

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

**A Thesis Submitted for the Degree of PhD at the University of Warwick**

<http://go.warwick.ac.uk/wrap/62061>

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

**The Role of Mitochondrial DNA In the Tumor  
Biology of Glioblastoma Multiforme and Multiple  
Myeloma**

**by**

**Ka Yu Yeung**

A thesis submitted in fulfilment of the requirements  
for the degree of  
**Doctor of Philosophy**



**April 2014**

## Table of Contents

List of Figures	14
List of Tables	20
Acknowledgements	23
Declaration	24
Abstract	25
Abbreviations	26
<b>CHAPTER 1: GENERAL INTRODUCTION</b>	<b>30</b>
1.1. Mitochondrial DNA and its role in energy production	31
1.2. Maintenance of mtDNA and its mode of replication	32
1.3. Overview of tumor biology	36
1.4. Heteroplasmy and mtDNA	38
1.5. Mitochondrial energetics	40
1.5.1. Glycolysis	41
1.5.2. Lactic acid fermentation	43
1.5.3. Citric acid cycle	45
1.5.4. Oxidative phosphorylation	47
1.5.4.1. Protein modelling analysis of the ETC proteins	50
1.5.4.1.1. Cytochrome bc1 complex	50
1.5.4.1.2. Cytochrome c oxidase complex	53
1.6. Warburg effect	55
1.7. Other effects of mitochondrial dysfunction in cancer	58
<b>1.8. MtDNA variants in cancer</b>	<b>58</b>
1.9. Multiple myeloma and mitochondria	62

1.10.	GBM and mitochondria	65
1.11.	Conclusion	67
1.12.	Hypothesis	68
1.13.	Aims and objectives	68
<b>CHAPTER 2: GENERAL METHODS</b>		<b>71</b>
2.1.	Cell culture	72
2.1.1.	Passaging of adherent cultures	72
2.1.2.	Passaging of non-adherent (suspension) cells	73
2.1.3.	Human foreskin cell line, BJ fibroblasts	74
2.1.4.	Human hepatocarcinoma cell line, HepG2	74
2.1.5.	Human multiple myeloma, U266 cells	74
2.1.6.	Human glioblastoma multiforme, HSR-GBM1/GBM L1/ GBM 6/GBM L2 cells	75
2.1.7.	Human glioblastoma multiforme, NO7-152 cells	75
2.1.8.	Human glioblastoma multiforme, GBM4/CSC-020/ CSC-014 cells	75
2.1.9.	Depletion and replenishment of mtDNA in U266 and HSR-GBM1 cells	76
2.1.10.	Inhibition of glycolysis using 2-deoxyglucose	76
2.2.	Isolation of plasma cells	76
2.3.	Fluorescent-activated cell sorting (FACS) on U266 cells	77
2.4.	DNA extraction	77
2.4.1.	DNA extraction from cells	78
2.4.2.	DNA extraction from tissues	79
2.5.	RNA extraction	79

2.5.1. RNA extraction from cultured cells	79
2.5.2. RNA extraction from tissue samples	81
2.6. Spectrophotometry	81
2.7. Reverse Transcription Polymerase Chain Reaction (RT-PCR)	82
2.8. Polymerase Chain Reaction (PCR)	82
2.8.1. Agarose gel electrophoresis	83
2.8.2. Gel documentation and analysis	84
2.9. PCR product purification	84
2.9.1. Direct purification of PCR products	84
2.9.2. Purification of PCR products from agarose gel following electrophoresis	85
2.10. Long range PCR	85
2.11. Real time PCR (qPCR)	86
2.11.1 Analysis of mtDNA copy number	88
2.11.2. Gene expression analysis	88
2.12. ATP production	88
2.13. Lactate production	89
2.14. Respirometry assessment on cells to determine their OXPHOS capacity	90
2.15. DNA cloning in bacteria	92
2.15.1 Plasmid purification	93
2.16. High resolution melt (HRM) analysis	94
2.17. Capillary sequencing	96
2.18. Next generation sequencing	97
2.18.1. Quantification of double-stranded DNA	98

2.18.2. Fragmentation of the amplicons from long PCR	99
2.18.3. End repair of the fragmented DNA and sample purification	100
2.18.4. Ligation of adapters, nick translation and DNA sample purification	103
2.18.5. Size-selection of the unamplified library	104
2.18.6. Amplification of the 200 base-read DNA libraries and purification	105
2.18.7. Quantification of the DNA libraries	106
2.18.8. Emulsion PCR on the Ion OneTouch instrument	108
2.18.9. Recovery of the template-positive ISPs	110
2.18.10. Enrich the template-positive ISPs	110
2.18.11. Rinsing of the enriched ISPs	111
2.18.12. Preparation of the enriched ISPs for sequencing	111
2.19. Detection of mtDNA variants	112
2.20. Predictive <i>in silico</i> analysis	113
2.21. Determining changes to the amino acid sequence	114
2.22. Determining susceptibility rates of the mtDNA regions to mutation	114
2.23. Structural analysis of proteins	115
2.24. Phylogenetic analysis	115
2.25. Statistical analysis	115
<b>CHAPTER 3: IDENTIFICATION OF MTDNA VARIANTS IN PATIENTS DIAGNOSED WITH MULTIPLE MYELOMA</b>	<b>117</b>
3.1. Introduction	118
3.2. Aims and hypothesis	121

3.3.	Preliminary experiments	<b>122</b>
3.3.1.	Optimisation of the long PCR.	<b>122</b>
3.3.2.	Optimisation of the Ion Torrent sequencing workflow	<b>126</b>
3.3.3.	Ion Torrent sequencing on control DNA samples	<b>130</b>
3.3.4.	Determining the sensitivity of variant detection on the Ion Torrent sequencer and HRM system.	<b>134</b>
3.4.	Methods	<b>137</b>
3.4.1.	HRM primers for the multiple myeloma mtDNA variants	<b>137</b>
3.4.2.	Ethics statement	<b>137</b>
3.5.	Results	<b>139</b>
3.5.1.	Detection of mtDNA variants in multiple myeloma tumors	<b>139</b>
3.5.2.	Verification of variants and screening of additional tumors using HRM	<b>139</b>
3.5.3.	Susceptibility of mtDNA regions to the development of mutations	<b>143</b>
3.5.4.	Identification of variants during mtDNA replenishment in U266 cells	<b>145</b>
3.5.5.	Susceptibility of depleted and recovered U266 cells to mutation	<b>148</b>
3.5.6.	MtDNA variants in U266 sub-populations of CD45+ and CD45- cells	<b>150</b>
3.5.7.	Susceptibility of CD45+ and CD45- cells to mutations	<b>155</b>
3.5.8.	The influence of GC content on the development of mutations	<b>156</b>
3.5.9.	<i>In silico</i> and protein structure analysis	<b>158</b>

3.5.9.1. Multiple myeloma tumor sample variants	<b>158</b>
3.5.9.2. CD45+ and CD45- sub-population variants	<b>161</b>
3.6. Discussion	<b>166</b>
3.6.1. Multiple myeloma variants	<b>166</b>
3.6.2. Impact of variants located in the non-coding regions of mtDNA	<b>166</b>
3.6.3. Replenishment of mtDNA and mtDNA variants	<b>168</b>
3.6.4. Additional considerations	<b>169</b>
3.6.5. Protein modelling analysis	<b>170</b>
3.6.6. Contribution of mtDNA variants to tumor biology	<b>173</b>
3.7. Conclusion	<b>174</b>
<b>CHAPTER 4: THE ROLE OF MTDNA IN THE MAINTENANCE OF MITOCHONDRIA AND MULTIPLE MYELOMA</b>	<b>175</b>
4.1. Introduction	<b>176</b>
4.2. Aims and hypothesis	<b>179</b>
4.3. Preliminary experiments	<b>181</b>
4.3.1. Standard curves for analysing ATP and lactate production.	<b>181</b>
4.3.2. Measuring oxygen consumption rates using respirometry.	<b>183</b>
4.4. Methods	<b>185</b>
4.4.1. Real-time PCR primers for gene expression and mtDNA copy number analysis	<b>185</b>
4.5. Results	<b>187</b>
4.5.1. MtDNA copy number in multiple myeloma cells and patient tumor samples	<b>187</b>
4.5.2. Oxygen consumption rates in the U266 cell line.	<b>187</b>

4.5.3.	Depletion of mtDNA in U266 cells.	<b>189</b>
4.5.4.	Expression of the mtDNA genes during depletion of U266 cells	<b>190</b>
4.5.5.	Expression of the mitochondrial transcription and replication genes	<b>192</b>
4.5.6.	Expression of the genes regulating multiple myeloma tumor biology	<b>197</b>
4.5.7.	ATP and lactate production during mtDNA depletion	<b>199</b>
4.5.8.	Recovery of mtDNA copy number in the presence of glycolysis inhibitor	<b>200</b>
4.5.9.	Changes in expression to the mtDNA genes during recovery with 2-DG	<b>203</b>
4.5.10.	Expression of the upstream mitochondrial regulators during recovery with 2-DG	<b>206</b>
4.5.11.	Expression of the mitochondrial transcription and replication genes during recovery with 2-DG	<b>209</b>
4.5.12.	Expression of the multiple myeloma tumor genes during recovery with 2-DG	<b>213</b>
4.5.13.	Determining the levels of apoptosis in mtDNA depleted and/or recovered U266 cells	<b>218</b>
4.6.	Discussion	<b>220</b>
4.6.1.	MtDNA copy number in normal and cancer cells	<b>220</b>
4.6.2.	Changes in the gene expression patterns during mtDNA depletion	<b>221</b>

4.6.3. Changes during mtDNA replenishment in the presence and absence of 2-DG	222
4.7. Conclusion	225
<b>CHAPTER 5: IDENTIFICATION OF MTDNA VARIANTS IN GLIOBLASTOMA MULTIFORME.</b>	<b>226</b>
5.1. Introduction	227
5.2. Aims and hypothesis	228
5.3. Methods	230
5.3.1. HRM primers used for the screening of GBM mtDNA variants	230
5.3.2. Culture conditions for the human neural stem cells	231
5.3.3. Normal brain biopsy samples	231
5.3.4. GBM patient tumor samples	231
5.3.5. Xenograft tumor models	231
5.4. Results	233
5.4.1. Identification of mtDNA variants in GBM cell lines	233
5.4.2. Validation of the GBM cell line variants using HRM analyses	237
5.4.3. Susceptibility of mtDNA regions to the development of mutations	239
5.4.4. Protein modelling of the GBM cell line variants	241
5.4.5. Identification of mtDNA variants in human neural stem cells	243
5.4.6. Identification of mtDNA variants in normal brain biopsy samples	246

5.4.7.	Susceptibility of the mtDNA regions to mutations within normal brain biopsy samples	249
5.4.8.	Phylogenetic analysis on the GBM cell lines and normal brain biopsy samples	252
5.4.9.	Identification of mtDNA variants in GBM patient tumor samples	253
5.4.10.	Susceptibility of the mtDNA regions to mutations for GBM patient tumor samples	255
5.4.11.	Protein modelling of the GBM tumor sample variants	260
5.4.12.	Comparative analyses between paired GBM tumor samples	261
5.4.13.	Variants in mtDNA depleted and recovered HSR-GBM1 cells	264
5.4.14.	Critical mtDNA variants required for the tumor biology of GBM	267
5.4.15.	Overview of the mtDNA variants identified across the cells and tissue samples	270
5.5.	Discussion	272
5.5.1.	MtDNA variants in the GBM cell lines, tumors and normal brain tissue	272
5.5.2.	Reliability of the Ion Torrent sequencing results	273
5.5.3.	The impact of GBM mtDNA variants on protein function	273
5.5.4.	MtDNA variants as initiators of GBM	274
5.6.	Conclusion	275
<b>CHAPTER 6: THE ROLE OF MTDNA IN GLIOBLASTOMA MULTIFORME</b>		<b>277</b>

6.1.	Introduction	<b>278</b>
6.2.	Aims and hypothesis	<b>281</b>
6.3.	Methods	<b>283</b>
6.3.1.	Differentiation of HSR-GBM1 cells into neurons and astrocytes	<b>283</b>
6.4.	Results	<b>284</b>
6.4.1.	MtDNA content in HSR-GBM1 cells recovered in conditioned media	<b>284</b>
6.4.2.	Determining the levels of apoptosis in mtDNA depleted and/or recovered HSR-GBM1 cells	<b>285</b>
6.4.3.	MtDNA gene expression changes in depleted HSR-GBM1 cells	<b>287</b>
6.4.4.	Expression of the upstream mitochondrial regulators	<b>287</b>
6.4.5.	Expression of the mitochondrial transcription and replication factors	<b>289</b>
6.4.6.	Expression of the mtDNA genes during recovery of HSR-GBM1 cells	<b>291</b>
6.4.7.	Expression of the upstream mitochondrial regulators during recovery	<b>294</b>
6.4.8.	Expression of the mitochondrial transcription and replication factors during recovery	<b>296</b>
6.4.9.	Changes in gene expression in differentiated astrocytes	<b>299</b>
6.4.10.	Expression of the mtDNA genes after differentiation into astrocytes and neurons	<b>301</b>

6.4.11. Expression of the upstream mitochondrial regulators after differentiation into astrocytes and neurons	302
6.4.12. Expression of the mitochondrial transcription and replication genes after differentiation into astrocytes and neurons	303
6.5. Discussion	306
6.5.1. Gene expression changes in HSR-GBM1 cells recovered in conditioned media	306
6.5.2. Gene expression analysis in astrocytes and neurons differentiated from HSR-GBM1 cells	307
6.6. Conclusion	309
<b>CHAPTER 7: GENERAL DISCUSSION</b>	<b>310</b>
7.1. The role of mtDNA in cancer stem cells and in tumor biology	311
7.2. MtDNA variants	312
7.3. Compensatory mechanisms during mtDNA depletion	313
7.4. The involvement of mtDNA in the initiation of GBM	314
7.5. The initiation of cancer from the transformation of normal cells	316
7.6. The correlation between mtDNA copy number and gene expression levels	317
7.6.1. Multiple myeloma	318
7.6.2. GBM	319
7.7. Bi-directional communication between mitochondria and the nucleus	320
7.8. Further evidence to support the role of mtDNA in the tumor biology of GBM	323

7.9. Summary	324
7.10. Further work	327
References	328
Appendices	352
i. Further information on sequence quality from the Ion Torrent	353
ii. Chemical composition of laboratory reagents	362
iii. Publications	371

## List of Figures

<b>Figure 1.1.</b> The human mitochondrial genome.	<b>31</b>
<b>Figure 1.2.</b> MtDNA replication.	<b>34</b>
<b>Figure 1.3.</b> Overview of the key processes of mitochondrial energetics.	<b>41</b>
<b>Figure 1.4.</b> Summary of the glycolysis pathway.	<b>43</b>
<b>Figure 1.5.</b> Lactic acid fermentation.	<b>44</b>
<b>Figure 1.6.</b> The citric acid cycle.	<b>46</b>
<b>Figure 1.7.</b> Schematic diagram of the mitochondrial ETC.	<b>48</b>
<b>Figure 1.8.</b> Diagram of the Complex III monomer, cytochrome $bc_1$ .	<b>51</b>
<b>Figure 1.9.</b> Cytochrome c oxidase (Complex IV) of the mitochondrial ETC	<b>53</b>
<b>Figure 1.10.</b> The Warburg effect.	<b>56</b>
<b>Figure 2.1.</b> Physical shearing of genomic DNA to produce random sized fragments.	<b>100</b>
<b>Figure 2.2.</b> End repair of blunt-ended DNA molecules.	<b>100</b>
<b>Figure 2.3.</b> DNA purification using the Agencourt AMPure XP kit.	<b>102</b>
<b>Figure 2.4.</b> Nick translation for the ligation of sequencing adaptors.	<b>103</b>
<b>Figure 2.5.</b> Emulsion PCR procedure.	<b>109</b>
<b>Figure 3.1.</b> Long PCR optimisation.	<b>124</b>
<b>Figure 3.2.</b> Long PCR amplification of human mtDNA.	<b>125</b>
<b>Figure 3.3.</b> Mechanical and enzymatic shearing of plasmid DNA libraries	<b>126</b>
<b>Figure 3.4.</b> Sequence coverage of enzymatically sheared MELAS plasmid DNA	<b>127</b>
<b>Figure 3.5.</b> Capillary sequencing of the fibroblast cell line containing the MELAS A3243G mutation.	<b>130</b>

<b>Figure 3.6.</b> Confirmation of MELAS wild-type and mutant inserts in recombinant plasmids.	<b>134</b>
<b>Figure 3.7.</b> HRM analysis of plasmid DNA mixed at specific mutational loads.	<b>135</b>
<b>Figure 3.8.</b> Fluorescent-activated cell sorting (FACS) analysis on U266 cells.	<b>151</b>
<b>Figure 3.9.</b> Location of the T45P mutation using the bovine model of Complex IV	<b>159</b>
<b>Figure 3.10.</b> Location of the CYT B P9L mutation using the bovine model	<b>160</b>
<b>Figure 3.11.</b> Mapping of the CYT B mutations P173L and T174S.	<b>161</b>
<b>Figure 3.12.</b> Modelling of the COX I mutations L150W and H291D in a bovine model	<b>162</b>
<b>Figure 3.13.</b> Mapping of the COX II variants within CD45+ and CD45- cells	<b>163</b>
<b>Figure 3.14.</b> Mapping of COX III variants within CD45+ and CD45- cells	<b>164</b>
<b>Figure 4.1.</b> Key factors in the tumor biology of multiple myeloma	<b>176</b>
<b>Figure 4.2.</b> Standard curve for measuring ATP production	<b>181</b>
<b>Figure 4.3.</b> Standard curve for measuring lactate production	<b>182</b>
<b>Figure 4.4.</b> Determining the optimal concentration of FCCP using respirometry.	<b>183</b>
<b>Figure 4.5.</b> MtDNA copy number for U266 cells and patient tumor samples.	<b>187</b>
<b>Figure 4.6.</b> MtDNA depletion of U266 cells for 30 days using ddC	<b>190</b>
<b>Figure 4.7.</b> Changes to the expression of the mtDNA-encoded genes during depletion	<b>191</b>

