygenase-2 and prostaglandin E2 synthesis in colonocytes via reactive oxygen species and activation of p38 MAPK', *Journal of Biological Chemistry*, 280(22), pp. 21237–21245. doi: 10.1074/jbc.M413842200.

Kim, J. M. *et al.* (2006) 'NF-κB activation pathway is essential for the chemokine expression in intestinal epithelial cells stimulated with *Clostridium difficile* toxin A', *Scandinavian Journal of Immunology*, 63(6), pp. 453–460. doi: 10.1111/j.1365-3083.2006.001756.x.

Klein, E. A. and Assoian, R. K. (2008) 'Transcriptional regulation of the cyclin D1 gene at a glance', *Journal of Cell Science*, 121(23), pp. 3853–3857. doi:

10.1016/j.physbeh.2017.03.040.

Klychnikov, O. I. *et al.* (2018) 'Discovery of a new Pro-Pro endopeptidase, PPEP-2, provides mechanistic insights into the differences in substrate specificity within the PPEP family', *Journal of Biological Chemistry*, 293(28), pp. 11154–11165. doi: 10.1074/jbc.RA118.003244.

Knippel, R. J. *et al.* (2018) 'Heme sensing and detoxification by HatRT contributes to pathogenesis during *Clostridium difficile* infection', *PLoS Pathogens*, 14(12), pp. 1–21. doi: 10.1371/journal.ppat.1007486.

Knippel, R. J. *et al.* (2020) 'Article *Clostridioides difficile* Senses and Hijacks Host Heme for Incorporation into an Oxidative Stress Defense System', *Cell Host and Microbe*, 28, pp. 1–11.

Kociolek, L. K., Polage, C. R. and Riley, T. V (2016) '*Clostridium difficile* — Diagnostic and Clinical Challenges Q & A', *Clinical Chemistry*, 62(2), pp. 310–314.

Komiya, Y. and Habas, R. (2008) 'Wnt signal transduction pathways', *Organogenesis*, 4(2), pp. 68–75. doi: 10.4161/org.4.2.5851.

Krivan, H. C. *et al.* (1986) 'Cell surface binding site for *Clostridium difficile* enterotoxin: Evidence for a glycoconjugate containing the sequence Ga1α1-3Galβ1-4GlcNAc', *Infection and Immunity*, 53(3), pp. 573–581. doi: 10.1128/iai.53.3.573-581.1986.

Kuijper, E. J. *et al.* (2008) 'Update of *Clostridium difficile* infection due to PCR ribotype 027 in europe, 2008', *Eurosurveillance*, 13(7–9), pp. 1–7.

Kulthong, K. *et al.* (2021) 'Transcriptome comparisons of *in vitro* intestinal epithelia grown under static and microfluidic gut-on-chip conditions with *in vivo* human epithelia', *Scientific Reports*, 11(1), pp. 1–13. doi: 10.1038/s41598-

021-82853-6.

Kushner, I. and Rzewnicki, D. L. (1994) 'The Acute Phase Response: General aspects', *Bailliere's Clinical Rheumatology*, 8(3), pp. 513–530. doi: 10.1016/S0950-3579(05)80113-X.

Kyne, L. *et al.* (2001) 'Association between antibody response to toxin A and protection against recurrent *Clostridium difficile* diarrhoea', *Lancet*, 357(9251), pp. 189–193. doi: 10.1016/S0140-6736(00)03592-3.

LaFrance, M. E. *et al.* (2015) 'Identification of an epithelial cell receptor responsible for *Clostridium difficile* TcdB-induced cytotoxicity', *Proceedings of the National Academy of Sciences of the United States of America*, 112(22), pp. 7073–7078. doi: 10.1073/pnas.1500791112.

LaMonte, G. M. *et al.* (2019) 'Dual RNA-seq identifies human mucosal immunity protein Mucin-13 as a hallmark of Plasmodium exoerythrocytic infection', *Nature Communications*, 10(1), pp. 1–13. doi: 10.1038/s41467-019-08349-0.

Larabee, J. L. *et al.* (2008) '*Bacillus anthracis* edema toxin activates nuclear glycogen synthase kinase 3β', *Infection and Immunity*, 76(11), pp. 4895–4904. doi: 10.1128/IAI.00889-08.

Larson, H. E. *et al.* (1978) '*Clostridium difficile* and the aetology of pseudomembranous colitis', *The Lancet*, 311(8073), pp. 1063–1066.

Lawley, T. D. *et al.* (2009) 'Antibiotic treatment of *Clostridium difficile* carrier mice triggers a supershedder state, spore-mediated transmission, and severe disease in immunocompromised hosts', *Infection and Immunity*, 77(9), pp. 3661–3669. doi: 10.1128/IAI.00558-09.

Lee, J. S. *et al.* (2015) 'Interleukin-23-Independent IL-17 Production Regulates Intestinal Epithelial Permeability', *Immunity*, 43(4), pp. 727–738. doi:

10.1016/j.immuni.2015.09.003.

Leffler, D. A. and Lamont, J. T. (2015) '*Clostridium difficile* Infection', *New England Journal of Medicine*, 372(16), pp. 1539–1548. doi: 10.1056/NEJMra1403772.

Lessa, F. C. *et al.* (2015) 'Burden of *Clostridium difficile* infection in the United States', *New England Journal of Medicine*, 372(9), pp. 825–834.

Leuzzi, R. *et al.* (2013) 'Protective efficacy induced by recombinant *Clostridium difficile* toxin fragments', *Infection and Immunity*, 81(8), pp. 2851–2860. doi: 10.1128/IAI.01341-12.

Li, N. *et al.* (2013) 'Roles of fibroblast growth factor-inducible 14 in hepatocellular carcinoma', *Asian Pacific Journal of Cancer Prevention*, 14(6), pp. 3509–3514. doi: 10.7314/APJCP.2013.14.6.3509.

Libby, J. M. and Wilkins, T. D. (1982) 'Production of antitoxins to two toxins of *Clostridium difficile* and immunological comparison of the toxins by cross-neutralization studies', *Infection and Immunity*, 35(1), pp. 374–376. doi: 10.1128/iai.35.1.374-376.1982.

