University of Warwick institutional repository: http://go.warwick.ac.uk/wrap

This paper is made available online in accordance with publisher policies. Please scroll down to view the document itself. Please refer to the repository record for this item and our policy information available from the repository home page for further information.

To see the final version of this paper please visit the publisher's website. Access to the published version may require a subscription.

Author(s): K. R. Collins and A. J. Easton
Article Title: Sequence variation in the haemagglutinin-neuraminidase gene of human parainfluenza virus type 3 isolates in the UK
Year of publication: 1995
Link to published version:
http://dx.doi.org/10.1017/S0950268800052201
Publisher statement: None

Epidemiol. Infect. (1995). **114**. 493–500
Copyright © 1995 Cambridge University Press

Sequence variation in the haemagglutinin-neuraminidase gene of human parainfluenza virus type 3 isolates in the UK

K. R. COLLINS AND A. J. EASTON*

Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK

(*Accepted 17 January 1995*)

SUMMARY

The sequence variation in a 934 base-pair region of the gene encoding the haemagglutinin-neuraminidase of five human parainfluenza virus type 3 (HPIV3) isolates was determined together with that of a prototype UK strain. All of the clinical isolates were from the Manchester area of the UK and were obtained in 1990, 1991 and 1993. The gene segment was amplified by the polymerase chain reaction using HPIV3-specific oligonucleotide primers. The nucleotide homology of the strains was high, around 99% and specific differences in the UK sequences when compared with that of the US prototype strain were identified. In addition, a number of isolate-specific differences were seen. No correlation was detected between the observed nucleotide mutations and the year of isolation, which supports the hypothesis that HPIV3 shows cocirculation of a heterogeneous population of viruses rather than varying with time in a linear fashion. However, the data suggested that geographically-defined genetic lineages of HPIV3 may exist.

INTRODUCTION

Human parainfluenza virus 3 (HPIV3) infections are a major cause of respiratory tract infections in young infants [1] and are second only to respiratory syncytial virus in incidence [2]. HPIV3 infection occurs world-wide and, in the United Kingdom at least, follows an epidemic pattern with peak incidence occurring in the summer months in weeks 21–32 [3]. Many sporadic cases of infection are observed between the peak periods and these may account for the earlier long-held view that HPIV3 was endemic [3]. Currently no effective vaccine is available.

Several studies have been carried out on the relatedness and possible evolution of the mammalian parainfluenza viruses, both by nucleotide sequence comparisons and by using panels of monoclonal antibodies (MAb). A panel of MAb raised against the haemagglutinin-neuraminidase (HN), fusion (F), nucleocapsid (N) and matrix (M) proteins of PIV3 strains showed little relationship between human and bovine viruses [4, 5], and of these proteins the most antigenically divergent was HN [6].

Previous investigations have concentrated on the antigenic [7] and nucleotide diversity of clinical isolates of HPIV3 from diverse geographical locations over a

* Corresponding author.

wide time scale. van Wyke Coelingh, Winter and Murphy [8, 9] sequenced the HN and F genes of six clinical isolates obtained from Texas and Washington USA and Australia between 1973–83. These strains were found to be highly conserved in both genes, with approximately 3% amino acid variability seen between isolates. The antigenic analysis was carried out on a larger sample number, and variation within the HN protein could sometimes be correlated with the geographic origin of the strain, although this was not observed with the smaller number of strains sequenced. The nucleotide diversity of the HPIV3 HN gene showed a similar pattern to that of influenza C virus, with an overall low rate of non-cumulative change and genetic heterogeneity rather than the progressive accumulation of mutations observed in the influenza A virus attachment gene.

It has recently been shown with HPIV1 that although cocirculation of strains is commonly seen, geographical lineages may develop [10]. However, it is not known whether such geographical lineages develop with HPIV3 infections. Comparison of the sequence data for UK isolates with other known HPIV3 sequences may provide information on whether such a geographical lineage has developed for HPIV3.

We report here an investigation of the sequence variability in the HN gene of HPIV3 within isolates obtained from a restricted area over a short time period to determine whether cocirculation of HPIV3 variants occurs or whether a geographically restricted monotypic variant has developed. The isolates examined were obtained over a period of three epidemics in one geographical area of the United Kingdom (Manchester).

MATERIALS AND METHODS

Virus strains

Clinical isolates, identified by direct immunofluorescence of throat swab material, were provided by Dr A. Turner of the Public Health Laboratory, Manchester. The isolates were supplied either as an aliquot of original sample material, or as tissue culture fluid from a first passage in cell culture (Table 1). Strain MK9, provided by the Public Health Laboratory, Colindale, London was used as a UK 'prototype' HPIV3 strain. Viruses were grown in BSC-1 monkey green kidney cells passaged twice or three times to provide sufficient material for analysis.