<b>Figure 4.8.</b> Gene expression changes to the upstream mitochondrial regulators	<b>193</b>
<b>Figure 4.9.</b> Expression of the mitochondrial transcription and replication genes	<b>195</b>
<b>Figure 4.10.</b> Gene expression changes to key tumorigenic genes during depletion	<b>198</b>
<b>Figure 4.11.</b> Changes to ATP and lactate production during mtDNA depletion	<b>200</b>
<b>Figure 4.12.</b> Recovery of mtDNA copy numbers in U266 cells after mtDNA depletion, cultured in the presence and absence of 2-deoxyglucose	<b>201</b>
<b>Figure 4.13.</b> Expression changes to the mtDNA genes during recovery with 2-DG, after 7 days of mtDNA depletion.	<b>204</b>
<b>Figure 4.14.</b> Expression changes to the mtDNA genes during recovery with 2-DG, after 14 days of mtDNA depletion.	<b>205</b>
<b>Figure 4.15.</b> Expression of the upstream mitochondrial regulators during recovery with 2-DG, after 7 days of depletion.	<b>207</b>
<b>Figure 4.16.</b> Expression of the upstream mitochondrial regulators during recovery with 2-DG, after 14 days of depletion.	<b>208</b>
<b>Figure 4.17.</b> Expression of the mitochondrial transcription and replication genes during recovery with 2-DG, after 7 days of mtDNA depletion.	<b>210</b>
<b>Figure 4.18.</b> Expression of the mitochondrial transcription and replication genes during recovery with 2-DG, after 14 days of mtDNA depletion.	<b>211</b>
<b>Figure 4.19.</b> Expression of the multiple myeloma tumorigenic genes during recovery with 2-DG, after 7 days of mtDNA depletion.	<b>215</b>

<b>Figure 4.20.</b> Expression of the multiple myeloma tumorigenic genes during recovery with 2-DG, after 14 days of mtDNA depletion.	<b>217</b>
<b>Figure 4.21.</b> Western blot analysis showing processivity of OPA1 in mtDNA depleted and/or recovered U266 cells.	<b>219</b>
<b>Figure 5.1.</b> Modelling of the V336M COX I mutation using the bovine model.	<b>241</b>
<b>Figure 5.2.</b> Modelling of the P223T COX II mutation using the bovine model.	<b>242</b>
<b>Figure 5.3.</b> Modelling of the hNSC G136D CYT B mutation using the bovine model.	<b>245</b>
<b>Figure 5.4.</b> Phylogenetic analysis on the GBM cell lines and normal brain biopsy samples	<b>252</b>
<b>Figure 5.5.</b> Modelling of the P305L CYT B mutation using the bovine model.	<b>260</b>
<b>Figure 5.6.</b> Modelling of the L168S COX II mutation using the bovine model.	<b>261</b>
<b>Figure 5.7.</b> The correlation between mtDNA depletion and the gain of mtDNA variants in HSR-GBM1 cells recovered in BALB/c nude mice.	<b>267</b>
<b>Figure 6.1.</b> MtDNA content in HSR-GBM1 cells depleted of their mtDNA and recovered in conditioned media	<b>284</b>
<b>Figure 6.2.</b> Western blot analysis for OPA1 expression in mtDNA depleted and/or recovered HSR-GBM1 cells.	<b>286</b>
<b>Figure 6.3.</b> Expression analysis on key mtDNA genes in depleted HSR-GBM1 cells	<b>287</b>

<b>Figure 6.4.</b> Gene expression of the upstream mitochondrial regulators in depleted HSR-GBM cells	<b>288</b>
<b>Figure 6.5.</b> Expression of the mitochondrial transcription and replication factors	<b>290</b>
<b>Figure 6.6.</b> Expression of the mtDNA genes during recovery of 7-day depleted HSR-GBM1 cells	<b>292</b>
<b>Figure 6.7.</b> Expression of the mtDNA genes during recovery of 14-day depleted HSR-GBM1 cells	<b>293</b>
<b>Figure 6.8.</b> Expression of the upstream mitochondrial regulators during recovery of 7-day depleted HSR-GBM1 cells	<b>294</b>
<b>Figure 6.9.</b> Expression of the upstream mitochondrial regulators during recovery of 14-day depleted HSR-GBM1 cells	<b>295</b>
<b>Figure 6.10.</b> Expression of the mitochondrial transcription and replication genes during recovery of 7-day mtDNA depleted HSR-GBM1 cells	<b>297</b>
<b>Figure 6.11.</b> Expression of the mitochondrial transcription and replication genes during recovery of 14-day mtDNA depleted HSR-GBM1 cells	<b>298</b>
<b>Figure 6.12.</b> Expression analysis for the mitochondrial maintenance genes in astrocytes	<b>300</b>
<b>Figure 6.13.</b> Expression of the mtDNA genes after differentiation into astrocytes and neurons	<b>301</b>
<b>Figure 6.14.</b> Expression of the upstream mitochondrial regulators after differentiation into astrocytes and neurons	<b>302</b>
<b>Figure 6.15.</b> Expression of the mitochondrial transcription and replication genes after differentiation into astrocytes and neurons	<b>305</b>

<b>Figure 7.1.</b> Model of bi-directional communication between the nucleus and mitochondria.	<b>321</b>
<b>Figure 7.2.</b> Proposed model for the role of mtDNA in the tumor biology of multiple myeloma and GBM cells.	<b>326</b>
<b>Figure A1.</b> Quality report on the raw sequencing data obtained from the Ion Torrent Suite Software	<b>354</b>
<b>Figure A2.</b> Read length distribution for the S219 myeloma sample before and after trimming	<b>355</b>
<b>Figure A3.</b> Distribution of trimmed read lengths for the S219 sample after mapping	<b>357</b>
<b>Figure A4.</b> Sequence coverage analysis using the myeloma S219 mapped alignment	<b>358</b>
<b>Figure A5.</b> Examining the relationship between the percentage numbers of sequence reads carrying a particular GC composition (S219)	<b>359</b>
<b>Figure A6.</b> Examining the relationship between sequence coverage and the percentage GC content in the S219 myeloma sample	<b>360</b>
<b>Figure A7.</b> Overall profile for sequence coverage across the mtDNA of the S219 myeloma sample	<b>361</b>

## List of Tables

<b>Table 1.1.</b> MtDNA variants reported in colorectal, breast and thyroid cancer.	<b>60</b>
<b>Table 2.1.</b> Sequences of the Ion ligation adaptors.	<b>103</b>
<b>Table 3.1.</b> Comparison of the automated and manual protocols for minimizing experimental variability	<b>129</b>
<b>Table 3.2.</b> Ion Torrent sequencing on a human fibroblast cell line containing the MELAS A3243G mutation.	<b>131</b>
<b>Table 3.3.</b> Ion Torrent sequencing on the human BJ fibroblast cell line.	<b>133</b>
<b>Table 3.4.</b> Determining the sensitivity levels on the Ion Torrent.	<b>136</b>
<b>Table 3.5.</b> Primers used for multiple myeloma variant screening on the HRM.	<b>138</b>
<b>Table 3.6.</b> MtDNA variants detected from multiple myeloma tumors.	<b>140</b>
<b>Table 3.7.</b> Screening additional multiple myeloma tumors using HRM.	<b>141</b>
<b>Table 3.8.</b> Susceptibility of mtDNA regions to the development of mutations.	<b>144</b>
<b>Table 3.9.</b> Fold change differences on the susceptibility of mtDNA regions to mutation.	<b>145</b>
<b>Table 3.10.</b> Identification of variants during mtDNA replenishment in U266 cells.	<b>146</b>
<b>Table 3.11.</b> Susceptibility of depleted and recovered U266 cells to mutation	<b>149</b>
<b>Table 3.12.</b> Fold change differences for the susceptibility of depleted and recovered U266 cells to the development of mutations.	<b>150</b>

<b>Table 3.13.</b> MtDNA variants in U266 sub-populations of CD45+ and CD45- cells	<b>152</b>
<b>Table 3.14.</b> Confirmation of CD45+ and CD45- variants using HRM analysis	<b>154</b>
<b>Table 3.15.</b> Determining the susceptibility of CD45+ and CD45- cells to mutations	<b>155</b>
<b>Table 3.16.</b> Fold change differences for the susceptibility of CD45+ and CD45- cells to mutation	<b>156</b>
<b>Table 3.17.</b> Calculating the GC content of each mtDNA variant region	<b>157</b>
<b>Table 4.1.</b> Primers used to examine the expression of genes and mtDNA copy number	<b>185</b>
<b>Table 4.2.</b> Respirometry measurements on U266, HSR-GBM1 and HepG2 cells.	<b>188</b>
<b>Table 5.1.</b> Primers used for GBM variant screening on the HRM	<b>230</b>
<b>Table 5.2.</b> MtDNA variants identified in the non-coding region of 12 GBM cell lines.	<b>234</b>
<b>Table 5.3.</b> MtDNA variants identified in the coding region of 12 GBM cell lines.	<b>236</b>
<b>Table 5.4.</b> Validation of the GBM cell line variants using HRM analyses.	<b>238</b>
<b>Table 5.5.</b> Susceptibility of mtDNA regions to the development of mutations.	<b>240</b>
<b>Table 5.6.</b> Fold change differences on the susceptibility of mtDNA regions to mutation.	<b>240</b>
<b>Table 5.7.</b> Identification of mtDNA variants in human neural stem cells	<b>244</b>

<b>Table 5.8.</b> Identification of mtDNA non-coding variants in normal brain biopsy samples	<b>247</b>
<b>Table 5.9.</b> Identification of mtDNA coding variants in normal brain biopsy samples	<b>248</b>
<b>Table 5.10.</b> Susceptibility of the mtDNA regions to mutations within normal brain biopsy samples	<b>250</b>
<b>Table 5.11.</b> Fold change differences on the susceptibility of mtDNA regions to mutation, for normal brain biopsy samples	<b>251</b>
<b>Table 5.12.</b> Identification of mtDNA non-coding variants in GBM patient tumor samples	<b>254</b>
<b>Table 5.13.</b> Identification of mtDNA coding variants in GBM patient tumor samples	<b>256</b>
<b>Table 5.14.</b> Susceptibility of the mtDNA regions to mutation for GBM patient tumor samples	<b>258</b>
<b>Table 5.15.</b> Fold change differences on the susceptibility of mtDNA regions to mutation, for GBM patient tumor samples	<b>259</b>
<b>Table 5.16.</b> Identification of mtDNA variants in paired samples from 3 individual GBM patients	<b>263</b>
<b>Table 5.17.</b> Identification of mtDNA variants in depleted HSR-GBM1 cells that have been recovered in BALB/c nude mice, using HRM analysis.	<b>265</b>
<b>Table 5.18.</b> Identification of mtDNA variants in human GBM tumor samples, using HRM analysis	<b>269</b>
<b>Table 5.19.</b> Table showing an overview of the sequencing results from the GBM cells, tumors and normal brain tissue	<b>270</b>
<b>Table (i).</b> Detailed sequencing results before and after trimming (S219)	<b>356</b>

## **Acknowledgments**

First, I would like to thank Professor Justin St John for his help and support throughout the duration of my PhD studies. I would also like to thank Professor Dimitris Grammatopoulos and Monash University for providing me with the opportunity to train overseas at the Monash Institute of Medical Research (MIMR) under the guidance of Professor Justin St John.

I am grateful to previous lab members Dr Richard Kelly and PhD student Adam Dickinson for their technical help with learning molecular and cellular biology techniques. I am also grateful to Jacqueline Johnson for sharing her expertise in cell culture. I would also like to thank Ms Vivien Vasic, Ms Jodee Gould and Dr Trevor Wilson at the Australian Cancer Research Foundation (ACRF) Centre for Cancer Genomic Medicine for their help and support with the next generation sequencing work, especially for their patience with the optimisation of the Ion Torrent workflow for sequencing of the mitochondrial genome. I am also grateful to our collaborators at The Alfred Hospital and at the Centre for Cancer Research (MIMR) for supplying samples for parts of my PhD project.

Finally, I am grateful to the Medical Research Council, UK, for providing me with a prestigious postgraduate scholarship and funding to support my 3 years of PhD training, and to Monash University for awarding me with the 2013 MIMR Education and Training Grant. Both of these financially supported my travels to the Stem Cell Biology 2013 Meeting at Cold Spring Harbor Laboratory and the Mitochondrial Disease: Translating Biology to New Treatments 2013 meeting at the Wellcome Trust Conference Centre, University of Cambridge.

## **Declaration**

I declare that this thesis is my own work except where it contains work based on collaborative research, in which case the contributions of others are stated. All materials contained in this thesis, which have been derived from the published or unpublished work of others, are acknowledged.

I confirm that this thesis presented to the University of Warwick has not been submitted previously for a degree at another university.

Ka Yu Yeung

April 2014

## Abstract

Cancer cells preferentially metabolise glucose via aerobic glycolysis (the Warburg effect), which is less energy efficient in terms of ATP production compared to oxidative phosphorylation (OXPHOS). Mitochondrial DNA (mtDNA) encodes proteins of the electron transfer chain and is crucial for functional OXPHOS. MtDNA exists as multiple copies in cells and, often in cancer, there is co-existence of mutant and wild-type mtDNA. There is evidence for mitochondria to contribute towards the tumor biology of multiple myeloma (MM) and glioblastoma multiforme (GBM). The mtDNA from both these cancer types were explored to determine its role in tumor biology.

Sequencing of MM cells and tumor samples using the Ion Torrent next generation sequencer identified Cytochrome C Oxidase and ATP 6 to contain critical variants that are capable of disrupting protein function. Gene expression analysis determined that glycolysis is essential to maintaining MM cell proliferation. Without glycolysis, there was up-regulation in the expression of tumor survival genes, which was only effective in MM cells that had sufficient mtDNA copy numbers above the mtDNA set point.

Sequencing of GBM cell lines, tumor and normal patient samples suggested that there is a predisposition of GBM tumors to acquire a set of GBM-specific mtDNA variants during tumor development. Conserved mtDNA regions, such as Cytochrome C Oxidase I, tend to be least susceptible to mutations. The presence of variants in these conserved regions carry more detrimental effects at the protein-level than at other mtDNA regions. Differentiation of GBM cells decreased the tumor phenotype, as assessed by gene expression analysis.

Altogether, this thesis provides support for the importance of mtDNA in tumor biology. The implications are that the variants identified could be used to screen MM and GBM tumors in a clinical diagnostic lab for the treatment of both these cancer types.

## Abbreviations

Acetyl-CoA	Acetyl-Coenzyme A
ADP	Adenosine Diphosphate
ANOVA	Analysis of Variance
ATP	Adenosine Triphosphate
CASP	Caspase
COX	Cytochrome C Oxidase
CYT B	Cytochrome B
DAD	Defender Against Death
ddC	Dideoxycytidine
2-DG	2-Deoxyglucose
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic Acid
EGF	Epidermal Growth Factor
ELISA	Enzyme-Linked Immunosorbent Assay
EMBL	The European Molecular Biology Laboratory
EMP	Embden-Meyerhof-Parnas
ETC	Electron Transfer Chain
FACS	Fluorescent-Activated Cell Sorting
FADH	Flavin Adenine Dinucleotide
FBS	Fetal Bovine Serum
FCCP	Trifluorocarbonylcyanide Phenylhydrazine
FGF	Fibroblast Growth Factor
GBM	Glioblastoma Multiforme
GITC	Guandine Isothiocyanate

GuCl	Guandine Hydrochloride
GSH	Glutathione
HIF	Hypoxia-Inducible Factor
HRM	High-Resolution Melt
HSP	Heavy-Strand Promoter
IMP	Insulin-like Growth Factor 2 mRNA Binding Protein 2
IL6	Interleukin 6
ISP	Ion Sphere Particles
IUPAC	International Union of Pure and Applied Chemistry
JAK	Janus-Activated Kinase
LB	Luria-Bertani
LSP	Light-Strand Promoter
MAPK	Mitogen-Activated Protein Kinase
MCL1	Myeloid Cell Leukemia 1
MgCl <sub>2</sub>	Magnesium Chloride
MMRF	Multiple Myeloma Research Foundation
MtDNA	Mitochondrial DNA
MTERF	Mitochondrial Specific Termination Factor
MTSSB	Mitochondrial Single-Stranded Binding Protein
NADH	Nicotinamide Adenine Dinucleotide
NaOH	Sodium Hydroxide
NCBI	National Center for Biotechnology Institute
ND	NADH Dehydrogenase
NEAA	Non-Essential Amino Acids
NRF	Nuclear Respiratory Factor

NSC	Neural Stem Cells
OXPPOS	Oxidative Phosphorylation
PBS	Phosphate-Buffered Saline
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PEP	Phosphoenolpyruvate
PGC1 $\alpha$	Peroxisome-Proliferator-Activated Receptor Gamma Co-activator-1-Alpha
PGM	Personal Genome Machine
PKM2	Protein Kinase M2
POLG	Polymerase Gamma
PPAR $\delta$	Peroxisome-Proliferator-Activated Receptor Gamma
RCSB	Research Collaboratory for Structural Bioinformatics
RITOLS	RNA-Incorporated Throughout the Lagging Strand
RNA	Ribonucleic Acid
ROS	Reactive Oxygen Species
rRNA	Ribosomal RNA
SOC	Super Optimal Broth with Catabolite Repression
SEM	Standard Error of the Mean
SIRT3	Sirtuin 3
SNP	Single Nucleotide Polymorphism
STAT3	Signal Transducer and Activator of Transcription 3
TE	Tris-EDTA
TFAM	Mitochondrial Transcription Factor A
TFB1M	Mitochondrial Transcription Factor B1

TFB2M	Mitochondrial Transcription Factor B2
tRNA	Transfer RNA
TWINKLE	Progressive External Ophthalmoplegia 1
UV	Ultraviolet
WHO	World Health Organisation
XBP1	X-Box Binding Protein 1
X-GAL	5-Bromo-4-Chloro-3-Indolyl $\beta$ -D-Galactopyranoside

# **Chapter 1: General introduction**



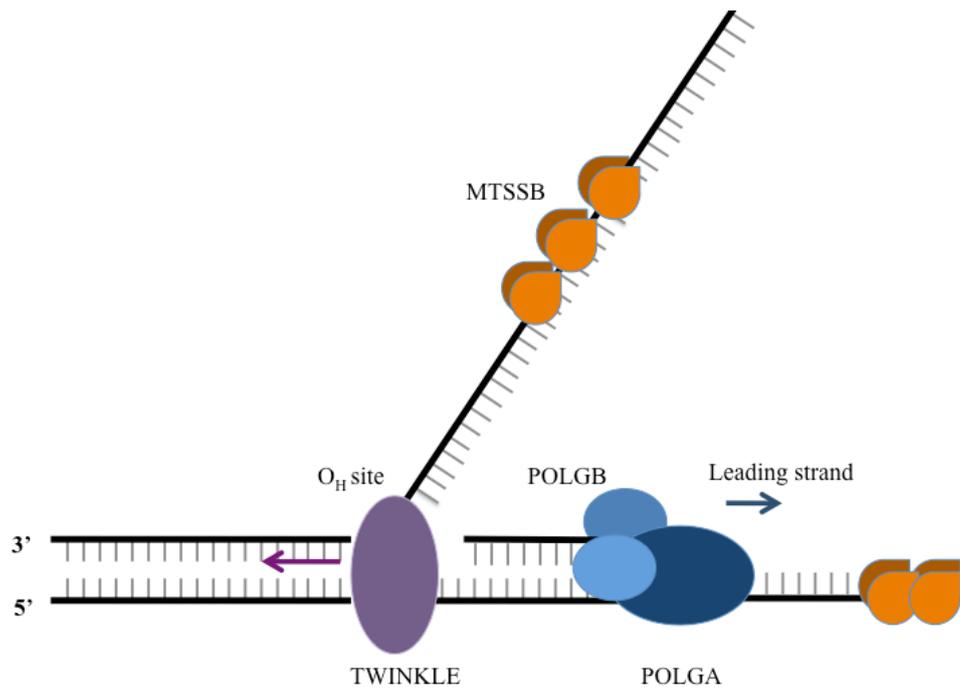
As a consequence of its small genome size, mitochondria rely on the import of nuclear-encoded subunits to complete the assembly of the ETC protein complexes. Complex II is the only protein of the ETC that is exclusively assembled from nuclear-encoded factors. These observations emphasize the necessity of communication between the nucleus and mitochondria in maintaining ETC function<sup>1,2</sup>.