Lima, B. B. *et al.* (2014) '*Clostridium difficile* toxin A attenuates Wnt/β-catenin signaling in intestinal epithelial cells', *Infection and Immunity*, 82(7), pp. 2680–2687. doi: 10.1128/IAI.00567-13.

Limsrivilai, J. *et al.* (2018) 'Systemic Inflammatory Responses in Ulcerative Colitis Patients and *Clostridium difficile* Infection', *Digestive Diseases and Sciences*, 63(7), pp. 1801–1810. doi: 10.1007/s10620-018-5044-1.

Linevsky, J. K. *et al.* (1997a) 'IL-8 release and neutrophil activation by *Clostridium difficile* toxin-exposed human monocytes', *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 1(1), pp. 1333–1340.

Linevsky, J. K. *et al.* (1997b) 'IL-8 release and neutrophil activation by Clostridium difficile toxin-exposed human monocytes', *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 1(1), pp. 1333–1340.

Liu, C. *et al.* (2014) 'Gastric de novo Muc13 expression and spasmolytic polypeptide-expressing metaplasia during *Helicobacter heilmannii* infection', *Infection and Immunity*, 82(8), pp. 3227–3239. doi: 10.1128/IAI.01867-14.

Liu, Y. H. *et al.* (2018) 'The ATP-P2X7 signaling axis is an essential sentinel for intracellular *Clostridium difficile* pathogen-induced inflammasome activation', *Frontiers in Cellular and Infection Microbiology*, 8(MAR), pp. 1–14. doi: 10.3389/fcimb.2018.00084.

Löffler, M. *et al.* (2005) 'Pyrimidine pathways in health and disease', *Trends in Molecular Medicine*, 11(9), pp. 430–437. doi: 10.1016/j.molmed.2005.07.003.

Lopez, C. A. *et al.* (2019) 'The Immune Protein Calprotectin Impacts *Clostridioides difficile* Metabolism through Zinc Limitation', *mBio*, 10(6), pp. e02289-19.

Lu, T., Li, Y. and Chen, T. (2013) 'Techniques for fabrication and construction of three-dimensional scaff | IJN', *International Journal of Nanomedicine*, 8, pp. 337–350.

Lyras, D. *et al.* (2009) 'Toxin B is essential for virulence of *Clostridium difficile*', *Nature*, 458(7242), pp. 1176–1179. doi: 10.1016/j.physbeh.2017.03.040.

Macfarlane, G. T., Macfarlane, S. and Gibson, G. R. (1998) 'Validation of a three-stage compound continuous culture system for investigating the effect of retention time on the ecology and metabolism of bacteria in the human colon', *Microbial Ecology*, 35(2), pp. 180–187. doi: 10.1007/s002489900072.

Madan, R. *et al.* (2014) 'Role of leptin-mediated colonic inflammation in defense

against *Clostridium difficile* colitis', *Infection and Immunity*, 82(1), pp. 341–349. doi: 10.1128/IAI.00972-13.

Maher, D. M. *et al.* (2011) 'Mucin 13: Structure, Function, and Potential Roles in Cancer Pathogenesis', *Mol Cancer Res*, 9(5), pp. 531–537. doi: 10.1158/1541-7786.MCR-10-0443.Mucin.

Major, M. B. *et al.* (2007) 'Wilms Tumor Suppressor WTX Negatively Regulates WNT/b-Catenin Signaling', *Science*, 316, pp. 1043–1046. doi: 10.1017/CBO9781107415324.004.

Makino, Y. *et al.* (2010) 'Apolipoprotein A-IV is a candidate target molecule for the treatment of seasonal allergic rhinitis', *Journal of Allergy and Clinical Immunology*, 126(6), pp. 1163–1169. doi: 10.1016/j.jaci.2010.06.031.

Mantis, N. J., Rol, N. and Corthésy, B. (2011) 'Secretory IgA's complex roles in immunity and mucosal homeostasis in the gut', *Mucosal Immunology*, 4(6), pp. 603–611. doi: 10.1038/mi.2011.41.

Mavromatis, C. H. *et al.* (2015) 'The co-transcriptome of uropathogenic *Escherichia coli*-infected mouse macrophages reveals new insights into host-pathogen interactions', *Cellular Microbiology*, 17(5), pp. 730–746. doi: 10.1111/cmi.12397.

Mazuski, J. E. *et al.* (2000) '*Clostridium difficile* toxins influence hepatocyte protein synthesis through the interleukin 1 receptor', *Archives of Surgery*, 135(10), pp. 1206–1211. doi: 10.1001/archsurg.135.10.1206.

McDermott, A. J. *et al.* (2016) 'Interleukin-23 (IL-23), independent of IL-17 and IL-22, drives neutrophil recruitment and innate inflammation during *Clostridium difficile* colitis in mice', *Immunology*, 147(1), pp. 114–124. doi: 10.1111/imm.12545.

McGlone, S. M. *et al.* (2012) 'The economic burden of *Clostridium difficile*', *Clinical Microbiology and Infection*, 18(3), pp. 282–289.

McGuckin, M. A. *et al.* (2011) 'Mucin dynamics and enteric pathogens', *Nature Reviews Microbiology*, 9(4), pp. 265–278. doi: 10.1038/nrmicro2538.

McKee, R. W. *et al.* (2013) 'The second messenger cyclic Di-GMP regulates *Clostridium difficile* toxin production by controlling expression of sigD', *Journal of Bacteriology*, 195(22), pp. 5174–5185. doi: 10.1128/JB.00501-13.

McKee, R. W. *et al.* (2018) 'Type IV Pili Promote *Clostridium difficile* Adherence and Persistence in a Mouse Model of Infection', *Infection and Immunity*, 86(5), pp. 1–13. doi: 10.1128/IAI.00943-17.

Moehle, C. *et al.* (2006) 'Aberrant intestinal expression and allelic variants of mucin genes associated with inflammatory bowel disease', *Journal of Molecular Medicine*, 84(12), pp. 1055–1066. doi: 10.1007/s00109-006-0100-2.

Morris, G. E. *et al.* (2014) 'A novel electrospun biphasic scaffold provides optimal three-dimensional topography for *in vitro* co-culture of airway epithelial and fibroblast cells', *Biofabrication*, 6(3), p. 035014. doi: 10.1088/1758-5082/6/3/035014.