PCR amplification and cloning

Cytoplasmic RNA was extracted from infected cells as described previously[11]. This was used to direct synthesis of cDNA using an oligo(dT) oligonucleotide primer and the cDNA was amplified by polymerase chain reaction (PCR) at 94 °C for 45 s, 54 °C for 45 s and 74 °C for 2 min for 30 cycles [12]. The primers used were HN2 and HN4 ($C_{938}ATCATCAGGCATAGAAG_{955}$ and $A_{1872}CTAATGAATAG_{1861}$ respectively). The nucleotide positions of the 5′ and 3′ residues are indicated. Following analysis of all of the sequences for the HN genes of HPIV1, HPIV2 and HPIV3 available on the GENEMBL database, the primers were designed to fall within totally conserved regions of all HPIV3 isolates sequenced to date and

Table 1. *HPIV3 isolates*

Isolate	Year of isolation	Age of patient	Sex
V90/02293	1990	1 month	Female
V90/02499	1990	1 year	Male
V91/01836	1991	4 months	Female
V91/02200	1991	4 months	Female
V93/03469	1993	8 months	Male

* Virus isolates used in this study.

amplified a fragment of 934 base pairs. In addition, the primers were also designed to decrease the possibility of amplifying any misdiagnosed HPIV1 or HPIV2 isolates by selecting areas of sequence variation between these viruses and HPIV3. PCR fragments were isolated from agarose gels and inserted into bacteriophage M13 DNA or the pCR™II cloning vector (InVitrogen BV) using standard techniques [13] either directly or following digestion with the restriction endonuclease *Bgl* II which recognized a single site within the amplified DNA. These were then amplified in *E. coli*.

DNA sequencing

Nucleotide sequence analysis was carried out using the dideoxy chain termination method [14]. The ^{35}S-dATP radioactive label was supplied by Amersham International plc. All other reagents were supplied in the Sequenase™ kit version 2·0 (United States Biochemical Corp.). The protocol was carried out in accordance with the manufacturers recommendations. All inserts were sequenced at least twice.

RESULTS AND DISCUSSION

Sequences were compared with the prototype strain of HPIV3 PI3/Wash/47885/57 [14]. As a consequence of the cloning and sequencing strategies employed, the region between residues 938–1331 of strain V90/02293 and residues 1347–1362 of strains V91/02200 and V93/03469 could not be determined unambiguously and have been omitted from subsequent analyses.

When the determined sequences were compared it was seen that no insertions or deletions had occurred in any of the isolates in the section of the HN gene examined. Compared with the American prototype HPIV3 strain, six mutations occurred in the same position in all UK isolates, and of these only one resulted in a coding change from serine to alanine (Table 2). In addition, each of the UK isolates contained unique differences. These are summarized in Table 3 together with the resulting coding changes where appropriate. Two mutations were seen which occurred in five out of six of the isolates; a $T \leftrightarrow C$ mutation at nucleotide position 1203 and an $A \leftrightarrow T$ mutation at nucleotide position 1518, but these did not correlate with the year of isolation as in both instances the mutations did not appear in a 1991 isolate, but were detected in the 1993 isolate.

A small number of mutations were observed to be isolate-specific in that certain nucleotide positions showed mutations in only one of the variants. The MK9 strain showed the highest level, having three mutations not observed in any of the

Table 2. *Comparison of common nucleotide differences seen in UK isolates compared to the American prototype strain*

Nucleotide position*	Nucleotide change	Amino acid change
969	C ↔ T (3)†	—
1456	C ↔ T (1)	—
1500	C ↔ T (3)	—
1542	A ↔ G (3)	—
1627	T ↔ G (1)	Ser ↔ Ala [512]
1836	C ↔ T ( )	Non-coding

* Nucleotide changes common to UK HPIV3 isolates compared to the American prototype strain PI3/Wash/47885/57.

† The sequence UK isolate V90/02293 is unknown at this position. The position of the nucleotide in the triplet codon is indicated in round brackets and the amino acid residue is given in square brackets.