MtDNA biogenesis, transcription and replication all require the activities of nuclear-encoded mitochondrial factors, which must be imported into the mitochondrion. These nuclear genes are either translated by ribosomes in the cytosol<sup>3</sup>, or co-translated during the import process into the mitochondrion<sup>4</sup>. Either way, these precursor proteins all contain a mitochondrial targeting signal, usually located on the N-terminus<sup>5</sup>. These signals are removed upon successful translocation into the organelle, and depending on the specificity of their target sequence, proteins will be directed to one of the four major intracellular compartments: the matrix, the inner membrane<sup>6</sup>, outer membrane<sup>7</sup> or intermembrane space<sup>8</sup>. In most cases, proteins are folded into their native conformation after translocation into the mitochondrion, as unfolded proteins are more able to efficiently pass through the translocase complexes that span across the inner (TIM) or outer (TOM) mitochondrial membrane<sup>9</sup>. The translocation process of these polypeptides is reliant on a supply of ATP<sup>10</sup>.

## 1.2. Maintenance of mtDNA and its mode of replication

Peroxisome-proliferator-activated receptor gamma co-activator-1-alpha (PGC1 $\alpha$ ) is widely considered as the key regulator of mitochondrial biogenesis. For example, the ectopic expression of PGC1 $\alpha$  in myotubes enhances the expression of key mitochondrial transcription factors, such as the nuclear respiratory factor (NRF) and mitochondrial transcription factor A (TFAM) genes<sup>11</sup>. In addition, SIRT3, a target of PGC1 $\alpha$ , which localizes to the mitochondria, can also be considered as the key regulator in mitochondrial biogenesis since SIRT3 mediates the effects of PGC1 $\alpha$  by maintaining mitochondrial biogenesis and reactive oxygen species (ROS) production<sup>12</sup>. PGC1 $\alpha$  expression is controlled by peroxisome-activated receptor  $\delta$  (PPAR $\delta$ )<sup>13</sup> and, once translated, PGC1 $\alpha$  activates nuclear respiratory factors 1 and 2 (NRF1 and NRF2), which stimulate the expression of a range of nuclear genes to promote mitochondrial transcription and replication. The NRFs regulate the expression of mitochondrial transcription factors A (TFAM), B1 (TFB1M) and B2 (TFB2M) by binding to their promoter regions<sup>14</sup>.

TFAM functions together with TFB2M to synthesise an RNA-DNA hybrid primer that can be used by the mitochondrial DNA polymerase complex, polymerase gamma (POLG), during mtDNA replication<sup>15,16</sup>. The POLG complex requires activity from subunits polymerase gamma A (POLGA) and B (POLGB) to exert its full function. POLGA carries catalytic capabilities, whereas POLGB is an accessory subunit supporting the synthesis of a nascent mtDNA strand by assisting in DNA binding and maintaining processivity of the polymerase<sup>17</sup> (**Figure 1.2**).



**Figure 1.2.** MtDNA replication. The nuclear-encoded replication factors work together to ensure efficient replication of mtDNA. The mtDNA specific helicase, TWINKLE, unwinds DNA at the origin of replication located on the heavy strand. The mitochondrial-specific DNA polymerase (POLG) complex utilizes the catalytic (POLGA) subunit, together with the two supporting subunits (POLGB), to synthesise a complementary mtDNA strand in the 5' → 3' direction. The mitochondrial single-stranded binding (MTSSB) proteins prevent the premature re-annealing of mtDNA single strands.

POLG also possesses 3' → 5' exonuclease activity and works closely with the mtDNA-specific helicase, TWINKLE, to separate double-stranded mtDNA into single-stranded mtDNA<sup>18</sup>. During this process, the mitochondrial single-stranded binding protein (MTSSB) is recruited to prevent the premature re-annealing of the two separated mtDNA strands prior to the complete synthesis of a nascent complementary strand from the mtDNA template. Mitochondrial specific termination factor (MTERF) mediates control on attenuating further

transcriptional activity of mtDNA<sup>19</sup>. Altogether, the combined functions of these nuclear-encoded factors are essential for efficient replication of mtDNA. Although TWINKLE has been associated primarily with unwinding of the double-stranded mtDNA during replication, structural analysis performed on TFAM has further suggested that it also possesses the capability to separate double-stranded mtDNA regions for transcription to take place<sup>20</sup>.

Several models for the replication of mtDNA have been proposed. The strand asynchronous method proposed that the initiation of mtDNA replication occurred at the promoter of the heavy strand, and upon successful replication of two-thirds of the mitochondrial genome, synthesis from the light strand promoter subsequently followed<sup>21</sup>. However, the theory was later challenged by others, based on the identification of double-stranded mtDNA regions during the replication process<sup>22</sup>.

The strand-coupled replication mechanism was then proposed, which is similar to the process of nuclear DNA replication<sup>23</sup>. It was thought that, during strand-synchronous mtDNA replication, double-stranded intermediates are produced, which comprise hybrid RNA-DNA regions. This led to the development of the RNA incorporated throughout the lagging strand (RITOLS) theory, which suggested that RNA intermediates formed part of the lagging strand during mtDNA replication<sup>24</sup>. However, based on the evidence to date, it is possible for mtDNA replication to occur via a combination of these mechanisms, with the strand-synchronous method predominating during periods of mtDNA replenishment after exposure to mtDNA stress<sup>25</sup>.

### 1.3. Overview of tumor biology

Tumor development refers to the transformation of normal cells into cancer cells via the accumulation of genetic and cellular changes, which result in the loss of responsiveness of these cells to growth regulatory signals<sup>26</sup>. Tumor growth normally occurs as a result of a net increase in the rates of cellular proliferation compared to the rates of apoptosis. Results from a colorectal cancer study provide evidence to support this, whereby rates of apoptosis and proliferation are both increased during development of the cancer, with cell proliferation having a greater influence on these cells than apoptosis<sup>27</sup>. However, there are some exceptions as with the case of lung cancer where it was suggested that tumor volume is associated with an increase in the rate of apoptosis and a decrease in the rate of proliferation, leading to the development of necrotic tissue<sup>28</sup>.

According to Hanahan and Weinberg, a malignant tumor has several characteristics. Apart from the acquired independence to proliferative and anti-proliferative signals, cells also have the ability to undergo self-renewal and metastasis and can maintain their growth through *de novo* angiogenesis for supplies of oxygen and nutrients to the tumor<sup>29</sup>.

When a cell undergoes division, there are 4 phases through which it must proceed. These are the G1 (cell growth), S (DNA synthesis), G2 (second cell growth) and M (mitotic) phases. In normal cells, there are a number of checkpoints within the cell cycle that function to protect the nuclear genome from DNA damage. The first checkpoint occurs prior to DNA synthesis during the initial growth of the cell in the G1 phase, which assesses for the presence of

sufficient nutrients, growth factors and the absence of DNA damage<sup>30</sup>. The second checkpoint functions between mitosis and DNA synthesis to check for the completion of DNA replication<sup>31</sup>, and another exerts control over the assembly of the spindle to ensure chromosomes are properly attached before cell division. After these checkpoints, the cell cycle is repeated<sup>32,33</sup> or the cell undergoes differentiation by leaving the cell cycle at the G0 resting phase<sup>34</sup>. However, dysregulation of these cell cycle checkpoints is thought to be a common cause of the onset of cancer development<sup>35,36</sup>.

It has been demonstrated that the gene encoding for p53, the protein that regulates progression of the cell cycle at the G1 phase, is often mutated in cancer<sup>37,38</sup>. Loss of function of this critical protein means that incompletely replicated or damaged DNA can progress through the cell cycle to be passed onto progenitor cells. Consequently, these cells will have characteristics of increased genomic instability, which would result in the gain of DNA sequence mutations, thereby potentially initiating the tumorigenic phenotype<sup>39</sup>.

Changes to tumor biology can be caused by the dysfunction to various genes involved in mitochondrial biogenesis. For example, signal transduction via the PGC1 $\alpha$  pathway has been reported to be up-regulated in type I endometrial cancer, compared to normal endometrial tissue<sup>40</sup>. Therefore, regulation of the PGC1 $\alpha$  molecule may be critical for the control of tumor development in some cancers. The role of SIRT3 has also been examined in oral squamous cell carcinoma cell lines, and was discovered to be highly expressed, which correlates with up-regulation in the expression of PGC1 $\alpha$  levels<sup>41</sup>.

From some studies, it has been suggested that mitochondria can influence progression of the cell cycle. For example, loss of the nuclear-encoded complex IV subunit, cytochrome c oxidase Va (CoVa) in *Drosophila melanogaster*, not only decreased ATP production by 60% of original levels, but also inhibited progression of the cell cycle at the G1 phase. This occurred via the activation of MAPK and p53 activity, which decreased levels of cyclin E required for cell cycle progression<sup>42</sup>. Since this discovery, mitochondria have been proposed to act via other signaling pathways. These could be stimulated by the increase in reactive oxygen species leading to the up-regulation of p27 levels, which can modulate progression of the cell cycle by inhibiting activity of the cyclin E-CDK2 complex<sup>43</sup>.

Mitochondria are also capable of secreting the protein survivin into the cytosolic compartment in an attempt to prevent the onset of apoptosis<sup>44</sup>. Defects in the genes that regulate mtDNA replication have been suggested as possible initiators of tumor development, since *POLG* is mutated in 63% of breast cancer cases<sup>45</sup>. Furthermore, the oncogenic protein, c-Myc, responsible for the control of cell size and in the rates of proliferation and apoptosis, is reported to directly regulate TFAM by binding to its promoter region, which could affect mitochondrial transcription<sup>46</sup>.

#### **1.4. Heteroplasmy and mtDNA**

Due to the limited presence of efficient proof-reading mechanisms during mtDNA replication, mtDNA is highly susceptible to the development of

mutations. Multiple mtDNA copies exist within a single mitochondrion, with more metabolically demanding tissues possessing greater numbers of mtDNA copy. In normal and cancer patients, a proportion of these mtDNA copies acquire mutations that exist alongside wild-type copies. This is known as heteroplasmy and is especially common in the pathogenesis of various diseases<sup>47,48</sup>. The presence of mtDNA variants in normal individuals is not unusual, as previous studies have demonstrated via a mathematical modeling approach that random mutations may arise in the absence of selective advantage or tumorigenic influences<sup>49,50</sup>. Such studies have also put forward the possibility that some homoplasmic mutations may have been acquired within mtDNA in this manner.

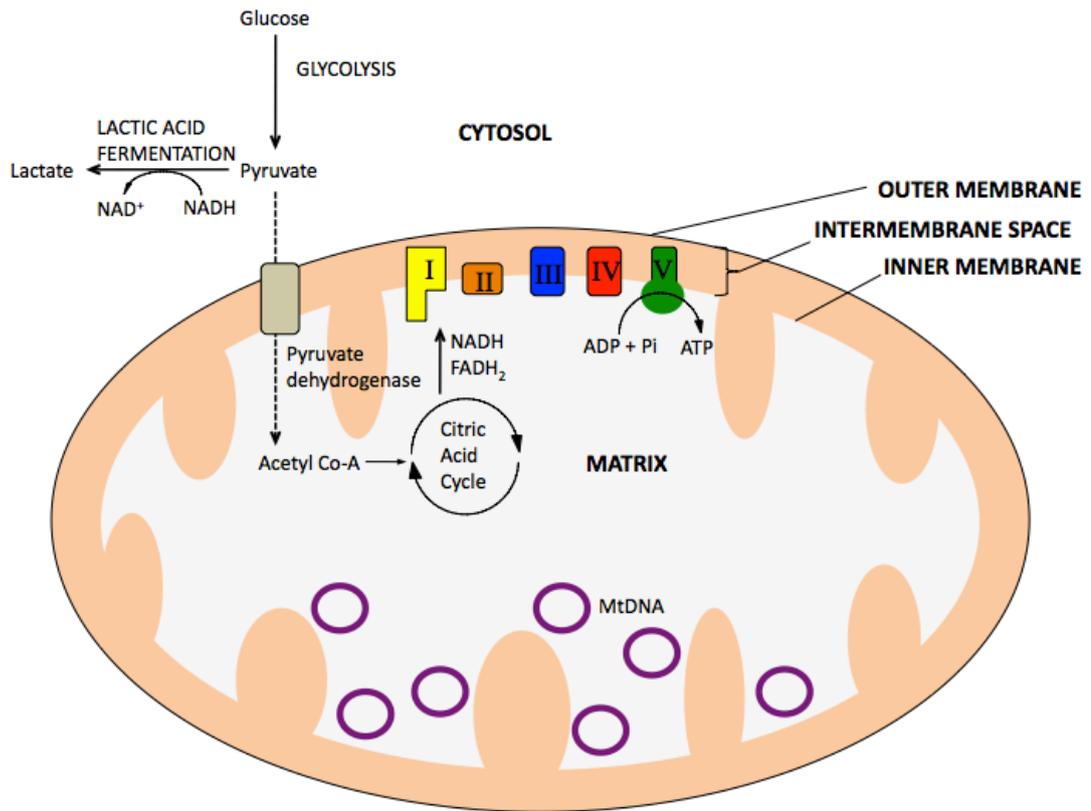
In the last few years, it has been increasingly accepted that normal cells naturally harbour heteroplasmic mutations<sup>51</sup>. However, in these cases, heteroplasmy tends to be reported at higher frequencies in the control regions of mtDNA than in the coding regions<sup>52</sup>. Sequence variants in the control D-loop region usually do not cause any detrimental effects, unless it resides in a region that is crucial for the binding of mitochondrial transcription and replication factors<sup>53</sup>. The diverse variability in the D-loop mtDNA sequences between individuals, explains why sequencing of the D-loop region alone, is commonly performed to determine the different mtDNA haplotypes that define normal individuals of varying ethnic backgrounds<sup>54-56</sup>.

Heteroplasmy in the coding regions of mtDNA is commonly thought to compromise the ability of the ETC to function, and therefore energy production required to maintain and support cellular metabolism. The concept of having a

critical threshold level for mtDNA mutations was first suggested by Howell in 1983, who hypothesized that beyond a certain proportion of mutant mtDNA copies, energy production via the electron transport chain declines to the point where cellular metabolism is also compromised<sup>57</sup>. Although, cybrid fusion experiments have supported the concept that mutant mtDNA enhances tumorigenicity of cancers, for example, in prostate cancer<sup>58</sup>, contradicting evidence remains for the involvement of mtDNA in tumor development<sup>59</sup>.

### **1.5. Mitochondrial energetics**

The key processes of mitochondrial energetics, which cover the pathways of glycolysis, the citric acid cycle and OXPHOS, are briefly described to outline the metabolic events taking place within normal cells (**Figure 1.3**). The processes of lactic acid fermentation and a discussion of the Warburg effect are also presented to explain the altered form of metabolism that is preferred by cancer cells.



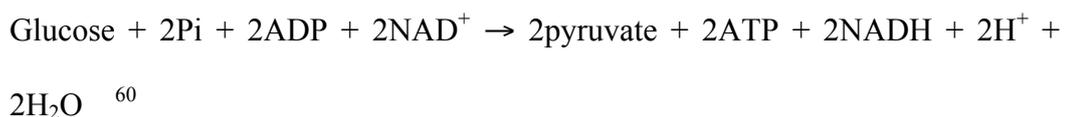
**Figure 1.3.** Overview of the key processes of mitochondrial energetics. Glycolysis takes place in the cytosol. The product pyruvate is imported into the mitochondrial matrix, where it is converted to acetyl coenzyme A (acetyl-coA). Acetyl-co A initiates the citric acid cycle to supply reducing factors  $\text{NADH}$  and  $\text{FADH}_2$  to the mitochondrial electron transfer chain, located on the inner mitochondrial membrane. ATP is generated from complex V of the electron transfer chain.