Mosesson, M. W. (2012) 'Fibrinogen and fibrin: Structure and functional aspects', *Thrombin: Function and Pathophysiology*, 3, pp. 1894–1904.

Moura, I. B. *et al.* (2019) 'Omadacycline Gut Microbiome Exposure Does Not Induce *Clostridium difficile* Proliferation or Toxin Production in a Model That Simulates the Proximal, Medial, and Distal Human Colon', *Antimicrobial Agents and Chemotherapy*, 63(2), pp. 1–11. doi: 10.1128/AAC.01581-18.

Muller, E. L., Pitt, H. A. and George, W. L. (1987) 'Prairie Dog Model for Antimicrobial Agent-Induced *Clostridium difficile* Diarrhoea', *Infection and*

Immunity, 55(1), pp. 198–200.

Mulligan, M. E. *et al.* (1993) 'Elevated levels of serum immunoglobulins in asymptomatic carriers of *Clostridium difficile*', *Clinical Infectious Diseases*, 16, pp. S239–S244. doi: 10.1093/clinids/16.Supplement_4.S239.

Muñoz-Miralles, J. *et al.* (2018) 'Indomethacin increases severity of *Clostridium difficile* infection in mouse model', *Future Microbiology*, 13(11), pp. 1271–1281. doi: 10.2217/fmb-2017-0311.

Na, X. *et al.* (2008) 'gp96 is a human colonocyte plasma membrane binding protein for *Clostridium difficile* toxin A', *Infection and Immunity*, 76(7), pp. 2862–2871. doi: 10.1128/IAI.00326-08.

Nakagawa, T. *et al.* (2016) 'Endogenous il-17 as a factor determining the severity of *Clostridium difficile* infection in mice', *Journal of Medical Microbiology*, 65(8), pp. 821–827. doi: 10.1099/jmm.0.000273.

Navalkele, B. D. and Chopra, T. (2018) 'Bezlotoxumab: An emerging monoclonal antibody therapy for prevention of recurrent *Clostridium difficile* infection', *Biologics: Targets and Therapy*, 12, pp. 11–21. doi: 10.2147/BTT.S127099.

Nawrocki, K. L. *et al.* (2016) 'CodY-dependent regulation of sporulation in *Clostridium difficile*', *Journal of Bacteriology*, 198(15), pp. 2113–2130. doi: 10.1128/JB.00220-16.

Ng, J. *et al.* (2010) '*Clostridium difficile* toxin-induced inflammation and intestinal injury are mediated by the inflammasome', *Gastroenterology*, 139(2), pp. 542–552. doi: 10.1053/j.gastro.2010.04.005.

Nottrott, S. *et al.* (2007) '*Clostridium difficile* toxin A-induced apoptosis is p53-independent but depends on glucosylation of Rho GTPases', *Apoptosis*, 12(8), pp. 1443–1453. doi: 10.1007/s10495-007-0074-8.

Nusrat, A. *et al.* (2001) '*Clostridium difficile* toxins disrupt epithelial barrier function by altering membrane microdomain localization of tight junction proteins', *Infection and Immunity*, 69(3), pp. 1329–1336. doi: 10.1128/IAI.69.3.1329-1336.2001.

O'Brien, J. B. *et al.* (2005) 'Passive immunisation of hamsters against *Clostridium difficile* infection using antibodies to surface layer proteins', *FEMS Microbiology Letters*, 246(2), pp. 199–205. doi: 10.1016/j.femsle.2005.04.005.

O'Connor, W. *et al.* (2009) 'A protective function for interleukin 17A in T cell–mediated intestinal inflammation', *Nature Immunology*, 10(6), pp. 603–609. doi: 10.1038/ni.1736.A.

Ogawa, A. *et al.* (2004) 'Neutralization of interleukin-17 aggravates dextran sulfate sodium-induced colitis in mice', *Clinical Immunology*, 110(1), pp. 55–62. doi: 10.1016/j.clim.2003.09.013.

Oliveira Paiva, A. M. *et al.* (2016) 'The Signal Sequence of the Abundant Extracellular Metalloprotease PPEP-1 Can Be Used to Secrete Synthetic Reporter Proteins in *Clostridium difficile*', *ACS Synthetic Biology*, 5(12), pp. 1376–1382. doi: 10.1021/acssynbio.6b00104.

Olling, A. *et al.* (2011) 'The repetitive oligopeptide sequences modulate cytopathic potency but are not crucial for cellular uptake of *Clostridium difficile* toxin A', *PLoS ONE*, 6(3). doi: 10.1371/journal.pone.0017623.

Onderdonk, A. B., Lowe, B. R. and Bartlett, J. G. (1979) 'Effect of environmental stress on *Clostridium difficile* toxin levels during continuous cultivation', *Applied and Environmental Microbiology*, 38(4), pp. 637–641. doi: 10.1128/aem.38.4.637-641.1979.

Ostos, M. A. *et al.* (2001) 'Antioxidative and antiatherosclerotic effects of human

apolipoprotein A-IV in apolipoprotein E-deficient mice', *Arteriosclerosis, Thrombosis, and Vascular Biology*, 21(6), pp. 1023–1028. doi: 10.1161/01.ATV.21.6.1023.

Ottlinger, M. E. and Lin, S. (1988) '*Clostridium difficile* toxin B induces reorganization of actin, vinculin, and talin in cultured cells', *Experimental Cell Research*, 174(1), pp. 215–229. doi: 10.1016/0014-4827(88)90156-5.

Pantaléon, V. *et al.* (2015) 'The *Clostridium difficile* protease Cwp84 Modulates both biofilm formation and cell-surface properties', *PLoS ONE*, 10(4), pp. 1–20. doi: 10.1371/journal.pone.0124971.

Pantosti, A. *et al.* (1989) 'Immunoblot analysis of serum immunoglobulin G response to surface proteins of *Clostridium difficile* in patients with antibiotic-associated diarrhea', *Journal of Clinical Microbiology*, 27(11), pp. 2594–2597. doi: 10.1128/jcm.27.11.2594-2597.1989.

Papatheodoroua, P. *et al.* (2011) 'Lipolysis-stimulated lipoprotein receptor (LSR) is the host receptor for the binary toxin *Clostridium difficile* transferase (CDT) Panagiotis', *PNAS*, 108(39), pp. 16422–16427. doi: 10.1073/pnas.1109772108.