Table 3. *Specific nucleotide sequence differences of UK isolates compared to the American prototype strain*

Virus isolate	Nucleotide position	Nucleotide* change	Amino acid change
MK9	1057	G ↔ C (1)	Val ↔ Leu [322]
	1203	T ↔ C (3)	—
	1378	C ↔ T (1)	His ↔ Tyr [429]
	1431	A ↔ C (3)	Arg ↔ Ser [446]
	1518	A ↔ T (3)	—
V90/02293	1179	C ↔ G (3)	Asp ↔ Glu [362]
	1203	T ↔ C (3)	—
	1346	A ↔ C (2)	Lys ↔ Thr [418]
	1518	A ↔ T (3)	—
V90/02499	1203	T ↔ C (3)	—
	1518	A ↔ T (3)	—
	1606	C ↔ T (1)	Pro ↔ Ser [505]
V91/02200	1203	T ↔ C (3)	—
	1606	C ↔ T (1)	Pro ↔ Ser [505]
V91/01836	1134	T ↔ C (3)	—
	1518	A ↔ T (3)	—
	1825	A ↔ G ( )	Non-coding
V93/03469	1130	A ↔ T (2)	Glu ↔ Val [346]
	1203	T ↔ C (3)	—
	1518	A ↔ T (3)	—

* Nucleotide changes specific to individual UK isolates compared to the American prototype strain PI3/Wash/47885/57. The position of the nucleotide in the triplet codon is indicated in round brackets and the amino acid residue is given in square brackets.

other isolates, all of which were coding changes. However, this UK prototype strain has been passaged in tissue culture more extensively than the other isolates and this may account for some of these mutations.

The overall level of homology with the 1957 prototype strain at the nucleotide level ranged from 98·8% (MK9)-99·0% (isolates V90/02499, V91/01836 and V93/03469). Similarly high levels of homology were found at the deduced amino acid level, from 98·6% (MK9) to 99·3% (isolates V90/02499, V91/01836 and V93/03469). This corresponds well with the percentage homology detected by van

Wyke Coelingh, Winter and Murphy [8] in the domain 2 region of six strains isolated over a 26-year time period, where the percentage homology between the isolates examined and the Washington 1957 prototype strain at the amino acid level ranged from 98·6–99·8%.

From the high level of homology seen at the amino acid level and the nature of the amino acid changes in the UK isolates it seems unlikely that any of these mutations would alter the integral characteristics of the protein such as hydrophobicity/charge and consequently, structure. The functionally important regions on the HPIV3 HN protein have not been well defined although they have been narrowed down to within the domain 2 region of the protein [8, 16], part of which was encompassed in this study. As might be expected, no mutations were found in any of the isolates within the fully conserved sequence NRKCS (amino acids 252–258) proposed as the site for neuraminidase activity [17, 18].

The data from this study, together with that previously reported suggests that the number of nucleotide differences are not directly proportional to the amount of time between isolation of the viruses, and that progressive mutations do not appear to accumulate. This is shown most clearly by the close relationship between the virus strain isolated in Washington in 1957 and strains isolated nearly 40 years later in the UK. However, it must be borne in mind that studies of this kind are normally only able to identify forward mutations. It is possible that certain nucleotides are particularly prone to point mutations and may oscillate around a given forward and backward sequence. Such oscillations may not be detected by this kind of analysis, giving an underestimation of the observed rate of mutation.

Figure 1 shows two dendrodram plots constructed from a pairwise nucleotide comparison of the parts of the HN gene of the UK HPIV3 isolates sequenced, together with previously reported HPIV3 HN sequence data. The comparisons were carried out in two parts to exclude regions of sequence ambiguity, especially in strain V90/02293. The distance on the horizontal axis is an indication of the level of sequence divergence of the strains. A striking feature of the two dendrograms is the apparently close relationship between the UK isolates, which appear to be more closely related at the nucleotide level than are the US isolates to each other. This may reflect the fact that all the UK isolates came from one city, whereas the US isolates were from two geographically distant cities. It is of interest to note that the 1957 Washington strain appears to be more closely related to the UK isolates from the 1990s than to US isolates from the late 1970s–early 1980s. The reasons for this divergence are unclear, and a study of HPIV3 variants isolated in the Washington area over the period 1957–79 may be of value. The Australian isolate is also quite clearly closer in nucleotide sequence to the UK isolates than to the US isolates. With the exception of the Texas 1982 strain, the US isolates fall into a lineage distinct from the UK/Australian isolates and prototype strains. The Texas 1982 isolate appears more closely related to the other US strains when comparing the 5′ end of the sequenced region than the 3′ end. The 3′-end region appears to be more closely related to that of the UK strains. This may suggest that domains exist within the HN protein, where one domain may mutate to become more similar to the US strains and the other more similar to the UK strains. Another possibility would be a recombination event between two strains of HPIV3, although currently there is no evidence to suggest that such