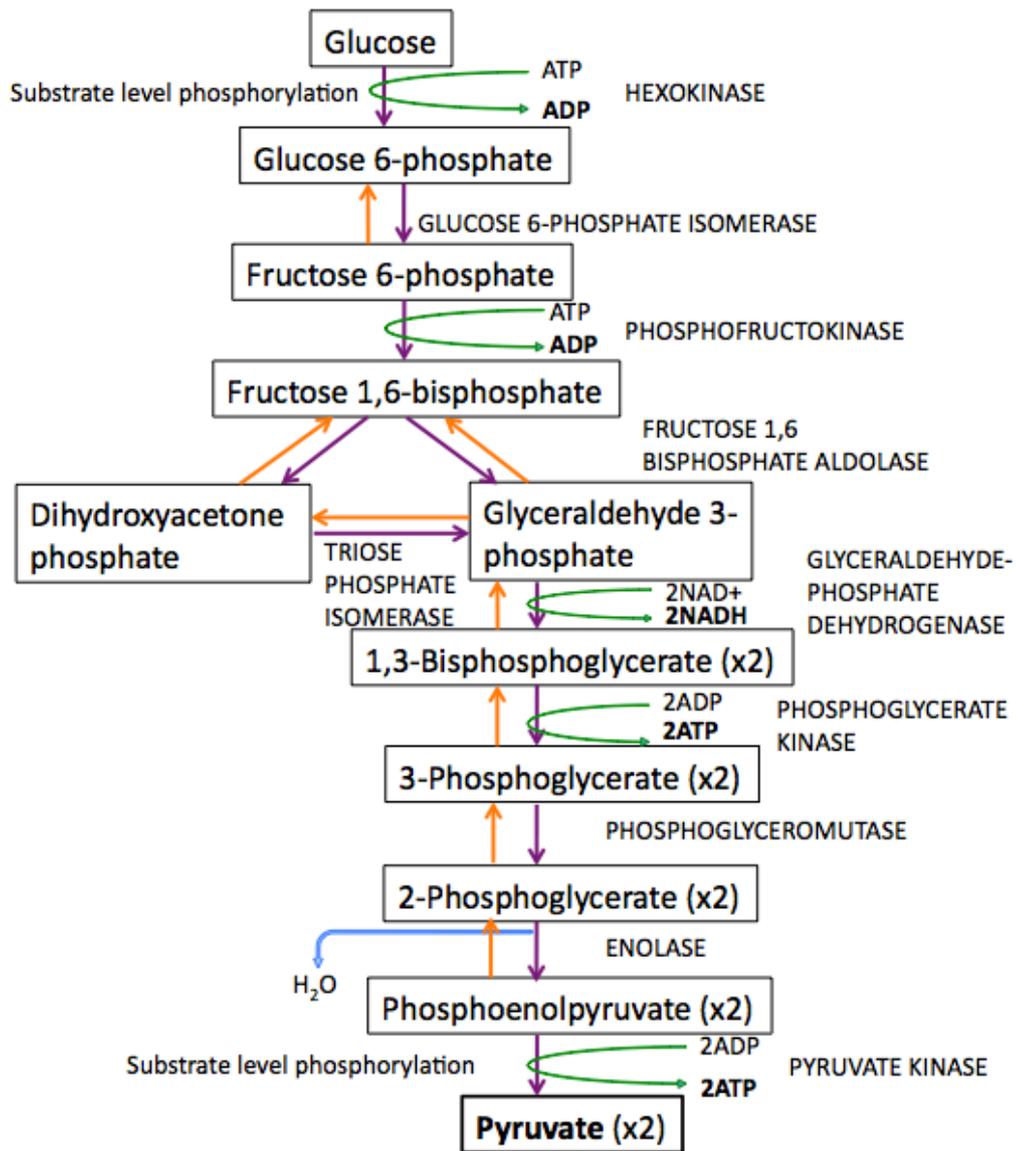
### 1.5.1. Glycolysis

Glycolysis (Embden-Meyerhof-Parnas (EMP) pathway) is a catabolic pathway commonly used in eukaryotes. During glycolysis, glucose enters cells in its unphosphorylated form since glucose 6-phosphate carries a negative charge that prevents its facilitated transport through the cell membrane<sup>60</sup>. Phosphorylation of glucose within the cell assists in its metabolism, as the presence of a phosphate

group causes instability of the glucose molecule<sup>60</sup>. In mammals, there are four types of hexokinases (I to IV), with only the subtype IV carrying substrate specificity for glucose<sup>61</sup>. Catalysis by hexokinase represents an irreversible reaction, responsible for the initial step of converting glucose to glucose-6-phosphate in the presence of ATP (**Figure 1.4**).

A series of reactions follow to produce fructose 1,6-bisphosphate, which undergoes a cleavage reaction via a class I aldolase to form two triose sugars: glyceraldehyde 3-phosphate and dihydroxyacetone phosphate<sup>62</sup>. Glyceraldehyde 3-phosphate can be further metabolised in the glycolytic pathway, which means that dihydroxyacetone phosphate must be converted to this isoform in order to be fully oxidized to pyruvate<sup>62</sup>. At this point, there are two molecules of glyceraldehyde 3-phosphate. The enzymatic steps that follow produce 4 ATP and 2 molecules of nicotinamide adenine dinucleotide (NADH). Therefore, glycolysis produces a net synthesis of 2 ATP molecules for each glucose molecule metabolized. The 2 NADH reducing factors generated can be used to donate electrons during OXPHOS to help establish a proton gradient across the mitochondrial inner membrane for ATP production. The overall reaction of glycolysis can be summarized in the following formula and in **Figure 1.4**, where all the metabolic intermediates from the pathway are shown:



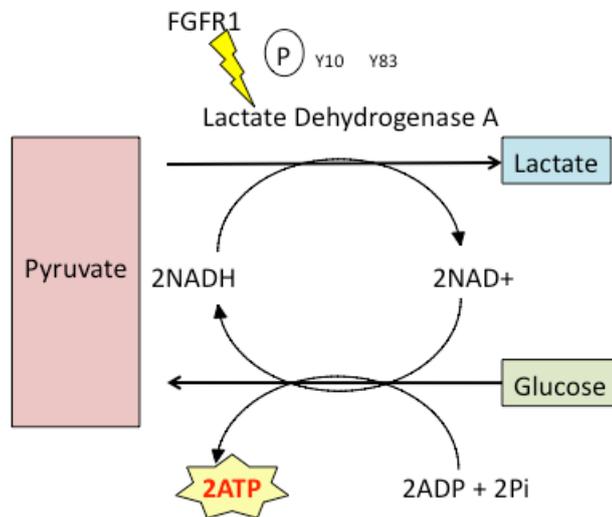
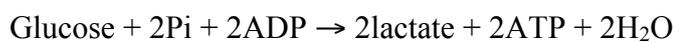


**Figure 1.4.** Summary of the glycolysis pathway. Glucose is metabolised to produce 2 molecules of pyruvate. Pyruvate can then be used as a substrate for the citric acid cycle upon conversion to acetyl coenzyme A in the presence of sufficient oxygen. In the absence of oxygen, pyruvate undergoes fermentation to produce lactate.

### 1.5.2. Lactic acid fermentation

Under low oxygen concentration, normal cells attempt to regenerate NAD<sup>+</sup> to be used as an electron acceptor in the glycolysis pathway. This occurs via the

catalysis of pyruvate into lactate via lactate dehydrogenase A, which is dependent on NADH produced during glycolysis. In cancer development, the receptor tyrosine kinase, FGFR1, is often over-expressed<sup>63</sup>. Its role is in the phosphorylation of lactate dehydrogenase A at tyrosine residues 10 and 83 of the protein structure<sup>64</sup>. This augments the catalytic ability of the enzyme and increases its binding capabilities to NADH. The reaction for the generation of lactate is represented in the formula below and is summarized in **Figure 1.5**:



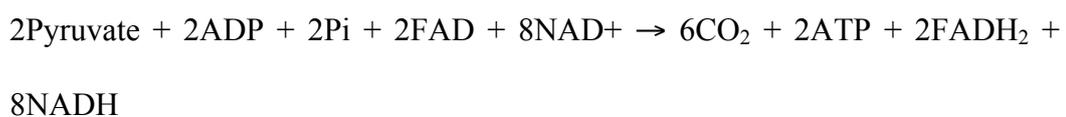
**Figure 1.5:** Lactic acid fermentation. Under conditions of limiting oxygen concentrations, pyruvate is converted into lactate via the enzyme lactate dehydrogenase A, which re-generates NAD<sup>+</sup> and stimulates further cycles of glycolysis to produce low levels of ATP. In cancer development, the over-expression of FGFR1 attaches phosphate groups to lactate dehydrogenase A at 2 positions on tyrosine 10 and 83 to increase the efficiency of catalysis.

In the presence of sufficient oxygen tension, there is a preference for normal cells to undergo further metabolism via the citric acid cycle rather than through lactic acid fermentation.

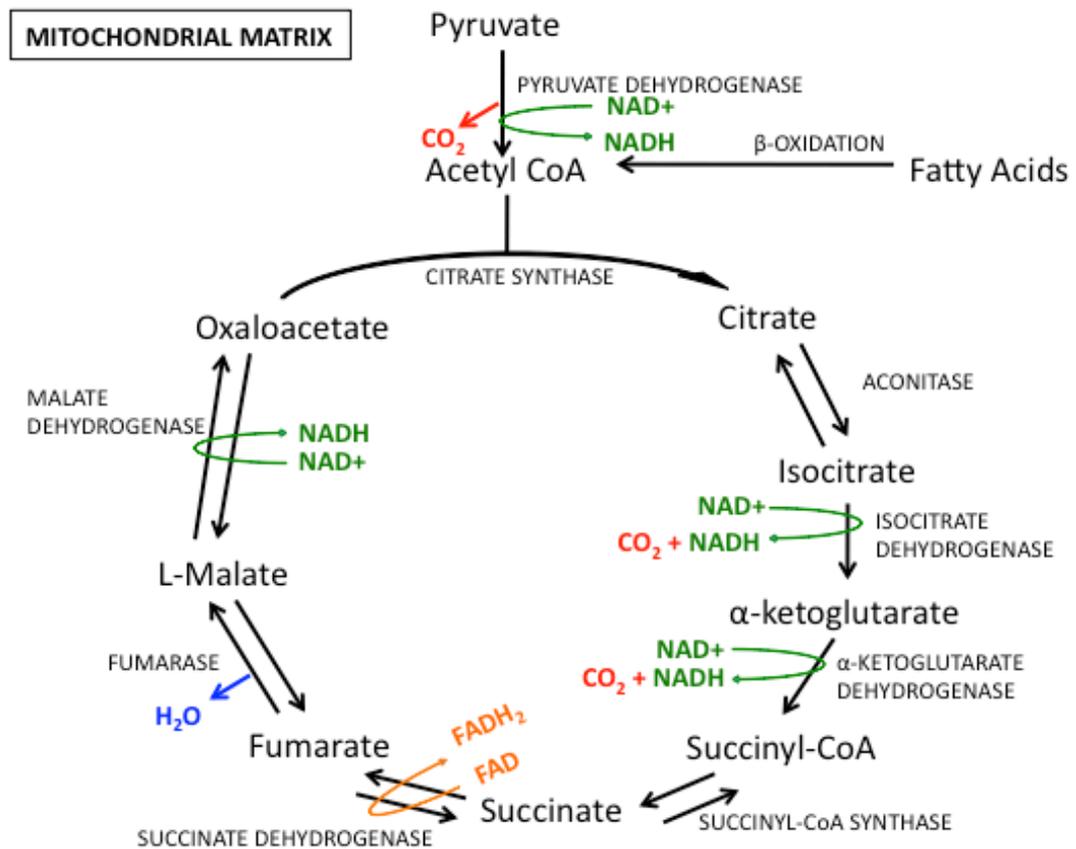
### 1.5.3. Citric acid cycle

Pyruvate is transported into the matrix of the mitochondrion, where it is subjected to the catalytic actions of pyruvate dehydrogenase to decarboxylate pyruvate to produce acetyl coenzyme A (acetyl Co-A) that can be used as a substrate for initiating reactions of the citric acid cycle<sup>65</sup>. Decarboxylation of acetyl-CoA results in the release of CO<sub>2</sub> as a product of the reaction. The first step of the citric acid cycle involves the combination of acetyl Co-A with oxaloacetate, a four-carbon molecule, to generate citrate and waste product coenzyme A (CoA) (**Figure 1.6**)<sup>66</sup>.

Completion of the citric acid cycle produces a number of reducing molecules including NADH and FADH<sub>2</sub> that are used during OXPHOS. The only reactions that are irreversible within the citric acid cycle are those catalysed by citrate synthase, isocitrate dehydrogenase and  $\alpha$ -ketoglutarate dehydrogenase<sup>67</sup>. The reactions that summarise the citric acid cycle are represented according to the following formula:



All intermediates produced upon completion of the citric acid cycle are shown in **Figure 1.6**.



**Figure 1.6.** The citric acid cycle. Oxidation of acetyl-Coenzyme A (acetyl-CoA) results in the production of reducing factors NADH and FADH<sub>2</sub> that can be used as high energy electron donors during the process of OXPHOS via the electron transfer chain.

The intermediate, citrate, has been suggested as a potential therapy for cancer<sup>68</sup>. It can be administered orally to cancer patients, and functions by inhibiting the actions of phosphofructokinase<sup>69</sup>, thereby blocking glycolysis. It is thought that, by inhibiting glycolysis, mitochondria will be forced to utilize other substrates, such as amino acids in place of glucose for their metabolic requirements<sup>70</sup>. The intermediates formed from the metabolism of amino acids could be used to stimulate the citric acid cycle, which would generate energy via OXPHOS<sup>71</sup>.

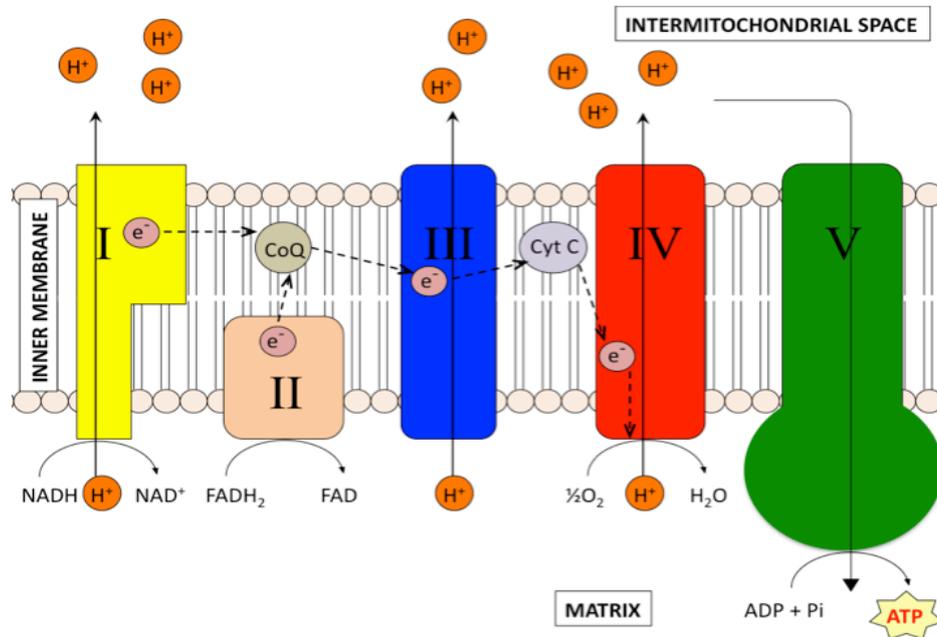
Pyruvate dehydrogenase activity induced by the oncogene BRAF<sup>V600E</sup> is frequently mutated in various cancers inclusive of melanoma. Recently, pyruvate dehydrogenase was thought to contribute to the onset of senescence in tumor cells<sup>72</sup>. It was proposed that activation of the enzyme increases the metabolism of pyruvate in the citric acid cycle to enhance respiration, which results in the increased production of reactive oxygen species<sup>72</sup>. This finding suggests that the over-activation of certain enzymes responsible for catalysis during the citric acid cycle could be detrimental to cancer cells.

On the other hand, recently, it was reported that gene expression levels of aconitase, a different enzyme of the citric acid cycle, which catalyses the conversion of citrate to isocitrate, is up-regulated under conditions of hypoxia, as observed in prostate cancer cells<sup>73</sup>. This contradicts the situation with pyruvate dehydrogenase, as up-regulation of aconitase, in this case, appears to support the growth of these tumor cells. Together, these results suggest that changes to the expression of different citric acid cycle enzymes can exert different effects upon the tumorigenicity of cancer cells.

#### **1.5.4. Oxidative phosphorylation**

The electron transfer chain is situated on the mitochondrial inner membrane and comprises 5 protein complexes (**Figure 1.7**). Complex II comprises entirely of nuclear-encoded subunits, whereas Complexes I, III, IV and V are assembled through a combination of nuclear-encoded and mtDNA- encoded subunits<sup>74</sup>. NADH is used as an electron donor to Complexes I, III and IV of the

mitochondrial electron transfer chain, whereas  $\text{FADH}_2$  supplies electrons to Complexes II, III and IV<sup>75</sup>.



**Figure 1.7.** Schematic diagram of the mitochondrial electron transfer chain (ETC). The sequential transfer of electrons across the 5 protein complexes reduces a molecule of oxygen to generate water at Complex IV within the mitochondrial matrix. During this process, an electrochemical gradient is formed across the mitochondrial inner membrane through the accumulation of protons in the intermembrane space. The electrochemical gradient is then used by Complex V to generate ATP from ADP and inorganic phosphate.

The sequential transfer of electrons from one complex to another, results in the translocation of protons into the intermembrane space from the mitochondrial matrix. However, Complex II, succinate dehydrogenase, which oxidizes  $\text{FADH}_2$  does not directly contribute to the generation of the proton gradient<sup>76</sup>. The electrochemical gradient is maintained as protons are unable to diffuse freely

across the mitochondrial membrane, relying instead on specific proton channels, otherwise known as the ATP synthase protein complex (Complex V). The ATP synthase utilizes this electrochemical gradient to produce ATP from ADP and inorganic phosphates<sup>77</sup>. However, excessive accumulation of the proton gradient may prevent further activity of the ETC, since this would stimulate inhibition of any further proton translocations as well as inhibit the oxidation of NADH<sup>75</sup>.

The coupling between activities of the ETC and the production of ATP via ATP synthase is referred to as OXPHOS. Proper functioning of the citric acid cycle and OXPHOS is important in, for example, the development of mouse ovarian follicles<sup>78</sup>. Decreased ability of any protein complex to function, impacts on the overall ability of the ETC to produce ATP. The first nuclear-encoded mutation identified in Complex I corresponded with the diagnosis of Leigh syndrome<sup>79</sup>. Since then, other Complex I disorders have been reported, which include Parkinson's disease caused by the presence of excessive oxidative stress, thereby resulting in the misassembly of Complex I<sup>80</sup>.

On the other hand, deficiency of Complex II has been associated with patients diagnosed with late-onset optic atrophy and ataxia<sup>81</sup>. Complex III disorders are thought to give rise to mitochondria myopathy<sup>82,83</sup>, and Complex IV defects to the progression of leukodystrophy<sup>84</sup>. Mis-assembly of one protein complex may impact upon the respiratory activity of another, a situation that arises in 25% of OXPHOS deficiencies<sup>85</sup>. This is exemplified by the requirement for Complex III to be fully assembled for the stability of Complex I, regardless of whether Complex III is functional or not<sup>86</sup>.