Paredes-Sabja, D. and Sarker, M. R. (2012) 'Adherence of *Clostridium difficile* spores to Caco-2 cells in culture', *Journal of Medical Microbiology*, 61(PART 9), pp. 1208–1218. doi: 10.1099/jmm.0.043687-0.

Paredes-Sabja, D., Shen, A. and Sorg, J. A. (2014) '*Clostridium difficile* spore biology: Sporulation, germination, and spore structural proteins', *Trends in Microbiology*, 22(7), pp. 406–416. doi: 10.1016/j.tim.2014.04.003.

Patel, J. C. and Galán, J. E. (2008) 'Investigating the Function of Rho Family GTPases during *Salmonella*/Host Cell Interactions', *Methods Enzymology*,

439(1), pp. 145–158. doi: 10.1016/S0076-6879(07)00411-9.Investigating.

Patil, R. and Blankenship, L. (2013) 'Proton Pump Inhibitors and *Clostridium difficile* Infection : Are We Propagating an Already Rapidly Growing Healthcare Problem?', *Gastroenterology Research*, 6(5), pp. 171–173.

Péchiné, S. *et al.* (2005) 'Immunological properties of surface proteins of *Clostridium difficile*', *Journal of Medical Microbiology*, 54(2), pp. 193–196. doi: 10.1099/jmm.0.45800-0.

Pechine, S. and Collignon, A. (2016) 'Immune responses induced by *Clostridium difficile*', *Anaerobe*, 41, pp. 68–78. doi: 10.1016/j.anaerobe.2016.04.014.

Peltier, J. *et al.* (2015) 'Cyclic-di-GMP regulates production of sortase substrates of *Clostridium difficile* and their surface exposure through ZmpI protease-mediated cleavage', *Journal of Biological Chemistry*, 290(40), pp. 24453–24469. doi: 10.1074/jbc.M115.665091.

Pepin, J. *et al.* (2005) 'Emergence of Fluoroquinolones as the Predominant Risk Factor for *Clostridium difficile* – Associated Diarrhea : A Cohort Study during an Epidemic in Quebec', *Clinical Infectious Diseases*, 41, pp. 1254–60.

Pereira, F. C. *et al.* (2013) 'The Spore Differentiation Pathway in the Enteric Pathogen *Clostridium difficile*', *PLoS Genetics*, 9(10). doi: 10.1371/journal.pgen.1003782.

Perelle, S. *et al.* (1997) 'Production of a complete binary toxin (actin-specific ADP- ribosyltransferase) by *Clostridium difficile* CD196', *Infection and Immunity*, 65(4), pp. 1402–1407. doi: 10.1128/iai.65.4.1402-1407.1997.

Pettit, L. J. *et al.* (2014) 'Functional genomics reveals that *Clostridium difficile* Spo0A coordinates sporulation, virulence and metabolism', *BMC Genomics*, 15(1), pp. 1–15. doi: 10.1186/1471-2164-15-160.

Polumuri, S. K. *et al.* (2012) 'Transcriptional Regulation of Murine IL-33 by TLR and Non-TLR Agonists', *The Journal of Immunology*, 189(1), pp. 50–60. doi: 10.4049/jimmunol.1003554.

Popoff, M. R. *et al.* (1988) 'Actin-specific ADP-ribosyltransferase produced by a *Clostridium difficile* strain', *Infection and Immunity*, 56(9), pp. 2299–2306. doi: 10.1128/iai.56.9.2299-2306.1988.

Pothoulakis, C. *et al.* (1993) 'Ketotifen inhibits *Clostridium difficile* toxin A-induced enteritis in rat ileum', *Gastroenterology*, 105(3), pp. 701–707. doi: 10.1016/0016-5085(93)90886-H.

Pothoulakis, C. (2000) 'Effects of *Clostridium difficile* Toxins on Epithelial Cell Barrier', *Annals New York Academy of Sciences*, 915, pp. 347–356.

Prevot, A. (1938) 'Etudes de systematic bacterienne', *Annual Institute Pasteur*, 61, pp. 72–91.

Prodromou, C. *et al.* (1997) 'Identification and structural characterization of the ATP/ADP-binding site in the Hsp90 molecular chaperone', *Cell*, 90(1), pp. 65–75. doi: 10.1016/S0092-8674(00)80314-1.

Proietti, M. *et al.* (2019) 'ATP released by intestinal bacteria limits the generation of protective IgA against enteropathogens', *Nature Communications*, 10(1), pp. 1–11. doi: 10.1038/s41467-018-08156-z.

Pron, B. *et al.* (1995) 'Chronic septic arthritis and osteomyelitis in a prosthetic knee joint due to *Clostridium difficile*', *European Journal of Clinical Microbiology and Infectious Diseases*, 14, pp. 599–601.

Purcell, E. B. *et al.* (2012) 'Cyclic diguanylate inversely regulates motility and aggregation in *Clostridium difficile*', *Journal of Bacteriology*, 194(13), pp. 3307–3316. doi: 10.1128/JB.00100-12.

van Putten, J. P. M. and Strijbis, K. (2017) 'Transmembrane Mucins: Signaling Receptors at the Intersection of Inflammation and Cancer', *Journal of Innate Immunity*, 9(3), pp. 281–299. doi: 10.1159/000453594.

Ranganathan, S. *et al.* (2019) 'Evaluating *Shigella flexneri* pathogenesis in the human enteroid model', *Infection and Immunity*, 87(4), pp. 1–14. doi: 10.1128/IAI.00740-18.

Rao, K. *et al.* (2014) 'The systemic inflammatory response to *Clostridium difficile* infection', *PLoS ONE*, 9(3). doi: 10.1371/journal.pone.0092578.

Rao, K. *et al.* (2015) '*Clostridium difficile* Ribotype 027: Relationship to Age, Detectability of Toxins A or B in Stool with Rapid Testing, Severe Infection, and Mortality', *Clinical Infectious Diseases*, 61(2), pp. 233–241. doi: 10.1093/cid/civ254.