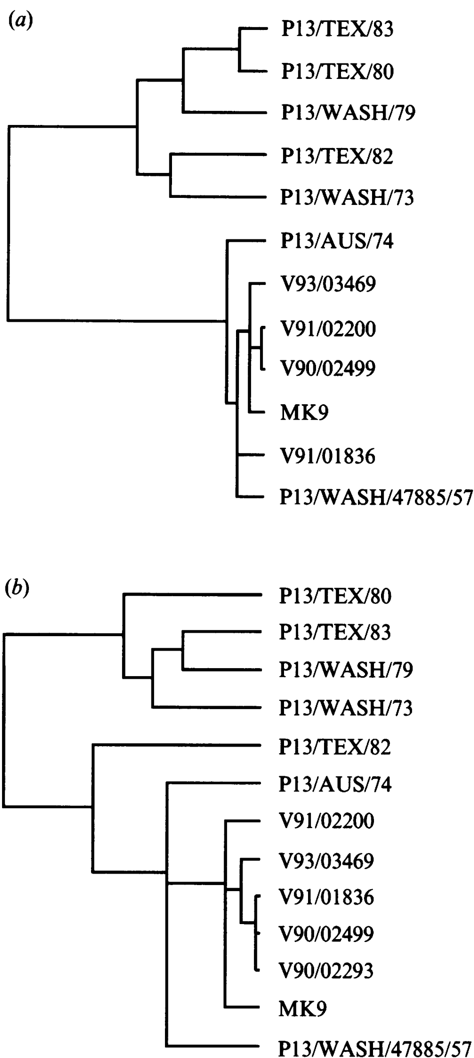

Fig. 1. Dendrogram plot of HPIV3 isolates based on the HN gene sequence. (*a*) Dendrogram based on nucleotide 938 to 1347. (*b*) Dendrogram based on nucleotides 1363 to 1871. The dendrogram was generated by pairwise comparison of the nucleotide sequences using the CLUSTAL program [19]. The sequence of PI3/Wash/47885/57 was taken from Elango and colleagues [15] and PI3/WASH/73, PI3/AUS/74, PI3/WASH/79, PI3/TEX/80, PI3/TEX/82 and PI3/TEX/83 were from van Wyke Coelingh, Winter and Murphy [8].

recombination events are found within the non-segmented, single stranded, negative sense RNA viruses.

The results illustrated in Fig. 1 show similarities to a recently reported study on the HN protein of HPIV1 [10] in which an evolutionary tree was constructed from 26 HPIV1 HN sequences spanning a 26-year time period suggesting co-circulation of multiple lineages, but with the proviso that some of the lineages were geographically limited. The level of variation seen in the HN gene of HPIV3 is similar to that reported for HPIV1 which showed only a maximum of 3·1% difference in the amino acid sequence of the predicted HN proteins of subgroup A

and subgroup B isolates over a 26-year period [20]. This is similar to the situation for influenza B and C viruses which show little sequence variation in HA genes and no correlation in the rate of change with time [21, 22]. From these analyses it appears that the parainfluenzaviruses and influenza B and C viruses do not undergo the rapid evolution of new HN variants such as is seen with HA variants of influenza A virus [23]. Similarly, populations of parainfluenzaviruses do not show the extensive sequence variation seen with human respiratory syncytial virus in which there may be as much as 20% variation in the amino acid sequence of attachment proteins of viruses of the same subgroup [24].

Comparison of the data presented here with that of van Wyke Coelingh and colleagues [8] supports the suggestions that co-circulation of lineages occurs with HPIV3 virus and that mutations in the HPIV3 HN gene do not correlate with the year of virus isolation. However, this study does suggest that dominance of geographically defined lineages may also occur in HPIV3 isolates. Analysis of a larger number of clinical isolates should further clarify this.

ACKNOWLEDGEMENTS

We would like to thank Dr A. Turner for kindly providing the isolates used in this work.