Glycolysis and the citric acid cycle collectively generate a total of 10 NADH and 2 FADH<sub>2</sub> reducing molecules. The whole process of glucose metabolism can therefore be illustrated using the following equation:



#### **1.5.4.1. Protein modeling analysis of the ETC proteins**

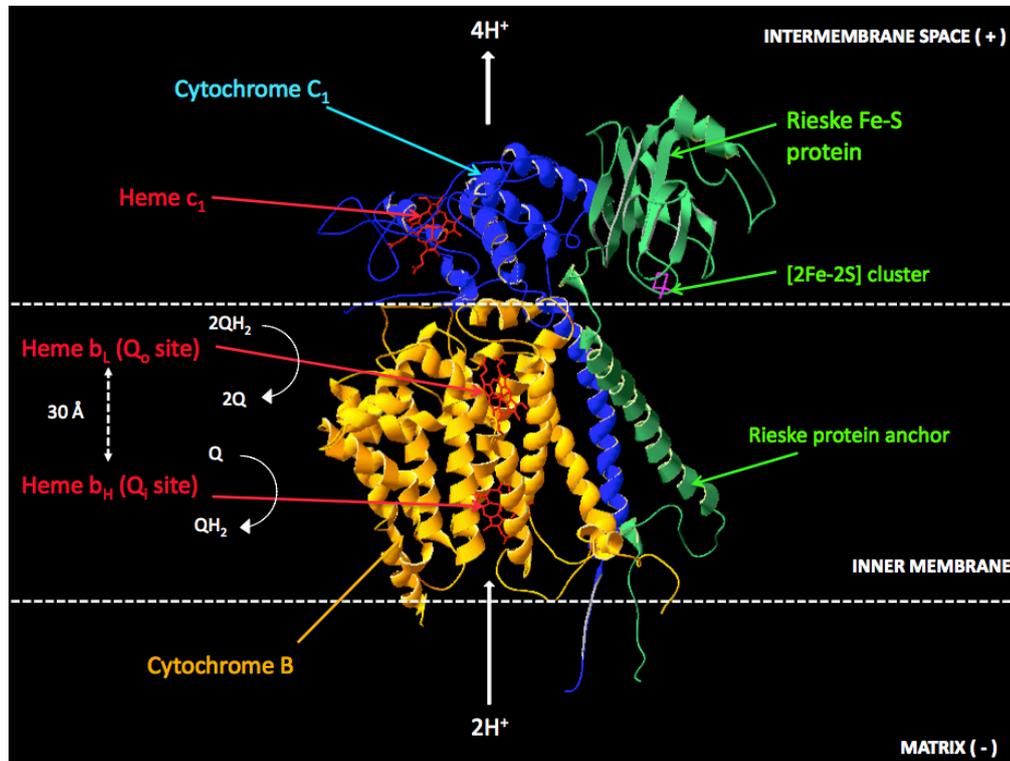
Currently, the molecular structures of human mitochondrial proteins have not yet been determined; as such protein structures from species closely related to human (sharing protein homology of over 70%) were used to identify the location of specific mutations. These proteins were identified through the Protein Data Bank (PDB) database, and the structures were downloaded from The Research Collaboratory for Structural Bioinformatics (RCSB) PDB database to be examined on the Swiss-PDB viewer (v.4.1.0) software. Only protein complexes III and IV of the human ETC were examined, since the bovine protein counterparts for these complexes were the only structures available that were over 70% homologous to the human protein complexes.

Before inferences can be drawn from these protein models, on the effect of any non-synonymous mutations, it was necessary to understand the interactions between different protein subunits and how these work together to create a working protein complex.

##### **1.5.4.1.1. Cytochrome bc<sub>1</sub> complex**

The cytochrome bc<sub>1</sub> complex represents Complex III of the mitochondrial respiratory chain (**Figure 1.8**). Its main function is in the electron transfer

processes from the hydrophobic carrier ubiquinol to cytochrome c, during which there is a redox reaction that is initiated from the oxidation of hydroquinones ( $\text{QH}_2$ ) to quinones (Q). This is followed by the reduction of cytochrome c<sup>87</sup>.



**Figure 1.8** Diagram of the Complex III monomer, cytochrome  $bc_1$ . The locations of the redox prosthetic groups in cytochrome b (heme  $b_L$  and  $b_H$ ) and in cytochrome  $c_1$  (heme  $c_1$ ) are shown. The main processes of the Q cycle are also represented, in addition to the net movement of the protons during catalysis.

Interaction between cytochromes b and  $c_1$  forms the cytochrome  $bc_1$  core complex. However, completion of the structure requires the presence of the Rieske iron-sulfur protein, which spans cytochrome b and  $c_1$  thereby connecting the 2 monomers<sup>88</sup>. The iron-sulfur cluster resides within the mobile C-terminus globular domain of the Rieske iron-sulfur protein<sup>89</sup>. Stabilisation of this redox

prosthetic group occurs only after the protein has formed a complete complex with cytochrome  $bc_1$ <sup>89</sup>. Formation of the cytochrome  $bc_1$  complex is made possible as the c-terminal helix of cytochrome  $c_1$  interacts with the fifth cytochrome b helix. The flexibility of the Fe/S protein head domain permits interaction with the cytochrome  $bc_1$  complex<sup>87</sup>, which is essential as the frequency at which the head domain moves controls the rates of catalysis<sup>90</sup>.

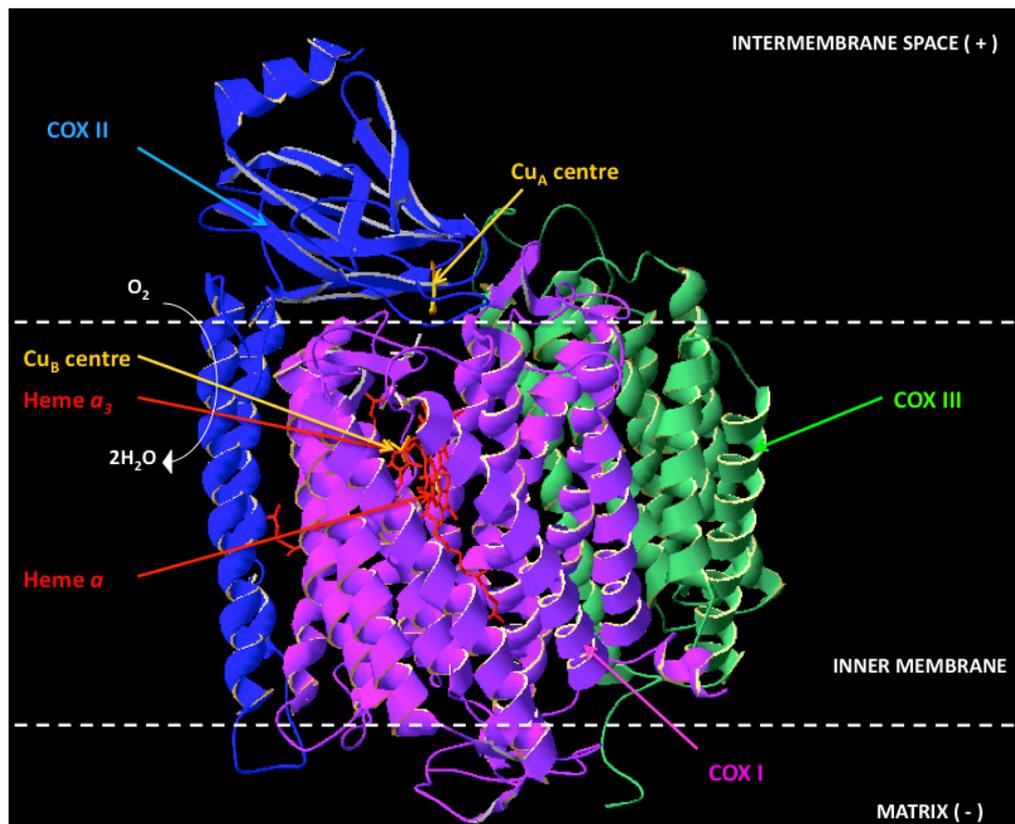
The Q cycle is responsible for carrying out the functions of cytochrome  $bc_1$  complex. Two catalytic sites exist that allow for oxidation of ubiquinol (QH<sub>2</sub>) and reduction of ubiquinone to ubiquinol (Q<sub>i</sub>). The Q<sub>o</sub> site and Q<sub>i</sub> site reside opposite one another on the positive and negative side of the inner membrane, respectively<sup>91</sup>. Although the two catalytic centers Q<sub>i</sub> and Q<sub>o</sub> are located apart from one another, they are actively coupled to each other<sup>88</sup>.

During two consecutive oxidation reactions of QH<sub>2</sub> at the Q<sub>o</sub> site, two electrons are lost. One electron is transferred to the Q<sub>i</sub> site to re-generate ubiquinol (QH<sub>2</sub>), which is then released from the enzyme complex<sup>92</sup>. The other electron is shuttled to the 2Fe2S cluster found in the Fe-S protein. This second electron is transferred from the Fe-S protein to cytochrome  $c_1$  during catalysis by movement of the head domain, from which it is further transferred to the hydrophobic carrier, cytochrome c, and shuttled to cytochrome c oxidase (Complex IV). During the Q cycle, by-products of superoxides can be formed from the oxidation of ubiquinol, as a result of electron leakage<sup>93</sup>. Nonetheless, the primary function of the Q cycle is in the generation of a membrane potential across the mitochondrial inner membrane. The net yield of 2 protons per catalytic cycle

contributes towards a proton gradient<sup>87</sup>, which is important for the production of ATP via ATP synthase (Complex V)<sup>92</sup>.

#### 1.5.4.1.2. Cytochrome c oxidase complex

Cytochrome c oxidase (COX) I-III together form the catalytic core of Complex IV<sup>94</sup> (**Figure 1.9**). Within the hydrophobic core of COX I, there are two heme A cofactors (*a* and *a*<sub>3</sub>)<sup>95</sup>. The flow of electrons begins from the hydrophobic carrier cytochrome c, which transfers electrons to the Cu<sub>A</sub> centre in the COX II subunit, and subsequently to the heme *a* site. From here, electrons are shuttled to the heme *a*<sub>3</sub>-Cu<sub>B</sub> bimetallic centre, where oxygen binds and becomes reduced to water<sup>96</sup>.



**Figure 1.9** Cytochrome c oxidase (Complex IV) of the mitochondrial ETC. Three subunits (COX I, II and III) comprise the catalytic core of Complex IV, with the two heme A cofactors, the copper centres and the reduction of oxygen to water.

COX II is highly hydrophilic compared to COX I and III<sup>97</sup>, and is structurally the smallest of all the subunits that comprise the catalytic core of cytochrome c oxidase<sup>98</sup>. On the other hand, COX III is absent of any prosthetic groups, with its structure being defined by the presence of seven membrane-spanning alpha helices.

Assembly of Complex IV occurs in a successive manner, initiating from COX I followed closely by recruitment of accessory factors such as heme groups, and subsequently by other subunits<sup>99</sup>. As a result, defects arising early during this process are thought to have a cumulative effect on the overall assembly, and thereby functionality of the protein complex<sup>95</sup>. Following the complete assembly of Complex IV, there appears to be no interaction between COX II and III, as they interact only through the transmembrane region of COX I<sup>100</sup>.

Quality control regarding assembly of the proteins occurs throughout protein biosynthesis, at both the subunit level and also when the complex is fully formed<sup>94</sup>. However, since there is thought to be a mutation threshold level, some defects in the mitochondrial respiratory chain are permitted, as non-mutated counterparts are believed to compensate for activity in the mutant proteins<sup>95</sup>.

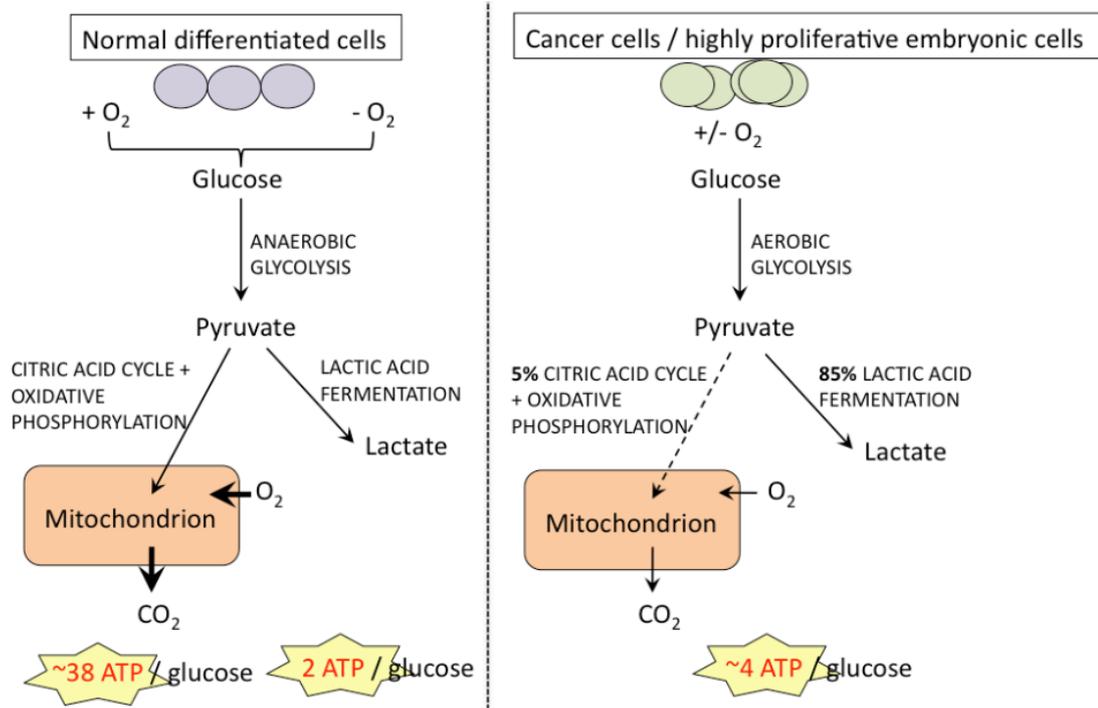
All subunits of Complex IV together with cytochrome B represent the mtDNA-encoded proteins that are most conserved throughout evolution. COX I is the most conserved region (47.2%) amongst the 13 mtDNA-encoded polypeptides, when comparing the human amino acid sequences to the yeast protein counterpart. This was followed by cytochrome B (32.4% protein homology) and

the remaining Complex IV subunits COX III (30.8% protein homology) and COX II (28.6% protein homology)<sup>101</sup>. Reasons for protein conservation for these mtDNA regions can be attributed to their roles as key components of the ETC.

In summary, cytochrome B has the important role of catalysing the transfer of electrons from the hydrophobic carrier ubiquinol to cytochrome c, which is coupled to the translocation of protons across the inner membrane of the mitochondrion. Within the cytochrome bc<sub>1</sub> complex, there are prosthetic groups and redox centres that collectively maintain activity of the Q cycle<sup>102</sup>. Complex IV catalyses the final stages of the ETC by reducing molecular oxygen to water, and as such, mutations to the complex subunits are more likely to be less tolerable. For example, the region within COX II that is involved with the shuttling of electrons from cytochrome c to the Cu<sub>A</sub> center has been reported to be very highly conserved<sup>102</sup>. Consequently, mtDNA variants detected within these regions have a greater chance of being predicted to cause protein impairment.

### **1.6. Warburg effect**

Frequently, in tumor biology, there is a metabolic switch from respiration via OXPHOS to the preferential use of glycolysis, even in the presence of sufficient oxygen concentration<sup>103</sup>. This process results in the Warburg effect (**Figure 1.10**). Although glycolysis yields fewer molecules of ATP per molecule of glucose, the generation of lactate regenerates and permits the recycling of reducing factors, such as NADH, that assist in the continual cycling of glycolysis. This results in the rapid generation of low levels of ATP, which supports the high levels of proliferation characteristic of these tumor cells<sup>104</sup>.



**Figure 1.10.** *The Warburg effect. Normal cells undergo anaerobic glycolysis and subsequently utilize pyruvate, the product of glycolysis, in different ways depending on the concentration of oxygen. In the presence of sufficient oxygen concentration, pyruvate undergoes complete oxidization to CO<sub>2</sub> during the process of OXPHOS. In the presence of limiting oxygen concentration, pyruvate is converted to lactate in order for glycolysis to continue. Metabolism via this pathway yields fewer molecules of ATP compared to conditions of high oxygen concentration, but is the preferred mode of metabolism in both cancer and highly proliferative cells regardless of the oxygen concentration available (Warburg effect).*

During development and in the event of tissue repair, rapidly proliferating embryonic stem cells have also been demonstrated to undergo metabolism via aerobic glycolysis<sup>105</sup>. There are several reasons for cancer cells to respire through aerobic glycolysis. The first is that the production of lactate reduces the local pH

levels, which may deter normal cells from migrating into the region thereby encouraging metastasis of the tumor<sup>106</sup>. Reduced oxygen availability might also be a factor in preventing the establishment of normal cells, since cancer cells are more readily able to adapt to these unfavorable conditions. Additionally, the secretion of lactate suppresses the host immune response by inhibiting the maturation of monocytes to form dendritic cells, which have a role in the initiation of specific T-cell immune response<sup>107</sup>. Lactate also prevents the release of cytokines from these dendritic cells as well as from the cytotoxic T lymphocytes, which are otherwise important for preventing the further growth of tumor cells<sup>108</sup>.

However, aerobic glycolysis is not universally applicable to all cancer cells, as there are exceptions where tumor cells retain the preferential use of OXPHOS<sup>109</sup>, for example, in the THP-1 leukemia cell line<sup>110</sup>. Nonetheless, elevated expression of factors, such as hypoxia-inducible factor (HIF) and c-Myc, are commonly observed in cancer cells, and has been demonstrated to correlate with increases in the expression of lactate dehydrogenase A<sup>111,112</sup>.

The balance between the activities of lactate dehydrogenase and pyruvate dehydrogenase dictates whether glycolysis or OXPHOS is used as the predominate method for cellular respiration<sup>113</sup>. However, it has been suggested that p32, a surface marker on cancer cells, may participate in the switch between glycolysis and OXPHOS<sup>114</sup>. The down-regulation of p32 was reported to promote OXPHOS over glycolysis, whereas its up-regulation was shown to promote the opposite<sup>114</sup>.

### **1.7. Other effects of mitochondrial dysfunction in cancer**

Apart from changes to mitochondrial energy metabolism, cancer cells also have abnormal iron metabolism. Iron is an essential component required for the heme and non-heme cofactors present in the protein complexes of the ETC<sup>115</sup> (also refer to section 1.5.4.1.1), and is also critical for the function of ribonucleotide reductase responsible for the synthesis of deoxyribonucleotides<sup>116,117</sup>. This means that dysfunction or decreased levels of iron would negatively impact on the energy production in the mitochondria. On the other hand, mitochondrial diseases such as, Alzheimer's or Parkinson's disease, can be characterized based on their over-accumulation of iron<sup>118</sup>. However, much work remains to be determined in understanding the exact location and degree of iron accumulation that causes pathogenesis in these diseases<sup>118</sup>.

Another effect of cancer on mitochondrial biology can be illustrated by the over expression of human microRNA, MiR-25, in ovarian and colon cancer cells. MiR-25 is believed to down-regulate levels of the mitochondrial calcium uniporter (MCU) and therefore decrease the uptake of mitochondrial calcium ions<sup>119</sup>. The consequence of this results in cancer cell survival, thereby providing another role for mitochondria in the control of tumor biology.

### **1.8. MtDNA variants in cancer**

The release of ROS by-products, caused by leakage of electrons during the sequential transfer across protein complexes of the ETC, can damage mtDNA by inducing oxidative base lesions, since mtDNA resides close to the ETC<sup>120</sup>. ROS production in cancer cells has the potential to be detrimental, since enhanced

ROS release combined with impaired mitochondrial antioxidant capacity in these cells, disrupt the integrity of mtDNA<sup>121</sup>. On the contrary, low levels of ROS production in normal cells, is believed to be important in mediating cell proliferation and in the control of cell signaling response<sup>122</sup>.