Recalde, D. *et al.* (2004) 'Human Apolipoprotein A-IV Reduces Secretion of Proinflammatory Cytokines and Atherosclerotic Effects of a Chronic Infection Mimicked by Lipopolysaccharide', *Arteriosclerosis, Thrombosis, and Vascular Biology*, 24(4), pp. 756–761. doi: 10.1161/01.ATV.0000119353.03690.22.

Rienksma, R. A. *et al.* (2015) 'Comprehensive insights into transcriptional adaptation of intracellular mycobacteria by microbe-enriched dual RNA sequencing', *BMC genomics*, 16(1), p. 34.

Riley, T. V. (1998) '*Clostridium difficile* : A pathogen of the nineties', *European Journal of Clinical Microbiology and Infectious Diseases*, 17(3), pp. 137–141. doi: 10.1007/s100960050037.

Robertson, J. *et al.* (2010) 'Peptidoglycan derived from *Staphylococcus epidermidis* induces Connexin43 hemichannel activity with consequences on the innate immune response in endothelial cells', *Biochemical Journal*, 432(1), pp.

133–143. doi: 10.1042/BJ20091753.

Rodriguez, C. *et al.* (2016) '*Clostridium difficile* infection: Early history, diagnosis and molecular strain typing methods', *Microbial Pathogenesis*, 97, pp. 59–78. doi: 10.1016/j.micpath.2016.05.018.

Rolfe, R. D. and Iaconis, J. P. (1983) 'Intestinal colonization of infant hamsters with *Clostridium difficile*', *Infection and Immunity*, 42(2), pp. 480–486. doi: 10.1128/iai.42.2.480-486.1983.

Rolfe, R. D. and Song, W. (1995) 'Immunoglobulin and non-immunoglobulin components of human milk inhibit *Clostridium difficile* toxin A-receptor binding', *Journal of Medical Microbiology*, 42(1), pp. 10–19. doi: 10.1099/00222615-42-1-10.

Rosa, R., Donskey, C. J. and Munoz-Price, L. S. (2018) 'The Intersection Between Colonization Resistance, Antimicrobial Stewardship, and *Clostridium difficile*', *Current Infectious Disease Reports*, 20(8), pp. 18–23. doi: 10.1007/s11908-018-0631-z.

Roula, D. *et al.* (2020) 'Apolipoprotein A-IV acts as an endogenous anti-inflammatory protein and is reduced in treatment-naïve allergic patients and allergen-challenged mice', *Allergy: European Journal of Allergy and Clinical Immunology*, 75(2), pp. 392–402. doi: 10.1111/all.14022.

Rousseau, C. *et al.* (2012) '*Clostridium difficile* carriage in healthy infants in the community: A potential reservoir for pathogenic strains', *Clinical Infectious Diseases*, 55(9), pp. 1209–1215. doi: 10.1093/cid/cis637.

Rubino, J. T. *et al.* (2016) 'Structural characterization of zinc-bound Zmp1, a zinc-dependent metalloprotease secreted by *Clostridium difficile*', *Journal of Biological Inorganic Chemistry*, 21(2), pp. 185–196. doi: 10.1007/s00775-015-

1319-6.

Ruiz de Morales, J. M. G. *et al.* (2020) 'Critical role of interleukin (IL)-17 in inflammatory and immune disorders: An updated review of the evidence focusing in controversies', *Autoimmunity Reviews*, 19(1). doi: 10.1016/j.autrev.2019.102429.

Rupnik, M., Wilcox, M. H. and Gerding, D. N. (2009) '*Clostridium difficile* infection: New developments in epidemiology and pathogenesis', *Nature Reviews Microbiology*. Nature Publishing Group, 7(7), pp. 526–536. doi: 10.1038/nrmicro2164.

Ryan, A. *et al.* (2011) 'A role for TLR4 in *Clostridium difficile* infection and the recognition of surface layer proteins', *PLoS Pathogens*, 7(6). doi: 10.1371/journal.ppat.1002076.

Sambol, S. P. *et al.* (1999) 'Colonization for the prevention of *Clostridium difficile* disease', *Anaerobe*, 5(3–4), pp. 195–199. doi: 10.1006/anae.1999.0291.

Sanderson, C. J. (1992) 'Interleukin-5, Eosinophils, and Disease', *Blood*, 79(12), pp. 3101–3109.

Savidge, T. C. *et al.* (2003) '*Clostridium difficile* Toxin B Is an Inflammatory Enterotoxinin Human Intestine', *Gastroenterology*, 125(2), pp. 413–420. doi: 10.1016/S0016-5085(03)00902-8.

Savio, L. E. B. *et al.* (2018) 'The P2X7 receptor in inflammatory diseases: Angel or demon?', *Frontiers in Pharmacology*, 9(FEB). doi: 10.3389/fphar.2018.00052.

Scaria, J. *et al.* (2011) '*Clostridium difficile* transcriptome analysis using pig ligated loop model reveals modulation of pathways not modulated *in vitro*', *Journal of Infectious Diseases*, 203(11), pp. 1613–1620. doi: 10.1093/infdis/jir112.

Schacherl, M. *et al.* (2015) 'Structural Basis of Proline-Proline Peptide Bond Specificity of the Metalloprotease Zmp1 Implicated in Motility of *Clostridium difficile*', *Structure*, 23(9), pp. 1632–1642. doi: 10.1016/j.str.2015.06.018.

Schüller, S. and Phillips, A. D. (2010) 'Microaerobic conditions enhance type III secretion and adherence of enterohaemorrhagic *Escherichia coli* to polarized human intestinal epithelial cells', *Environmental Microbiology*, 12(9), pp. 2426–2435. doi: 10.1111/j.1462-2920.2010.02216.x.

Schwan, C. *et al.* (2009) '*Clostridium difficile* toxin CDT induces formation of microtubule-based protrusions and increases adherence of bacteria', *PLoS Pathogens*, 5(10). doi: 10.1371/journal.ppat.1000626.

Semenyuk, E. G. *et al.* (2014) 'Spore Formation and Toxin Production in *Clostridium difficile* Biofilms', *PLoS ONE*, 9(1), p. e87757. doi: 10.1371/journal.pone.0087757.

Semenyuk, E. G. *et al.* (2015) 'Analysis of bacterial communities during *Clostridium difficile* infection in the mouse', *Infection and Immunity*, 83(11), pp. 4383–4391. doi: 10.1128/IAI.00145-15.