REFERENCES

1. Glezen W, Frank AL, Taber LH, Kasel JA. Parainfluenza virus type 3: Seasonality and risk of infection and reinfection in young children. J Infect Dis 1984; **150**: 851–7.
2. Black FL. Epidemiology of paramyxoviridae. In: Kingsbury DW, ed. The paramyxoviruses. New York: Plenum Publishing Corporation, 1991: 509–29.
3. Easton AJ, Eglin RP. Epidemiology of parainfluenza virus type 3 in England and Wales over a ten-year period. Epidemiol Infect 1989; **102**: 531–5.
4. Ray R, Compans RW. Monoclonal antibodies reveal extensive antigenic differences between the hemagglutinin-neuraminidase glycoproteins of human and bovine parainfluenza 3 viruses. Virology 1986; **148**: 232–6.
5. Rydbeck R, Love A, Orvell C, Norrby E. Antigenic analysis of human and bovine parainfluenza virus type 3 strains with monoclonal antibodies. J Gen Virol 1987; **68**: 2153–60.
6. Klippmark E, Rydbeck R, Shibuta H, Norrby E. Antigenic variation of human and bovine parainfluenza virus type 3 strains. J Gen Virol 1990; **71**: 1577–80.
7. van Wyke Coelingh KL, Winter CC, Murphy BR. Antigenic variation in the hemagglutinin-neuraminidase protein of human parainfluenza type 3 virus. Virology 1985; **143**: 569–82.
8. van Wyke Coelingh KL, Winter CC, Murphy BR. Nucleotide and deduced amino acid sequence of hemagglutinin-neuraminidase genes of human type 3 parainfluenza viruses isolated from 1957 to 1983. Virology 1988; **162**: 137–43.
9. van Wyke Coelingh KL, Winter CC, Tierney EL, London WT, Murphy BR. Antibody responses of humans and nonhuman primates to individual antigenic sites of the hemagglutinin-neuraminidase and fusion glycoproteins after primary infection or reinfection with parainfluenza type 3 virus. J Virol 1990; **64**, 3833–43.
10. Hetherington SV, Watson AS, Scroggs RA, Portner A. Human parainfluenza virus type 1 evolution combines cocirculation of strains and development of geographically restricted lineages. J Infect Dis 1994; **160**: 248–52.
11. Ling R, Easton AL, Pringle CR. Sequence analysis of the 22K, SH and G genes of turkey rhinotracheitis and their intergenic regions reveals a gene order different from that of other pneumoviruses. J Gen Virol 1992; **73**: 1709–15.
12. Saiki RK, Gelfand DH, Stoffel S, et al. Primer-directed enzymatic amplification of dna with a thermostable dna polymerase. Science 1989; **239**: 487–97.

13. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning, a laboratory manual, 2nd ed. Cold Spring Harbor: Cold Spring Harbor Press, 1989.
14. Sanger F, Coulsen AR, Barrell BG, Smith AJH, Roe BA. Cloning in single stranded bacteriophage as an aid to rapid DNA sequencing. J Mol Biol 1980; **143**: 161–78.
15. Elango N, Coligan JE, Jambou RC, Venkatesan S. Human parainfluenza type 3 virus hemagglutinin-neuraminidase glycoprotein: nucleotide sequence of mRNA and limited amino acid sequence of the purified protein. J Virol 1986; **57**: 481–9.
16. Morrison T, Portner A. Structure, function, and intracellular processing of the glycoproteins of paramyxoviridae. In: Kingsbury DW, ed. The paramyxoviruses. New York: Plenum Publishing Corporation, 1991: 347–83.
17. Jorgensen ED, Collins PL, Lomedico PT. Cloning and nucleotide sequence of newcastle disease virus hemagglutinin-neuraminidase mRNA: Identification of a putative sialic acid binding site. Virology 1987; **156**: 12–24.
18. Colman PM, Hoyne PA, Lawrence MC. Sequence and structure alignment of paramyxovirus hemagglutinin-neuraminidase with influenza virus neuraminidase. J Virol 1993; **67**: 2972–80.
19. Higgins DG, Sharp PM. CLUSTAL: A package for performing multiple sequence alignment on a microcomputer. Gene 1988; **73**: 237–44.
20. Henrickson KJ, Savatski LL. Genetic variation of human parainfluenza virus type 1 haemagglutinin-neuraminidase: Analysis of 12 clinical isolates. J Infect Dis 1992; **166**: 995–1005.
21. Yamashita M, Krystal M, Fitch WM, Palese P. Influenza B virus evolution – co-circulating lineages and comparison of evolutionary pattern with those of influenza A and influenza C viruses. Virology 1988; **163**: 112–22.
22. Buonagurio DA, Nakada S, Desselberger U, Krystal M, Paleses P. Noncumulative sequence changes in the hemagglutinin genes of influenza C virus isolates. Virology 1985; **146**: 221–32.
23. Raymond FL, Caton AJ, Cox NJ, Kendal AP, Brownlee GG. The antigenicity and evolution of influenza H1 hemagglutinin, from 1950–1957 and 1977–1983: 2 pathways from one gene. Virology 1986; **148**: 275–87.
24. Cane PA, Matthews DA, Pringle CR. Analysis of respiratory syncitial virus strain variation in successive epidemics in one city. J Clin Microbiol 1994; **32**: 1–4.