MtDNA sequence variants have been reported in various cancer types as potential contributors towards the onset of cancer (**Table 1.1**). Those thought to have roles in breast cancer include variants at positions 12308bp<sup>123</sup>, 9055bp (G→A), 10398bp (A→G) and 16519bp (T→C)<sup>124</sup>. On the contrary, mtDNA variants predicted to reduce the risk of breast cancer have also been proposed, which include those at positions 3197bp (T→C) and 13708bp (G→A). These have been suggested to stabilize the protein structures where these mutations reside<sup>124</sup>.

Differences in mtDNA haplotypes of individuals also have an influence on the susceptibility of developing cancer, an association which has been shown with several cancer types including, prostate cancer<sup>125</sup> as well as breast cancer<sup>123,124</sup>. MtDNA haplotypes refer to inherited changes in the mtDNA sequence of individuals that allow for the determination of their maternal ancestry, and therefore identify these individuals as originating from a particular ethnic background or region<sup>126</sup>. For Chinese people belonging to sub-haplogroups D, D4a and D5, it is thought that their mitochondrial genome increases their risk to oesophageal cancer<sup>127</sup>. Furthermore, in a separate study, it was reported that within the D-loop, the presence of the 150bp (C→T) variant positively associated with the probability of developing human papillomavirus infection and cervical

**Table 1.1. MtDNA variants reported in colorectal, breast and thyroid cancer**

Cancer	Reported variants	MtDNA region
Colorectal <sup>128</sup>	T1651C	tRNA <sup>Val</sup>
	G2758A, T2352C, A2706G, A2768G, T2885C, T2914G, 3107delG	16s rRNA
	C3316T/A, A3759G	ND1
	A4769G, G5178T	ND2
	G6260A, G7028A, T7055C/G, 7335insC, G7256A, T7146A	COX I
	C7498T	tRNA <sup>Ser</sup>
	C7521T	tRNA <sup>Asp</sup>
	G8206A, G8251A	COX II
	A8618G, T8784C, A8860G, G9055A	ATP 6
	A10398G, C10400T	ND3
	T11641C	ND4
	C12633A	ND5
	G16233A, G16266A, G16294A, T16309G, A16325G, C16390T	D-loop
	9-bp Ins5892C, 8269-9bp del	Non-coding
Breast <sup>124,129</sup>	G85A, T194C/T, C195T/C, C199T, C204T, A207G, C295T, 2-bp Ins303C, T210C/T, Ins311C, C16069T, C16114T/C, G16145A/G, T16189C, A16293G, T16519C	D-loop
	G709A, T710C, C1243T/C, A1555G	12s rRNA
	G1719A, G2706A, T3197C	16s rRNA
	T3394C, G3849A/G	ND1
	T4216C, A4295G, T4323C/T	tRNA <sup>Iso</sup>
	T4336C	tRNA <sup>Gln</sup>
	C4499T/C, A4529T, G4580A, G4665A/G, A4917G, A5240T, G5460A	ND2
	C7028T	COX I
	G8251A, A8498G/A, G8994A, G9055A	ATP 6
	T9885A/T	COX III
	T10034C	tRNA <sup>Gly</sup>
	A10398G	ND3
	A11768G/A	ND4
	A12308G	tRNA <sup>Leu</sup>
	T12636C, A12642G, C12852T/C, G13263A, G13368A, T13397A, T13398A, T13674G/T, G13708A	ND5
	T14470C	ND6
	C14766T	tRNA <sup>Glu</sup>
	A15607G, A15655G/A, C15700T, G15755T/G, T15783C, A15824G	CYT B
	A15924G/A	tRNA <sup>Thr</sup>
Thyroid <sup>130</sup>	Ins3571C, 3331del242bp, T3949C, G3392A, T4222C, G4148A	ND1
	G5185A, G4720A, G4975A, G4831A	ND2
	A8836G, A8836G, G8839A	ATP6
	G10537A	ND4L
	11084delCA, 11038delA, 10885delT, G11403A, G12056A, G11475A, T11613C	ND4
	T12797C, G13414A, A13870T, Ins13235T, T13271C, G13042A	ND5
	T15209C, T15674C, T15843C, T15813G	CYT B

cancer, independently of mtDNA copy number within a population of Chinese women<sup>131</sup>. It was thought that the C150T variant promoted increased mitochondrial replication or increased ATP production, which was beneficial for tumor cell survival<sup>132</sup>.

Statistics for colon cancer estimate that defective mtDNA could be responsible for up to 70% of cases<sup>133</sup>, whereas for gastric cancer, predictions were in the range of 5 to 37%. In samples representative of head and neck cancer, a total of 9 mtDNA base substitutions were observed, together with one base insertion across the entire genome<sup>134</sup>. Identification of these mtDNA rearrangements can be determined using non-invasive techniques on patients, for example through the analysis of saliva samples<sup>135</sup>. Likewise, bladder cancer mtDNA variants can also be identified using urinary patient samples<sup>135</sup>.

Aside from mtDNA variants arising as a result of sequence base substitutions, deletions are also commonly observed in cancers. Large-scale deletions, such as the 4977bp deletion located between mtDNA positions 8470bp and 13477bp disrupts the activity of multiple mtDNA genes that are important for OXPHOS within the ETC. In this context, genes that are disrupted include both *ATPase 6* and *ATPase 8* subunits of Complex V, *Cytochrome C Oxidase III (COX III)* of Complex IV and several of the Complex I subunits<sup>136</sup>. This deletion has been associated with lung<sup>136</sup>, thyroid<sup>137</sup>, breast<sup>138</sup>, oral<sup>139</sup> and gastric cancer<sup>140</sup> and is thought to be a relatively common mutation amongst cases of tumor development.

Overall, there is evidence to suggest that mtDNA variants could lead to the onset of cancer. However, there remain a few exceptions. For example, it has been demonstrated that, with the progression of head and neck squamous cell cancer, mtDNA variants become established at a much later stage of tumor development with levels of mutation reported between 21 – 51%, primarily within the D-loop region<sup>141</sup>. There was also the preferential establishment of coding region variants within ATPase 6, NADH Dehydrogenase 2 and 5 (ND2 and ND5), Cytochrome B (CYT B) and COX III<sup>142</sup>. Together, this suggests that in some cases, mtDNA variants arise as a consequence of tumor development. Furthermore, the development of these mtDNA variants was thought to occur independently of mtDNA copy number in human cancer cells<sup>143</sup>.

The mtDNA of Glioblastoma multiforme (GBM) and multiple myeloma cells were examined to determine whether these cancers share the presence of known mtDNA variants identified within other cancer types. The purpose of this was to investigate whether certain mtDNA variants predispose cells to develop cancer. This could explain why cancer cells preferentially switch their metabolism to glycolysis during tumor establishment. Multiple myeloma and GBM were chosen for this analysis, as there is an increasing interest in the involvement of mitochondria for both cancers.

### **1.9. Multiple myeloma and mitochondria**

Multiple myeloma is a type of bone marrow cancer affecting plasma cells. It represents one of the more common haematological disorders, and is frequently preceded by a premalignant stage known as, monoclonal gammopathy of

unknown significance (MGUS). Although MGUS is asymptomatic, it is relatively common, affecting more than 3% of individuals over 50<sup>144</sup>, with numbers rising to 8% in those over the age of 80<sup>144</sup>. However, the risk of MGUS progression into fully malignant multiple myeloma is very low, estimated at 1% per year after diagnosis<sup>145</sup>. In the UK alone, over 4,000 new cases of multiple myeloma arise per year, which is in addition to the current estimate of approximately 20,000 to 30,000 individuals already suffering from the effects of this cancer<sup>146</sup>. Many of the symptoms displayed by affected patients appear to be non-specific<sup>146</sup> and amongst these, approximately 80% of cases involve degradation of the bone<sup>147</sup>.

Development of myeloma involves a multi-step process. Transformation towards a malignant cell type is normally initiated from a combination of altered genetic and environmental events that accumulate and cause change in the functionality of a plasma cell<sup>148</sup>. The resultant myeloma cell invades cavities in the bone, and secretes monoclonal proteins that are abnormal forms of immunoglobulin. These proteins eventually crowd out the functional immunoglobulins secreted by the normal plasma cells, resulting in the onset of symptoms, such as reduced immunity, leading to infection in the affected patient<sup>149</sup>. Proliferation rates for multiple myeloma cells tend to be dependent on the stage of advancement, with newly transformed cells having a slower rate of proliferation than those at a later of stage of progression<sup>150</sup>.

Currently, there are no effective treatments that exist for the complete elimination of multiple myeloma. Instead, therapy aims to prolong the survival of

affected patients<sup>151</sup>. One of the key features of multiple myeloma is its ability to undergo recurrent relapse, such that with each relapse period, the biology of the cancer changes from its original state of diagnosis. Alterations could include differences in the way the cancer responds to chemotherapy or differences in their behavior *in vivo*, with one study in particular demonstrating changes to the nuclear genomic DNA sequence variants depending on the stage of tumor progression<sup>152</sup>.

A number of antioxidant systems reside within the mitochondria, and one type of defence mechanism relies on the import of cytoplasmic glutathione (GSH) into the organelle<sup>153</sup>. The import of GSH maintains the mitochondrial redox state to prevent the onset of dysfunction and cell death<sup>154</sup>. Since mitochondrial function is often perturbed in cancer, an anti-tumor drug, Imexon, has been developed for use as a potential therapy to selectively eliminate cancer cells whose dysfunction originates in the mitochondria. Imexon functions by sequestering molecules, such as cysteine, that are essential for the production of GSH<sup>155</sup>. Physically, Imexon exerts changes to the ultrastructural arrangement of mitochondria. Moreover, it has been reported to increase cellular susceptibility to ROS damage and imbalance of the membrane potential that spans across the inner mitochondrial membrane<sup>155</sup>. Although cellular cytotoxic response of Imexon has been compared across a range of multiple myeloma cell lines, which include the RPMI 8226, U266, NCI-H929 and NB-4 cells, it has been demonstrated that differences exist between each in terms of their sensitivity to the drug. Despite this, all cell lines investigated were able to show reduced ability to survive in the presence of the drug<sup>155</sup>.

Other mechanisms that have evolved within the mitochondrion to counteract ROS accumulation include enlargement in the overall size of the mitochondrion, which is thought to increase mitochondrial internal surface area to dilute the effects of ROS<sup>155</sup>. Aside from alterations to mitochondria as a whole, Imexon was discovered to specifically target mtDNA, whilst leaving nuclear DNA intact at low dosages<sup>155</sup>. However, to date, not many studies exist that focus on mtDNA and any changes that may be induced in the genome during tumor development.

### **1.10. GBM and mitochondria**

Glioblastoma multiforme (GBM) is a cancer of the brain that originates from abnormal proliferation of glial cells. GBM can arise in one of two ways; *de novo* initiation of the disease is referred to as its primary state, whereas GBM that arises as a result of de-differentiation of a less aggressive tumor is known as a secondary tumor. GBM is the most common malignant tumor, estimated to affect 80% of adults diagnosed with primary brain cancer<sup>156</sup>. Prognosis for affected patients remains poor, as GBM is a category IV brain tumor, classified as the most severe of its kind according to the World Health Organisation (WHO) standards<sup>157</sup>. Median survival rates for these patients are around 12 months<sup>158</sup>, and in 90% of cases the tumor re-establishes itself even after treatment with intense combination therapy<sup>159</sup>. As such, current treatment of GBM focuses largely on palliative care<sup>160</sup>.

Necrosis of local brain tissue is often observed in GBM<sup>161</sup> and, at the genomic level, there has been reported genomic instability of these tumor cells, making it

difficult to define the underlying cause of the tumor<sup>157</sup>. Investigations conducted into the microenvironment of glioma cells provided evidence for functional differences between cells located at the periphery of the sphere and those within its centre<sup>162</sup>, supporting the concept of heterogeneity for these cell types. GBM is also known to acquire mutations during the course of its tumor development<sup>163</sup>.

One study attempting to define the cause of tumor establishment in GBM was able to show that expression of protein kinase M2 (PKM2) was positively correlated with the grade of GBM. PKM2 participates in the glycolytic pathway converting phosphoenolpyruvate (PEP) to pyruvate by catalyzing the transfer of a phosphate group from PEP to a molecule of ADP<sup>164</sup>, and functions to restrict the production of pyruvate in order to divert resources towards stimulating proliferation via the generation of macromolecules<sup>165</sup>. It was also determined that GBM cells are highly reliant on OXPHOS, as disruption to this process negatively impacted upon clonal expansion of the GBM cells<sup>166</sup>. This was identified via examination of insulin-like growth factor 2 mRNA binding protein 2 (IMP2), which normally delivers mitochondrial mRNAs to their protein complexes for assembly<sup>166</sup>.

Alantolactone is a sesquiterpene lactone found in the roots of *inula helenium* L.<sup>167</sup>. Its anti-cancer properties for GBM cells has demonstrated potential towards the development of future therapies to treat this malignant disease<sup>168</sup>. It was discovered that, in the presence of this compound, cell proliferation was prevented and apoptosis was stimulated in GBM cells, an effect that reported

time and dose dependence. The mechanism of apoptosis was determined to occur via depletion of glutathione (GSH) and the increase in ROS, as well as loss of the mitochondrial membrane potential, which is similar to the effects that can be induced by Imexon for the selective elimination of these cancer cells<sup>168</sup>.

### **1.11. Conclusion**

Transformation of normal cells into the malignant phenotype involves the dysregulation of many collaborating factors at the nuclear and mitochondrial level. MtDNA is unique to the mitochondrion, and has an important role in controlling the metabolism of cells. In cancer, the processes that govern mitochondrial energetics are believed to be compromised according to the Warburg effect, which states that there is the preferential utilization of glycolysis regardless of the availability of oxygen. The phenotypic changes that arise as a consequence of these metabolic changes are thought to provide cancer cells with adaptability to establish a tumor in regions that are considered sub-optimal for the growth of normal cells.

There is some support for the concept that mitochondria could have a role in the initiation of tumor development for some cancer types. MtDNA variants have been identified in a number of cancers and amongst these, several have been thought to promote or prohibit cancer development, depending on the specific variant involved. Given that targeting of the mitochondria has been suggested as a potential form of treatment against GBM and multiple myeloma, it would be feasible to examine the contribution of the mitochondrial genome to the development of these cancers.

### **1.12. Hypothesis**

MtDNA has a role in the initiation of tumor development in GBM and multiple myeloma, with mtDNA variants disrupting the functional ability of the ETC protein complexes. In this respect, it was hypothesized that variants would prevent the proper assembly of protein complexes that comprise the ETC, or would carry impairment in their ability to shuttle electrons between the protein complexes, or disrupt the translocation of protons across the inner mitochondrial membrane. Failure to establish the electrochemical gradient required for the production of ATP via Complex V, would compromise ATP production and force cancer cells to use glycolysis as their primary source of metabolism, thereby following the Warburg effect.

### **1.13. Aims and objectives**

#### **2. To identify novel mtDNA variants that could be responsible for initiating multiple myeloma.**

To do this, next generation sequencing will be used, and the variants identified will be analysed by bioinformatics and other *in silico* tools to predict the severity of the variants identified and their likelihood of causing protein impairment (Chapter 3).

#### **3. To determine the importance of mtDNA in maintaining the tumorigenic phenotypes of multiple myeloma cells.**

Multiple myeloma cells will be depleted of their mtDNA in culture and analyses will be conducted to determine the changes in the expression profiles for genes involved with mitochondrial biogenesis, transcription and replication, as well as

genes associated with its tumor biology. Metabolic assays will be performed to measure the oxygen consumption of these cancer cells. ATP and lactate assay measurements will also be taken to define any changes to the tumor phenotype of these cells after induced depletion of mtDNA (Chapter 4).

**4. The outcomes from Aim 1 and 2 will be tested in another model, using GBM cells. Therefore, the aim is to identify mtDNA variants that could be responsible for initiating GBM and to determine whether certain tumor types harbour specific mtDNA variants.**

GBM cells and patient tumor samples will undergo next generation sequencing to determine whether the mtDNA variants are the same or different to those identified from sequencing of the multiple myeloma cells. GBM tumor variants will be compared with mtDNA variants identified from the sequencing of normal brain biopsy samples, and GBM cell line variants will be compared with the non-tumorigenic human neural stem cells. To determine the importance of mtDNA in the establishment of a tumor, GBM cells will be depleted of their mtDNA in culture and transplanted into immunodeficient mice for the recovery of mtDNA copy number to examine whether the presence of mtDNA is required for the establishment of a tumor *in vivo* (Chapter 5).

**4. To determine the contribution of mtDNA to the tumorigenic phenotype of GBM cells.**

MtDNA depletion will be performed and cells subsequently analysed for changes in the expression levels of key genes involved with mitochondrial biogenesis,

transcription and replication. GBM cells differentiated into astrocytes and neurons will also be analysed for changes to the expression levels of these genes to determine whether the tumorigenicity of GBM cells is lost upon differentiation into mature cell types (Chapter 6).

## **Chapter 2: General methods**

For chemical compositions of laboratory reagents, refer to Appendix (i).

## **2.1. Cell culture**

All culturing was performed under sterile conditions in the Top-Safe 1.2 (Class II) Bio-cabinet (LAF Technologies Pty Ltd., Victoria, Australia). Generally, the procedure for passaging of the cells was as follows:

### **2.1.1. Passaging of adherent cultures**

Media was aspirated and rinsed with Dulbecco's Phosphate-Buffered Saline (PBS) in the absence of  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$ . TrypLE Express (Gibco, Carlsbad, CA, USA) was used as a cell dissociation agent and cells were incubated for 3 minutes at  $37^{\circ}\text{C}$ . To inactivate TrypLE activity, cell culture media containing all the essential supplements to fully support cell growth (complete media) was added to the culture at a ratio of 1:3 (TrypLE:media), and transferred to a 15ml conical tube (BD Biosciences, Franklin Lakes, NJ USA)\*. Cells were centrifuged for 3 minutes at 14,000rpm (Eppendorf centrifuge 5702; Eppendorf AG, Hamburg, Germany). The supernatant was aspirated and cells were resuspended in 1-2ml complete media in preparation for the cell count.

\* Until this point, the procedure outline has been specific to passaging for adherent cultures. The steps after this are shared with the protocol for non-adherent cells.

### 2.1.2. Passaging of non-adherent (suspension) cells

Cells in suspension were transferred to a 15ml conical tube (BD Biosciences). The culture flask was rinsed with PBS and subsequently transferred into the 15ml conical flask containing the cell suspension. Cells were centrifuged for 3 minutes at 14,000rpm, after which the supernatant was aspirated and the cells resuspended in 1ml TrypLE Express for incubation at 37°C for 3 minutes. All remaining steps of the procedure follow the same protocol as above (section 2.1.1).