Shabana, L. *et al.* (2018) 'A 3D intestinal tissue model supports *Clostridioides difficile* germination, colonization, toxin production and epithelial damage', *Anaerobe*, 50(3), pp. 85–92. doi: 10.1016/j.anaerobe.2018.02.006.A.

Shan, M. *et al.* (2013) 'Mucus enhances gut homeostasis and oral tolerance by delivering immunoregulatory signals', *Science*, 342(6157), pp. 447–453. doi: 10.1126/science.1237910.

Sheng, Y. H. *et al.* (2011) 'The MUC13 cell-surface mucin protects against intestinal inflammation by inhibiting epithelial cell apoptosis', *Gut*, 60(12), pp. 1661–1670. doi: 10.1136/gut.2011.239194.

Sheng, Y. H. *et al.* (2013) 'MUC1 and MUC13 differentially regulate epithelial inflammation in response to inflammatory and infectious stimuli', *Mucosal Immunology*, 6(3), pp. 557–568. doi: 10.1038/mi.2012.98.

Sherer, K. *et al.* (2007) 'Lethal and edema toxins in the pathogenesis of Bacillus anthracis septic shock: implications for therapy', *American Journal of Respiratory and Critical Care Medicine*, 175(3), pp. 211–221.

Silwal, P. *et al.* (2018) 'AMP-Activated protein kinase and host defense against infection', *International Journal of Molecular Sciences*, 19(11). doi: 10.3390/ijms19113495.

Simpson, L. L., Maksymowych, A. B. and Hao, S. (2001) 'The Role of Zinc Binding in the Biological Activity of Botulinum Toxin', *Journal of Biological Chemistry*, 276(29), pp. 27034–27041. doi: 10.1074/jbc.M102172200.

Singhal, A. *et al.* (2014) 'Metformin as adjunct antituberculosis therapy', *Science Translational Medicine*, 6(263). doi: 10.1126/scitranslmed.3009885.

Skibinski, G., Elborn, J. S. and Ennis, M. (2007) 'Bronchial epithelial cell growth regulation in fibroblast cocultures: the role of hepatocyte growth factor', *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 293(1), pp. 69–76. doi: 10.1152/ajplung.00299.2006.

Smith, A. C. *et al.* (2006) 'O-glycosylation of serum IgA1 antibodies against mucosal and systemic antigens in IgA nephropathy', *Journal of the American Society of Nephrology*, 17(12), pp. 3520–3528. doi: 10.1681/ASN.2006060658.

Smith, R. S. *et al.* (1997) 'Fibroblasts as sentinel cells. Synthesis of chemokines and regulation of inflammation.', *The American journal of pathology*, 151(2), pp. 317–22. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9250144%0Ahttp://www.pubmedcent

ral.nih.gov/articlerender.fcgi?artid=PMC1858004.

Snyder, M. L. (1937) 'Further studies on *Bacillus difficilis* (Hall and O'Toole)', *Journal of Infectious Diseases*, 60(2), pp. 223–231. doi: 10.1093/infdis/60.2.223.

Soavelomandroso, A. P. *et al.* (2017) 'Biofilm structures in a mono-associated mouse model of Clostridium difficile infection', *Frontiers in Microbiology*, 8, pp. 1–10. doi: 10.3389/fmicb.2017.02086.

Sonnenburg, J. L. *et al.* (2005) 'Glycan foraging *in vivo* by an intestine-adapted bacterial symbiont', *Science*, 307(5717), pp. 1955–1959. doi: 10.1126/science.1109051.

Sorg, J. A. and Sonenshein, A. L. (2008) 'Bile salts and glycine as cogerminants for *Clostridium difficile* spores', *Journal of Bacteriology*, 190(7), pp. 2505–2512. doi: 10.1128/JB.01765-07.

Sorg, J. A. and Sonenshein, A. L. (2009) 'Chenodeoxycholate is an inhibitor of *Clostridium difficile* spore germination', *Journal of Bacteriology*, 191(3), pp. 1115–1117. doi: 10.1128/JB.01260-08.

Sorg, J. A. and Sonenshein, A. L. (2010) 'Inhibiting the initiation of *Clostridium difficile* spore germination using analogs of chenodeoxycholic acid, a bile acid', *Journal of Bacteriology*, 192(19), pp. 4983–4990. doi: 10.1128/JB.00610-10.

Spari, D. and Beldi, G. (2020) 'Extracellular ATP as an inter-kingdom signaling molecule: Release mechanisms by bacteria and its implication on the host', *International Journal of Molecular Sciences*, 21(15), pp. 1–14. doi: 10.3390/ijms21155590.

Stabler, R. A. *et al.* (2009) 'Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium', *Genome Biology*, 10(9), pp. 1–15. doi: 10.1186/gb-

2009-10-9-r102.

Stanimirov, B., Stankov, K. and Mikov, M. (2012) 'Pleiotropic functions of bile acids mediated by the farnesoid X receptor', *Acta Gastro-Enterologica Belgica*, 75(4), pp. 389–398.

Steele, J. *et al.* (2010) 'Piglet Models of Acute or Chronic *Clostridium difficile* Illness', *The Journal of Infectious Diseases*, 201(3), pp. 428–434. doi: 10.1086/649799.

Steele, J. *et al.* (2012) 'Systemic dissemination of *Clostridium difficile* toxins A and B is associated with severe, fatal disease in animal models', *Journal of Infectious Diseases*, 205(3), pp. 384–391. doi: 10.1093/infdis/jir748.

Steiner, T. S. *et al.* (1997) 'Fecal lactoferrin, interleukin-1β, and interleukin-8 are elevated in patients with severe *Clostridium difficile* colitis', *Clinical and Diagnostic Laboratory Immunology*, 4(6), pp. 719–722. doi: 10.1128/cdli.4.6.719-722.1997.

Sun, J. *et al.* (2005) 'Crosstalk between NF-κB and β-catenin pathways in bacterial-colonized intestinal epithelial cells', *American Journal of Physiology - Gastrointestinal and Liver Physiology*, 289(1 52-1), pp. 129–137. doi: 10.1152/ajpgi.00515.2004.

Sun, X. *et al.* (2009) 'Essential role of the glucosyltransferase activity in *Clostridium difficile* toxin-induced secretion of TNF-α by macrophages', *Microbial Pathogenesis*, 46(6), pp. 298–305. doi: 10.1038/jid.2014.371.

Sun, X. and Hirota, S. A. (2015) 'The roles of host and pathogen factors and the innate immune response in the pathogenesis of *Clostridium difficile* infection', *Molecular Immunology*. doi: 10.1016/j.molimm.2014.09.005.

Taek, N. S. *et al.* (2012) '*Clostridium difficile* Toxin A Inhibits Erythropoietin

Receptor-Mediated Colonocyte Focal Adhesion Through Inactivation of Janus Kinase-2', *Journal of Microbiology and Biotechnology*, 22(12), pp. 1629–1635. doi: 10.4014/jmb.1207.07063.

Takeoka, A. *et al.* (1991) 'Purification and characterization of S layer proteins from *Clostridium difficile*', *Journal of General Microbiology*, 137(2), pp. 261–267. doi: 10.1099/00221287-137-2-261.

Tam, J. *et al.* (2020) 'Intestinal bile acids directly modulate the structure and function of *C. difficile* TcdB toxin', *Proceedings of the National Academy of Sciences of the United States of America*, 117(12), pp. 6792–6800. doi: 10.1073/pnas.1916965117.

Tao, L. *et al.* (2016) 'Frizzled are colonic epithelial receptors for *Clostridium difficile* toxin B', *Nature*, 538(7625), pp. 350–355. doi: 10.1016/j.physbeh.2017.03.040.

Tao, L. *et al.* (2019) 'Sulfated glycosaminoglycans and low-density lipoprotein receptor contribute to *Clostridium difficile* toxin A entry into cells', *Nature Microbiology*, 4(10), pp. 1760–1769. doi: 10.1038/s41564-019-0464-z.

Tasteyre, A. *et al.* (2001) 'Role of FliC and FliD Flagellar Proteins of *Clostridium difficile* in Adherence and Gut Colonization', *Infection and Immunity*, 69(12), pp. 7937–7940. doi: 10.1128/IAI.69.12.7937.

Thanissery, R., Winston, J. A. and Theriot, C. M. (2017) 'Inhibition of spore germination, growth, and toxin activity of clinically relevant *C. difficile* strains by gut microbiota derived secondary bile acids', *Anaerobe*, 45, pp. 86–100. doi: 10.1016/j.anaerobe.2017.03.004.

Theriot, C. M., Bowman, A. B. and Young, V. B. (2016) 'Antibiotic-induced alterations of the gut microbiota alter secondary bile acid production and allow

for *Clostridium difficile* spore germination and outgrowth in the large intestine', *mSphere*, 1(1), pp. e00045-15.

Triantafilou, K. *et al.* (2001) 'Fluorescence recovery after photobleaching reveals that LPS rapidly transfers from CD14 to hsp70 and hsp90 on the cell membrane', *Journal of Cell Science*, 114(13), pp. 2535–2545.

Trunk, T., S. Khalil, H. and C. Leo, J. (2018) 'Bacterial autoaggregation', *AIMS Microbiology*, 4(1), pp. 140–164. doi: 10.3934/microbiol.2018.1.140.

Tucker, K. D. and Wilkins, T. D. (1991) 'Toxin A of *Clostridium difficile* binds to the human carbohydrate antigens I, X, and Y', *Infection and Immunity*, 59(1), pp. 73–78. doi: 10.1128/iai.59.1.73-78.1991.

Tulli, L. *et al.* (2013) 'CbpA: A Novel Surface Exposed Adhesin Of *Clostridium difficile* Targeting Human Collagen', *Cellular Microbiology*, 15(10), pp. 1674–1687. doi: 10.1111/cmi.12139.

Underwood, S. *et al.* (2009) 'Characterization of the sporulation initiation pathway of *Clostridium difficile* and its role in toxin production', *Journal of Bacteriology*, 191(23), pp. 7296–7305. doi: 10.1128/JB.00882-09.

Vaninsberghe, D. *et al.* (2020) 'Diarrhoeal events can trigger long-term *Clostridium difficile* colonization with recurrent blooms', *Nature Microbiology*, 5, pp. 642–650.

Vardakas, K. Z. *et al.* (2012) 'Treatment failure and recurrence of *Clostridium difficile* infection following treatment with vancomycin or metronidazole: A systematic review of the evidence', *International Journal of Antimicrobial Agents*, 40(1), pp. 1–8. doi: 10.1016/j.ijantimicag.2012.01.004.

Venteclef, N. *et al.* (2011) 'Metabolic nuclear receptor signaling and the inflammatory acute phase response', *Trends in Endocrinology and Metabolism*,

22(8), pp. 333–343. doi: 10.1016/j.tem.2011.04.004.

Voth, D. E. and Ballard, J. D. (2005) '*Clostridium difficile* toxins: mechanism of action and role in disease', *Clinical microbiology reviews*, 18(2), pp. 247–263.

Vowinkel, T. *et al.* (2004) 'Apolipoprotein A-IV inhibits experimental colitis', *The Journal of Clinical Investigation*, 114(2), pp. 260–269.

Waligora, A. J. *et al.* (1999) '*Clostridium difficile* Cell Attachment Is Modified by Environmental Factors', *Applied and Environmental Microbiology*, 65(9), pp. 4234–4238. doi: 10.1128/aem.65.9.4234-4238.1999.

Waligora, A. J. *et al.* (2001) 'Characterization of a cell surface protein of *Clostridium difficile* with adhesive properties', *Infection and Immunity*, 69(4), pp. 2144–2153. doi: 10.1128/IAI.69.4.2144-2153.2001.

Walsh, S. V. *et al.* (2001) 'Rho kinase regulates tight junction function and is necessary for tight junction assembly in polarized intestinal epithelia', *Gastroenterology*, 121(3), pp. 566–579. doi: 10.1053/gast.2001.27060.

Wang, J. *et al.* (2018) 'Ceragenin CSA13 Reduces *Clostridium difficile* Infection in Mice by Modulating the Intestinal Microbiome and Metabolites', *Gastroenterology*, 154(6), pp. 1737–1750. doi: 10.1053/j.gastro.2018.01.026.