All cells were counted using a haemocytometer counting chamber of 0.1mm depth by 0.0025mm<sup>2</sup> (Improved Neubauer with bright-line; Blau Brand, Wertheim, Germany) as cultures were passaged, using a glass coverslip of 0.40mm. Cell viability was assessed using the trypan blue exclusion assay (Sigma, St Louis, MO) at a 1:1 cell to trypan blue ratio. The calculation for total cell numbers were as follows, based on viable cell counts:

i) Average cell count per grid on hemocytometer X dilution factor of cells

= value 1

ii) Value 1 X 10<sup>4</sup> = value 2

iii) Value 2 X cell resuspension volume (ml) = total number of cells in culture

The percentage viability of cells was calculated using the following formula:

= [1.00 – (number of blue, nonviable cells / total number of cells) X 100]

All cells were visualized using the Primo Vert inverted microscope (Carl Zeiss Microimaging, Gottingen, Germany) and all cell lines were maintained under conditions of 37°C with 5% CO<sub>2</sub>. Details for the maintenance of each cell line used are outlined below.

### **2.1.3. Human foreskin cell line, BJ fibroblasts**

BJ fibroblasts (American Type Culture Collection, number CRL-2522, Manassas, VA, USA) were maintained as a monolayer culture in 1x Dulbecco's Modified Eagle Medium (DMEM) high glucose, containing 10% Fetal Bovine Serum (FBS), 2mM GlutaMAX, 0.1mM non-essential amino acids (NEAA), 1mM sodium pyruvate and 25U of penicillin/streptomycin (all from Gibco). Fibroblasts were passaged every 3 to 4 days, with cultures seeded at  $8 \times 10^4$  cells/cm<sup>2</sup> at the beginning of each passage cycle.

### **2.1.4. Human hepatocarcinoma cell line, HepG2**

HepG2 cells were cultured as a monolayer in DMEM media supplemented with 10% FBS, 2mM GlutaMAX and 25U of penicillin/streptomycin (All from Gibco). Cells were passaged every 3 days and seeded at a density of  $8 \times 10^4$  cells/cm<sup>2</sup>.

### **2.1.5. Human multiple myeloma, U266 cells**

Human multiple myeloma U266 cells<sup>169</sup> were cultured as a single cell suspension in RPMI 1640 media, supplemented with 10% fetal bovine serum (FBS), 2mM GlutaMAX and 1mM sodium pyruvate (all from Gibco). Cells were subcultured at a density of  $8 \times 10^4$  cells/cm<sup>2</sup> and passaged every 3 to 4 days.

#### **2.1.6. Human glioblastoma multiforme, HSR-GBM1 / GBM L1 / GBM L2 cells**

Human glioblastoma multiforme (GBM) cell, HSR-GBM1, GBM L1 and GBM L2 were cultured as neurospheres in DMEM/F12 media supplemented with 2% StemPro neural supplement (Gibco, Carlsbad, CA, USA), 20ng/ml basic fibroblast growth factor (bFGF) and 20ng/ml epidermal growth factor (EGF; both Millipore, Billerica, MA). Cells were passaged every 3-4 days using TrypLE Express (Gibco).

#### **2.1.7. Human glioblastoma multiforme, NO7-152 cells**

NO7-152 cells represent another type of GBM cells. These adherent cells are cultured as a monolayer and required media composed of DMEM high glucose, 4% non-essential amino acids (NEAA), 2mM GlutaMAX and 20% FBS (all from Gibco). Cells were passaged in the presence of TrypLE Express.

#### **2.1.8. Human glioblastoma multiforme, GBM 4 / CSC-020 / CSC-014 cells**

Additional GBM cells, GBM 4, CSC-020 and CSC-014 were cultured in knockout DMEM/F12 media supplemented with 2% StemPro neural supplement, EGF and bFGF (both at concentrations 20ng/ml) with added 2mM GlutaMAX (all from Gibco). The GBM4 cells were maintained as an adherent culture, whereas CSC-020 and CSC-014 were maintained as neurospheres similar to the HSR-GBM1, GBM L1 and L2 cell lines. Passage of the GBM 4, CSC-020 and CSC-014 cells was performed every 3-4 days in the presence of TrypLE Express (Gibco).

### **2.1.9. Depletion and replenishment of mtDNA in U266 and HSR-GBM1 cells**

U266 cells underwent mtDNA depletion using 10 $\mu$ M of 2',3'-dideoxycytidine (ddC), in the presence of 50 $\mu$ M of uridine (both from Sigma, St Louis, MO), which were supplemented to the culture media. Supplies of both factors were replenished daily through half media changes for up to 30 days, with the cells undergoing a complete media change after 3 or 4 days. Collected samples for analysis were rinsed with PBS (Gibco) prior to storage of the cell pellet at -80°C.

For replenishment of mtDNA, U266 cells were cultured in complete RPMI 160 media in the absence of ddC and uridine. HSR-GBM1 cells were cultured in conditioned media during mtDNA replenishment after a period of mtDNA depletion. Conditioned media was collected from HSR-GBM1 cells that had been cultured for 2 days in fresh culture media supplemented with growth factors, and stored at -20°C until use.

### **2.1.10. Inhibition of glycolysis using 2-deoxyglucose**

U266 cells were depleted of their mtDNA in culture and allowed to recover in complete media together with the presence of 5mM 2-deoxyglucose.

## **2.2. Isolation of plasma cells**

Bone marrow mononuclear cells were isolated from the bone marrow of patients diagnosed with relapse and/or refractory stage multiple myeloma. Isolation was performed using Ficoll-Paque Plus (Amersham Biosciences, Rydalmere, Australia) followed by treatment with red blood cell lysis buffer (154.98mM

ammonium chloride, 126.61mM EDTA, 9.99mM potassium bicarbonate). Cells were then cultured for 24 hours in complete U266 media after which, plasma cells were magnetically separated using the CD138 microbeads kit (Miltenyi Biotec, North Ryde, Australia).

### **2.3. Fluorescent-activated cell sorting (FACS) on U266 cells**

U266 cells were subjected to FACS based on their expression of the cell surface marker CD45. Cells were harvested and resuspended in 100 $\mu$ L FACS buffer consisting of PBS supplemented with 5% FBS. Cells were then labeled using the mouse anti-human CD45 antibody conjugated to FITC (BD Biosciences) by adding antibodies at a 1:100 dilution and incubating the mixture at 4°C in the dark for 15 minutes. Cells were then washed three times with FACS buffer and centrifuged for 5 minutes each at 300g. The supernatant was aspirated and the cell pellet was re-suspended in 100 $\mu$ L FACS buffer, for analysis on the FACSAria Cell Sorter (BD Biosciences). CD45 positive and negative fractions were collected into two sterile 15ml polypropylene conical tubes (BD Biosciences), and cultured in U266 media as necessary.

### **2.4. DNA extraction**

Extraction of DNA was performed on either cell samples stored at -80°C at a density of  $2 \times 10^6$ , or on 25mg of tissue samples. The procedure was undertaken using the DNeasy Blood & Tissue kit (Qiagen).

#### **2.4.1. DNA extraction from cells**

Cell pellets were thawed and resuspended in a volume of 200 $\mu$ l of PBS (Gibco), after which 20 $\mu$ L of proteinase K (stock 20mg/ml) was added. 4 $\mu$ L of RNase A (stock concentration 100mg/ml) was added and the sample left to incubate for 2 minutes at room temperature♦, before adding 200 $\mu$ L of lysis Buffer AL. This was mixed by pulse-vortexing for 15 seconds (Gallenkamp Spinmix vortex shaker, Gallenkamp, Sussex, UK). The mixture was incubated for 10 minutes at 56°C using the Dri-Block DB-2A heating block (Techne, Staffordshire, UK), after which 200 $\mu$ L of 100% ethanol was added and mixed by pulse-vortexing for 15 seconds.

♦ Until this point, the procedure outlined was specific to the DNA extraction procedure for cultured cells.

The sample was then applied to the silica gel membrane of a DNeasy mini spin column and centrifuged for 1 minute at 8000rpm (Heraeus Biofuge Pico, Kendro, Germany). The flow-through containing debris was discarded and the spin column retaining DNA was placed in a new 2ml collection tube. 500 $\mu$ L of Buffer AW1 wash buffer was added and centrifuged for 1 minute at 8000rpm. The collection tube containing flow-through was again discarded and the column placed in a new collection tube. 500 $\mu$ l of Buffer AW2, a wash buffer containing ethanol, was then applied as a second wash step. This was centrifuged for 3 minutes at 13,000rpm to dry the membrane. The collection tube was replaced with a fresh collection tube and the column was centrifuged for 1 minute at 13,000rpm to remove any residual ethanol traces, before eluting DNA under

centrifugation at 13,000rpm for 1 minute in 30 $\mu$ L of sterile water. Products were stored at -20°C.

#### **2.4.2. DNA extraction from tissues**

25mg of frozen tissue was cut using a sterile scalpel under sterile conditions, and stored in 100 $\mu$ L of Buffer ATL. The sample was pulse-homogenized using the Pro 200 tissue homogenizer (Pro Scientific Inc., Oxford CT, USA) for several seconds each time, until the tissue aggregates could no longer be observed. 80 $\mu$ L of Buffer ATL was added to the tissue lysate, followed by 20 $\mu$ L of proteinase K (stock 20mg/ml). This was mixed by pulse-vortexing (Gallenkamp Spinmix vortex shaker) for 15 seconds before overnight incubation of the sample at 56°C on a heat block (Techne). After incubation, 4 $\mu$ L of RNase A (stock 100mg/ml) was added and mixed by vortexing. The sample was left to incubate at room temperature for 2 minutes, then vortexed before continuing with the DNA extraction procedure outlined above, as for cultured cells.

#### **2.5. RNA extraction**

RNA was extracted from either cell samples stored under conditions of -80°C at a density of  $2 \times 10^6$  or from tissue samples. The protocol for the RNeasy mini kit (Qiagen) was followed.

##### **2.5.1. RNA extraction from cultured cells**

For cultured cells, 350 $\mu$ l of Buffer RLT containing denaturing guanidine isothiocyanate (GITC) was added to a thawed cell pellet for lysis and instant inactivation of RNases that may degrade and disrupt the integrity of the RNA

sample. This was pulse-vortexed for 15 seconds (Gallenkamp Spinmix vortex shaker)Δ.

Δ the procedure outlined until this step was specific to RNA extraction for cultured cells. Methodology after this was shared with the RNA extraction procedure for tissue samples.

Cell homogenization was performed using the QIAshredder (Qiagen) by applying the lysate to the column. This was centrifuged at 13,000rpm for 2 minutes and 1 volume of 70% ethanol was then added and mixed by pipetting. The sample was transferred to an RNeasy spin column and centrifuged briefly for 15 seconds at 13,000rpm. The flow-through was discarded and 350μl of Buffer RW1 was applied to the spin column to selectively isolate RNA of >200bp. This was centrifuged for 15 seconds at 13,000rpm to wash the column membrane, and the resulting flow-through was discarded.

To perform the on-column DNase digestion step, 10μl of DNase I solution (2.727 Kunitz units/μl) was carefully mixed with 70μl of Buffer RDD. This was added to the column membrane and left for 15 minutes at room temperature for on-column DNA digestion, before further addition of 350μl Buffer RW1 to the spin column. The column was centrifuged at 13,000rpm for 15 seconds, the flow-through was discarded and 500μL Buffer RPE, a mild washing buffer, was added to the spin column. This was centrifuged for 15 seconds at 13,000rpm. The flow-through was discarded afterwards and another 500μL of Buffer RPE was applied to the membrane, and centrifuged for 2 minutes at 13,000rpm. To ensure no

carry-over of ethanol, the column was centrifuged for an additional minute at 13,000rpm, before eluting RNA under centrifugation at 13,000rpm for 1 minute using 30 $\mu$ L of RNase-free water. RNA was stored at -80°C.

### **2.5.2. RNA extraction from tissue samples**

30mg of frozen tissue was cut under sterile conditions using a clean scalpel. Tissue was homogenised using the Pro 200 tissue homogenizer (Pro Scientific Inc.) in 600 $\mu$ L Buffer RLT containing beta-mercaptoethanol (Sigma) diluted in a ratio of 1:100 (6 $\mu$ L). Homogenisation was conducted for no longer than 10 seconds each time, until the tissue was completely lysed. The lysate was transferred into a QIAshredder, and centrifuged for 2 minutes at 13,000rpm, continuing with the steps as outlined above for RNA extraction in cultured cells.

### **2.6. Spectrophotometry**

Quantification of DNA and RNA templates was performed using the NanoDrop 1000 spectrophotometer (Thermo Scientific, Wilmington, USA). Prior to its use, the instrument was calibrated with sterile water to provide a blank reading. The NanoDrop 1000 (version 3.7) was used as the analytical software for this instrument.

For the 260/280nm absorbance ratio, a value of 1.8 was considered good quality for DNA and a value of around 2.0 was considered good quality for RNA. Values from 1.8 to 2.2 for the 260/230nm absorbance ratio were considered acceptable for further use.

## **2.7. Reverse Transcription Polymerase Chain Reaction (RT-PCR)**

For synthesis of cDNA, a two-step reverse transcriptase (RT) PCR was performed. 1µg RNA template was added to 540ng Oligo (dT)<sub>18</sub> primer (Bioline) made to a final volume of 20µl with sterile water. The reaction was incubated at 70°C for 5 minutes, using the PTC-200 Peltier Thermal Cycler (Global Medical Instrumentation, Minnesota, USA), and allowed to cool to 4°C before addition of further components to the RNA-primer mix, which involves addition of 5µl of 5x reaction buffer, 10mM dNTPs, 40U of RiboSafe RNase inhibitor and 200U of enzyme BioScript™ (all from Bioline) to a final reaction volume of 30µl in sterile water. The sample underwent thermal cycling at 42°C for 60 minutes and 70°C for 10 minutes with products held at 4°C at the end of the reaction. cDNA was stored at -20°C.

## **2.8. Polymerase Chain Reaction (PCR)**

50ng of template was used for each PCR reaction, in the presence of the following: 1X NH<sub>4</sub> Reaction Buffer, 1.5mM Magnesium Chloride (MgCl<sub>2</sub>), 1mM dNTPs, 2.5U of BioTaq DNA polymerase (all from Bioline) and 25µM each of the forward and reverse primer. The total reaction volume was made up to 50µL by the addition of sterile water.

Cycling conditions for conventional PCR using the PTC-200 Peltier Thermal Cycler (Global Medical Instrumentation) were as follows: 95°C for 5 minutes, then 35 cycles of 94°C for 30 seconds, specific primer annealing temperature for 30 seconds (Refer to individual chapter methods for primer details) and 72°C for

30 seconds. Final extension of the products was performed at 72°C for 10 minutes with end products held at 4°C.

The formula used for calculation of the optimal annealing temperature of primers was based on the GC content of the short primer sequence, taking into account nucleotide length of the primer:

$$T_m = \%GC \times 0.41 + 64.9 - (600 / n) \quad 170$$

### **2.8.1. Agarose gel electrophoresis**

DNA products were mixed with 6X Loading Dye in a ratio of 5:1, and loaded onto wells of a 1.5%Φ agarose (Bioline) gel made with 1X Tris-acetate-ethylenediamine tetraacetic acid (TAE) buffer, and stained with SYBR Safe (Invitrogen) for product visualisation under UV radiation. Agarose gels were moulded using a gel tray and comb set (Plaztek Scientific, Melbourne, Australia). Gel electrophoresis was performed under conditions of 100V for 50 minutes using the PowerPac 300 voltage power supply (Bio-Rad, New South Wales, Australia) in the presence of 1X TAE running buffer.

Φ A 1.5% agarose gel was prepared for products anticipated to be between 400-1000bp in size. Fragments below 400bp, were resolved on a 2% agarose gel and, for products of long range PCR, producing >8kb DNA fragments, a 0.7% agarose gel was prepared.

### **2.8.2. Gel documentation and analysis**

DNA products following gel electrophoresis were observed under UV illumination, using the MiniBIS Pro (DNR Bio-Imaging Systems, Jerusalem, Israel) with images acquired through GelCapture (version 7.0.15), DNR's acquisition and analyses software. UV exposure of the products was set at 0.7 with settings for the gain and brightness adjusted as needed.

## **2.9. PCR product purification**

### **2.9.1. Direct purification of PCR products**

PCR products between 100bp to 10kb in size were purified using the QIAquick PCR purification kit (Qiagen), only if the product appeared to be specifically generated from amplification for the gene of interest as assessed by agarose gel electrophoresis. The protocol was as follows: Buffer PBI containing pH indicator was added at a ratio of 5:1 to the volume of PCR product that was required for purification. In cases where the pH was found to be sub-optimal, 10 $\mu$ L of 3M sodium acetate (pH 5.2) salt was added to adjust pH conditions. The mixture was applied to a QIAquick spin column and centrifuged for 1 minute at 13,000rpm. The flow-through was discarded and 750 $\mu$ L of wash buffer (Buffer PE) containing ethanol was added and left to stand at room temperature for 5 minutes before centrifugation for 1 minute at 13,000rpm. The flow-through was discarded and the column was again centrifuged for 1 minute at 13,000rpm, before the final elution phase, which was performed in 30 $\mu$ L of sterile water.

### **2.9.2. Purification of PCR products from agarose gel following electrophoresis**

PCR products between 70bp up to 10kb were purified following gel electrophoresis using the QIAquick Gel Extraction Kit (Qiagen). In brief, 600µl of solubilisation buffer (Buffer QG) was added directly to the excised gel slice stored in a 1.5ml Eppendorf tube. This was incubated at 56°C for 10 minutes until the gel slice had been completely dissolved in buffer. If the product for purification was less than 400bp or above 4kb in size, 200µL of isopropanol was added to the buffer-gel mix. The contents of the tube was transferred to a QIAquick spin column and centrifuged for 1 minute at 13,000rpm. The flow-through was discarded and 750µl of wash Buffer PE was applied to the column, which was allowed to incubate at room temperature for 5 minutes before centrifugation for 1 minute at 13,000rpm. The empty column was again centrifuged at 13,000rpm for 1 minute, before eluting the DNA in 30µL sterile water.

### **2.10. Long range PCR**

Long range PCR was used to amplify the whole mitochondrial genome in two overlapping fragments:

Long 1: forward primer: 5'-GACGGGCTCACATCACCCCATAA-3', reverse primer: 5'-GCGTACGGCCAGGGCTATTGGT-3'.

Long 2: forward primer: 5'-GCCACAACCTCCTCGGACTCCT-3', reverse primer: 5'-GGTGGCTGGCACGAAATTGACC-3'.

Reactions were performed using the PTC-200 Peltier Thermal Cycler (Global Medical Instrumentation). 50ng genomic DNA template was used in the presence of 1X High Fidelity PCR Buffer, 0.2mM each of dNTPs, 2mM MgSO<sub>4</sub>, 0.2μM each of the forward and reverse primer, and 1.0 unit of Platinum Taq High Fidelity DNA Polymerase in a total volume of 50μL sterilized water.