Wang, T. *et al.* (2017) 'Knockdown of the differentially expressed gene TNFRSF12A inhibits hepatocellular carcinoma cell proliferation and migration in vitro', *Molecular Medicine Reports*, 15(3), pp. 1172–1178. doi: 10.3892/mmr.2017.6154.

Wann, E. R., Gurusiddappa, S. and Höök, M. (2000) 'The fibronectin-binding MSCRAMM FnbpA of *Staphylococcus aureus* is a bifunctional protein that also binds to fibrinogen', *Journal of Biological Chemistry*, 275(18), pp. 13863–13871. doi: 10.1074/jbc.275.18.13863.

Warny, M. *et al.* (1994) 'Human antibody response to *Clostridium difficile* toxin A in relation to clinical course of infection', *Infection and Immunity*, 62(2), pp. 384–389. doi: 10.1128/iai.62.2.384-389.1994.

Warny, M. *et al.* (2000) 'p38 MAP kinase activation by *Clostridium difficile* toxin A mediates monocyte necrosis, IL-8 production, and enteritis', *Journal of Clinical Investigation*, 105(8), pp. 1147–1156. doi: 10.1172/JCI7545.

Wegner, A. and Aktories, K. (1988) 'ADP-ribosylated actin caps the barbed ends of actin filaments', *Journal of Biological Chemistry*, 263(27), pp. 13739–13742.

Wells, J. E. and Hylemon, P. B. (2000) 'Identification and characterization of a bile acid 7α-dehydroxylation operon in Clostridium sp. strain TO-931, a highly active 7α-dehydroxylating strain isolated from human feces', *Applied and Environmental Microbiology*, 66(3), pp. 1107–1113. doi: 10.1128/AEM.66.3.1107-1113.2000.

Westermann, A. J. *et al.* (2016) 'Dual RNA-seq unveils noncoding RNA functions in host-pathogen interactions', *Nature*. Nature Publishing Group, 529(7587), pp. 496–501. doi: 10.1038/nature16547.

Westermann, A. J., Barquist, L. and Vogel, J. (2017) 'Resolving host–pathogen interactions by dual RNA-seq', *PLoS Pathogens*, 13(2), pp. 1–19. doi: 10.1371/journal.ppat.1006033.

Westermann, A. J., Gorski, S. A. and Vogel, J. (2012) 'Dual RNA-seq of pathogen and host', *Nature Reviews Microbiology*. Nature Publishing Group, 10(9), pp. 618–630. doi: 10.1038/nrmicro2852.

Williams, S. J. *et al.* (2001) 'MUC13, a Novel Human Cell Surface Mucin Expressed by Epithelial and Hemopoietic Cells', *Journal of Biological Chemistry*, 276(21), pp. 18327–18336. doi: 10.1074/jbc.M008850200.

Wilson, K. H. (1983) 'Efficiency of various bile salt preparations for stimulation of *Clostridium difficile* spore germination', *Journal of Clinical Microbiology*, 18(4), pp. 1017–1019. doi: 10.1128/jcm.18.4.1017-1019.1983.

Wilson, K. H. and Perini, F. (1988) 'Role of competition for nutrients in suppression of *Clostridium difficile* by the colonic microflora', *Infection and Immunity*, 56(10), pp. 2610–2614. doi: 10.1128/iai.56.10.2610-2614.1988.

Wilson, K. H. and Sheagren, J. N. (1983) 'Antagonism of toxigenic *Clostridium difficile* by nontoxigenic *C. difficile*', *Journal of Infectious Diseases*, 147(4), pp. 733–736. doi: 10.1093/infdis/147.4.733.

Wu, X. *et al.* (2008) 'Rac1 activation and subsequent β-catenin phosphorylation controls nuclear localization of β-catenin during canonical Wnt signaling', *Cell*, 133(2), pp. 340–353. doi: 10.1038/jid.2014.371.

Wullt, M. *et al.* (2012) 'IgG antibody response to toxins A and B in patients with *Clostridium difficile* infection', *Clinical and Vaccine Immunology*, 19(9), pp. 1552–1554. doi: 10.1128/CVI.00210-12.

Xia, Y. *et al.* (2000) '*Clostridium difficile* toxin A excites enteric neurones and suppresses sympathetic neurotransmission in the guinea pig', *Gut*, 46(4), pp. 481–486. doi: 10.1136/gut.46.4.481.

Yang, W. *et al.* (2015) 'Overlapping and unique signatures in the proteomic and transcriptomic responses of the nematode *Caenorhabditis elegans* toward pathogenic *Bacillus thuringiensis*', *Developmental and Comparative Immunology*, 51(1), pp. 1–9. doi: 10.1016/j.dci.2015.02.010.

Yoch, D. C. and Valentine, R. C. (1972) 'Ferredoxins and Flavodoxins of Bacteria', *Annual Review of Microbiology*, 26, pp. 139–162.

Yoshino, Y. *et al.* (2013) '*Clostridium difficile* flagellin stimulates toll-like

receptor 5, and toxin B promotes flagellin-induced chemokine production via TLR5', *Life Sciences*, 92(3), pp. 211–217. doi: 10.1016/j.lfs.2012.11.017.

Yu, H. *et al.* (2017) 'Cytokines Are Markers of the *Clostridium difficile*-Induced Inflammatory Response and Predict Disease Severity', *Clinical and Vaccine Immunology*, 24(8), pp. 1–11.

Yuan, P. *et al.* (2015) 'Chondroitin sulfate proteoglycan 4 functions as the cellular receptor for *Clostridium difficile* toxin B', *Cell Research*, 25(2), pp. 157–168. doi: 10.1038/cr.2014.169.

Zackular, J. P. *et al.* (2017) 'Dietary Zinc Alters the Microbiota and Decreases Resistance to *Clostridium difficile* Infection', *HHS Access*, 22(11), pp. 1330–1334. doi: 10.1038/nm.4174.Dietary.

Zhao, H. *et al.* (2015) 'Compound 13, ana1-selective small molecule activator of AMPK, inhibits *Helicobacter pylori*-induced oxidative stresses and gastricepithelial cell apoptosis', *Biochemical and Biophysical Research Communications*, 463(4), pp. 510–517. Available at: http://dx.doi.org/10.1016/j.bbrc.2015.05.059.