Conditions for the amplification reaction were as follows: 94°C for 2 minutes, then 35 cycles of: 94°C for 15 seconds, 63°C for 30 seconds, 68°C for 8 minutes 45 seconds∇, followed by a final extension phase of 68°C for 10 minutes, with products kept at 4°C until further use.

∇ Platinum Taq High Fidelity DNA Polymerase has processivity of extending 1kb of sequence for every 1 minute of the extension time.

Purification of long PCR fragments followed the QIAquick PCR purification kit protocol, as described in section 2.9.1.

### **2.11. Real time PCR (qPCR)**

For the generation of real-time PCR standards, products of 150-450bp underwent (1.5%) agarose (Bioline) gel electrophoresis with SYBR Safe (Invitrogen) and were extracted from the gel following the manufacturer's instructions for the QIAquick Gel Extraction Kit (Qiagen). Spectrophotometric analysis (section 2.6) was performed to assess the approximate concentration and quality of the sample prior to use for real-time PCR. High quality samples were diluted to a final starting concentration of 1ng/μl, from which ten-fold serial dilutions ( $1 \times 10^{-2}$  –  $1 \times 10^{-8}$  ng/μL) were generated to produce a correlation coefficient (efficiency;  $R^2$ )

value that is informative of how well the template has been converted to the amplified product during the exponential phase of each cycle, which was generally in the range of 0.86 (mtDNA copy number) to 0.98 (*POLGB*).

Standards were run in duplicate and samples in triplicate, together with water-only negative controls (in triplicate per run). Samples for each gene of interest were assessed twice on independent run cycles to minimize inter-experimental variability. For each reaction, 20ng template/standard, 5 $\mu$ M of both forward and reverse primers all in 1x SensiMix<sup>TM</sup> SYBR (Bioline) was used in a total reaction volume of 10 $\mu$ L. Cycling of the reactions utilized the 72-well Rotor Gene (RG)-3000 instrument (Corbett Research, Cambridge, UK). Data were collected and analysed using software Rotor Gene 6.

Collection of data was performed as previously described<sup>171</sup>. Briefly, acquisition of data was obtained from the FAM/SYBR channel during the extension step. After cycling of the samples and standards, melt analysis was performed with ramping ranging from 47°C to 99°C at 1°C increments to measure product dissociation. To eliminate interfering fluorescence from primer dimers, a second acquisition step was added, set to the temperature at which the specific product starts to denature.

The standard conditions for all genes consisted of 45 cycles of the following: 95°C for 15 seconds, specific annealing temperature (Table 4.1) for 15 seconds, 72°C for 15 seconds and the appropriate temperature for second acquisition for 15 seconds.

### 2.11.1. Analysis of mtDNA copy number

MtDNA copy number per cell was determined by calculating the ratio of mtDNA to nuclear DNA, divided by 2, as there are two copies of the nuclear genome. Normalisation for copy number analysis was performed against endogenous  $\beta$ -globin levels.

### 2.11.2. Gene expression analysis

Calculations for relative gene expression were performed using the Pfaffl method<sup>172</sup> on cDNA. The formula for the Pfaffl equation:

$$\text{Expression ratio} = \frac{(E_{\text{target}})^{\Delta C_{\text{t target (control-treated)}}}}{(E_{\text{ref}})^{\Delta C_{\text{t ref (control-treated)}}}}$$

Where E represents the efficiency of the run, and  $\Delta C_{\text{t}}$  represents the change in Ct value. The Ct value is the value where the threshold level intersects with the amplification curves generated from the real-time PCR run. The threshold level was manually set at 0.1 on the Rotor Gene 6 software, which was set within the exponential phase of the amplification, above the background level of fluorescence detection.  $\beta$ -actin was used as a reference gene against which expression of the gene of interest was measured. For quality control, only reactions of high efficiency values (above 0.88) were considered acceptable.

### 2.12. ATP production

Fresh cells were harvested, resuspended in complete culture media and quantified using a hemocytometer. ATP measurements were performed using the ATPLite assay kit (Perkin Elmer, Zaventem, Belgium). Standard dilutions of

ATP from  $10^{-3}$  mM to  $10^{-6.5}$  mM at 0.5 mM intervals were prepared from a stock of known concentration (10 mM) provided with the assay kit. Standards were run as duplicates on a white 96-well plate (Nunc, Roskilde, Denmark).

All test samples were seeded at 50,000 cells per well, and were run in triplicate alongside negative controls. 50  $\mu$ L of cell lysis buffer was added to each of the wells. The plate was mixed by mechanical shaking at 700 rpm for 5 minutes (room temperature). Substrate buffer containing luciferase and D-luciferin was added to all wells and the plate was again mixed for 5 minutes at 700 rpm. The samples were dark adapted by covering the plate with foil for 15 minutes, following which luminescence was quantified on the FLUOstar Optima plate reader (BMG Labtech, Ortenberg, Germany; v. 2.20). Data analysis for the concentration of ATP within the samples was performed using the MARS data analysis software v.1.10, based on data extrapolated from the standard curve generated within the software program.

### **2.13. Lactate production**

A sample of culture media was obtained from cell cultures 24 hours after a complete media change with and without ddC. A colorimetric system was used to measure quantitatively lactate content, as provided by the Lactate Assay Kit II (BioVision, San Francisco, CA, USA).

The protocol was as follows: a sample of 2  $\mu$ L collected media was used for each well of a clear 96-well plate (Nunc) made to total volume of 50  $\mu$ L with Lactate Assay Buffer. Samples and a negative control were all prepared in triplicate sets

on the plate, whereas standards ranging from 0 to 10nmol generated at dilutions of 2nmol intervals were plated in duplicate. For each well, the following reagents were added: 46µL of Lactate Assay Buffer, 2µL of Lactate Substrate Mix and 2µL of Lactate Enzyme Mix, so that all wells contained a total volume of 100µL. The negative control also contained the reaction mix but in the absence of the enzyme. The plate was subjected to shaking at 700rpm for 5 minutes followed by incubation in the dark for 30 minutes at room temperature.

Absorbance was measured on the FLUOstar Optima plate reader (BMG Labtech) under 450nm wavelength. Data collection and analysis based on extrapolation from a standard curve were performed as for the ATP assay above. The calculation of lactate concentration is represented by the formula:

$$\text{Lactate concentration} = \frac{\text{Amount of lactic acid derived from the linear regression curve (nmol/}\mu\text{L)}}{\text{Total volume of the sample originally within the culture (}\mu\text{L)}}$$

#### **2.14. Respirometry assessment on cells to determine their OXPHOS capacity**

Real-time measurements of oxygen capacity within cells were performed using the Oxygraph-2K high-resolution respirometry system (Oroboros Oxygraph-2K, Innsbruck, Austria) in conjunction with the integrated DatLab (v3.1) analysis software. Prior to the experiment, the two polarographic oxygen sensors from chambers A and B (2cm<sup>3</sup> each) were calibrated in 1X Hanks balanced salt solution (HBSS) (Gibco) to normal atmospheric oxygen concentrations (oxygen saturation levels) and to conditions of zero oxygen levels using a 5% solution of sodium hydrosulfite (Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub>; Sigma) in 1x HBSS.

After calibration, both instrument chambers were rinsed thoroughly using several washes of 100% ethanol, 70% ethanol, water and 1X HBSS. The chambers were then filled with another aliquot of 1X HBSS, which was allowed to warm to 37°C, whilst continuously stirred at 750rpm. Once both chambers reached atmospheric oxygen concentration, the chamber stoppers were lowered to create a closed chamber containing exactly 2ml of 1X HBSS.

Cells were harvested fresh and quantified using a haemocytometer,  $2 \times 10^6$  cells of which were rinsed with PBS before being resuspended in 30 $\mu$ L of 1x HBSS (Gibco). Using a glass syringe,  $2 \times 10^6$  cells were transferred each into the closed instrument chambers (A and B). A basal measurement of oxygen respiration was taken. Oxygen consumption rates were expressed as specific oxygen ( $O_2$ ) flux per cell (pmol  $O_2$  per million cells (Mil) per second(s))<sup>173</sup>. Oligomycin (5mg/ml) was then added to block complex V (ATP synthase) activity and determine the rate of uncoupled respiration that arises as a result of proton leak from the electron transfer chain. Maximal respiratory capacity was assessed using carbonyl cyanide p-(trifluoromethoxy)-(phenylhydrazine) (FCCP) (Sigma), following titration of the inhibitor at various concentrations between 10-100nM. Lastly, 5mM of antimycin A, an inhibitor to Complex III, was used to determine background levels of respiration that arise from non-mitochondrial sources. In between the addition of cells and each of the inhibitors, there was a waiting interval of 5 to 10 minutes to ensure complete equilibration before each step.

### 2.15. DNA cloning

Conventional PCR reactions were set up to amplify the region of interest in preparation for insertion into a plasmid vector. PCR products were gel purified and ligations into the pCR2.1 vector (Invitrogen) were performed immediately after PCR using reagents supplied with the TA Cloning Kit (Invitrogen): 1x ligation buffer, 50ng of pCR2.1 vector, 4 Weiss Units of T4 DNA ligase together with a 1:1 vector:insert ratio for the PCR product, to a final volume of 10µl with sterile ddH<sub>2</sub>O. The formula below was used to calculate the amount of PCR product required for the ligation reaction with the pCR2.1 vector:

$$X = \frac{Y \times 50}{3900}$$

Where X represents the amount (µL) of PCR product of Y (bp) that is required for a ligation at a 1:1 vector:insert molar ratio. The value 50 represents the concentration (ng) of the pCR 2.1 vector, and the value 3900 represents the total size of the pCR 2.1 vector.

Ligations were incubated at 12.5°C overnight, and were confirmed by gel electrophoresis with M13 primers (Forward: 5'→3' CTGGCCGTCGTTTTAC; Reverse: 5'→3' GTCATAGCTGTTTCCTG) designed to span the insert region. For long-term storage, the ligation product was kept at -20°C conditions.

Once the ligation efficiency had been confirmed, 50µl of 20mg/ml 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-GAL) solution (Bioline) was added and evenly spread across Luria-Bertani (LB) agar plates containing 100µg/ml ampicillin (Sigma-Aldrich).

DH5 $\alpha$  competent bacterial (*E. coli*) cells from the TA Cloning Kit (Invitrogen) were thawed on ice for 5 minutes and 2 $\mu$ l of ligation product was added directly to this. The mixture was left to incubate for 30 minutes on ice, then the cells were heat-shocked at 42°C for 30 seconds, and transferred immediately back onto ice. 250 $\mu$ l of room temperature S.O.C media was added to the cells, and the bacteria were incubated at 37°C for 1 hour with shaking at 225rpm in tubes that were permissive to O<sub>2</sub> exchange. 50 $\mu$ l of transformed bacteria was then plated onto LB ampicillin plates that had been previously coated with X-gal. Plates were inverted and left for overnight incubation at 37°C.

White colonies were selected and screened by PCR using M13 primers. Products of the anticipated molecular size were excised from the gel and purified prior to capillary sequencing. Sequences were checked for the presence of the gene region of interest, and also for orientation of the insert. Colonies of interest were expanded further overnight at 37°C with shaking at 220rpm, in 5ml LB broth containing 100 $\mu$ g/ml ampicillin (Sigma-Aldrich).

#### **2.15.1. Plasmid purification**

Plasmids were purified using the QIAprep Spin Miniprep Kit (Qiagen). 1.5ml of bacterial culture expanded in LB broth overnight (section 2.15) was transferred to a 1.5ml Eppendorf tube and centrifuged for 15 minutes at 8000rpm. The resulting bacterial pellet was resuspended in 300 $\mu$ l of Buffer P1 (resuspension buffer) containing 100 $\mu$ g/ml RNase A. 300 $\mu$ l of lysis buffer (Buffer P2) and

mixed thoroughly by inverting the tube 6 times. The mixture was then incubated at room temperature for 5 minutes.

300µl of Buffer P3 containing potassium acetate was added and mixed by inversion 6 times. The tube was then subjected to incubation on ice for 5 minutes and centrifuged at 13,000rpm for 10 minutes. The supernatant containing plasmid DNA was carefully extracted and applied to a QIAprep spin column. This was then centrifuged for 1 minute at 13,000rpm, and the flow-through was discarded.

The column was washed with 500µl of Buffer PB containing isopropanol and guanidine hydrochloride (GuCl). The column was centrifuged for 1 minute at 13,000rpm and the resultant flow-through discarded. 750µl Buffer PE containing ethanol and Tris-Cl was applied to the spin column and centrifuged for 1 minute at 13,000rpm. The flow-through was discarded and the empty column was centrifuged to dry the membrane. Elution was performed in 30µl using sterile water by centrifuging for 1 minute at 13,000rpm.

#### **2.16. High resolution melt (HRM) analysis**

All primers used for HRM curve analysis generated an amplicon size of no more than 250bp. For each primer pair specific to the gene of interest, an optimization step was performed using gradient PCR on the MultiGene™ Gradient Thermal Cycler (Labnet International) to determine the optimal annealing temperature at which primers will bind to and therefore amplify the template DNA. All HRM reactions were conducted in triplicate on a 96-well plate, with 10ng/µl template

DNA, 2.5 $\mu$ M of forward and reverse primers, 1x HRM master mix containing LCGreen® Plus+ (TrendBio) and sterile water to a final volume of 10 $\mu$ l. 20 $\mu$ l of mineral oil (Sigma) was added to each sample well to prevent evaporation of the sample, and a plastic seal was then applied over the plate after which the plate was centrifuged for 1 minute at 1000rpm.

The PCR protocol for primer optimization was as follows: 95°C for 2 minutes, 45 cycles of 94°C for 30 seconds and 60-72°C for 30 seconds, followed by 1 cycle at 94°C for 30 seconds (to allow for heteroduplex formation), and lastly the final products were held at 25°C before further processing. The 12 temperatures that were tested during optimization included 60°C, 60.7°C, 61.3°C, 62.9°C, 64.7°C, 66.4°C, 67.2°C, 68.7°C, 70.1°C, 71.3°C, 71.6°C and 72°C.

The plate containing the amplified products was analysed using the LightScanner (Idaho Technologies, Salt Lake City, Utah) with data collection performed within the LightScanner Instrument & Analysis software with Call-IT 2.0 (V.2.0.0.1331). The starting temperature for the melt analysis was set to 70°C with the range ending at 96°C with increases at 0.1°C increments. Upon completion of the run, 7 $\mu$ l of the sample from the optimization run was taken out from the wells to represent each of the different temperatures (60-72°C) and analysed by gel electrophoresis on a 2.5% 1x TAE agarose gel (at voltage 100V) to determine the optimal primer binding temperature. Once identified, the above method was applied to test samples but instead of using a specific annealing temperature for the PCR protocol, with the same settings being applied to the LightScanner instrument.

### **2.17. Capillary sequencing**

Samples up to 1000bp in length were analysed with the 16-capillary 3130xl Genetic Analyser fitted with a 80cm array (Applied biosystems). The amount of template used for sequencing depended on fragment size: 10ng for products between 100-200bp, 25ng for those between 200-500bp and 50ng for fragments 500-1000bp. Template was added to 3.2 $\mu$ M of either the forward or reverse primer, and the total volume was made up to 16 $\mu$ l with sterile water.

Processing of the samples was conducted at The Gandel Charitable Trust Sequencing Centre (Monash Institute of Medical Research, Clayton), using the BigDye® Terminator v3.1 Cycle Sequencing kit, which uses dideoxy terminators tagged with one of four different fluorescent dyes to generate products of different length that can be separated according to size by electrokinetic injection. The procedure is as follows:

For each reaction 1 $\mu$ l of BigDye v3.1 was added, together with 3.5 $\mu$ l of 5x BDT buffer to 15.5 $\mu$ l of sequencing template, as prepared above. The total reaction volume of 20 $\mu$ l was then subjected to the following cycle sequencing conditions: 96°C for 1 minute, 25 cycles of 96°C for 10 seconds, 50°C for 5 seconds and 60°C for 4 minutes. Products were then held at 4°C until the purification stage.

Purifying extension products were performed using ethanol/sodium acetate precipitation. Firstly, an ethanol/sodium acetate mix was prepared by adding 2 $\mu$ l of sodium acetate (pH4.6) to 50 $\mu$ l of 95% ethanol. The 20 $\mu$ l sequencing reaction from the thermal cycler was added to the 52 $\mu$ l ethanol/sodium acetate mix. This

was vortexed briefly and left to precipitate for 15 to 30 minutes at room temperature. The mix was then centrifuged at 13,000rpm for 20 to 30 minutes, and the supernatant carefully aspirated until all ethanol had been removed. 250 $\mu$ l of 70% ethanol was then added, vortexed and centrifuged for 10 to 15 minutes at 13,000rpm. Again, the supernatant was aspirated to remove all traces of ethanol, and the tubes were left to dry on a heat block set at 90°C for 2 to 5 minutes. The sample pellet was stored at 4°C and protected from light until electrophoresis was performed.

For electrophoresis, the purified sample pellet was resuspended in 22 $\mu$ l of deionised formamide to maintain single-stranded DNA. This was vortexed and centrifuged briefly before heating to 90°C for 2 minutes, after which the sample was transferred immediately onto ice. The sample was then loaded on to a 96-well plate, before placing it into the 16-capillary 3130xl Genetic Analyser. Data were analysed using the Foundation Data Collection software version 3.0, and parameters suitable for the polymer type (POP-7) were also set using this software.

The digital data were visualized in the form of trace sequence files using the software, 4Peaks, version 1.7.2 (Neefix laboratories).

## **2.18. Next generation sequencing**

Procedures describing the library preparation are covered in sections 2.18.1 through to 2.18.7. Sections 2.18.8 onwards describe the procedures for performing automated emulsion PCR, recovery and enrichment of the amplified

DNA in preparation for sequencing on the Ion Torrent Personal Genome Machine (PGM).

### **2.18.1 Quantification of double-stranded DNA**

Overlapping products from long PCR covering the entire mitochondrial genome were quantified using the Qubit 2.0 Fluorometer. For each sample, 1 or 2 $\mu$ L of DNA were mixed with the Qubit Working solution (prepared by diluting the Qubit reagent 1:200 in Qubit buffer) and made to a total volume of 200 $\mu$ L, vortexed and then incubated for 2 minutes at room temperature in the dark. The Qubit instrument was calibrated beforehand using the set of standards provided with the Qubit assay kit to produce a linear regression graph plotting fluorescence against concentration. The standards were prepared by adding 10 $\mu$ L of each standard to 190 $\mu$ L of Qubit Working Solution. The concentration of the test samples was then measured taking into account the volume of the sample and fluorescence emitted. The formula for calculating the concentration is:

$$\text{Sample concentration } (\mu\text{g/mL}) = \text{QF value} \times (200/x)$$

Where the QF value refers to the fluorescence value retrieved after assessment using the Qubit fluorometer, and the value x refers to the amount of sample (in  $\mu$ L) added to the Qubit Working Solution.

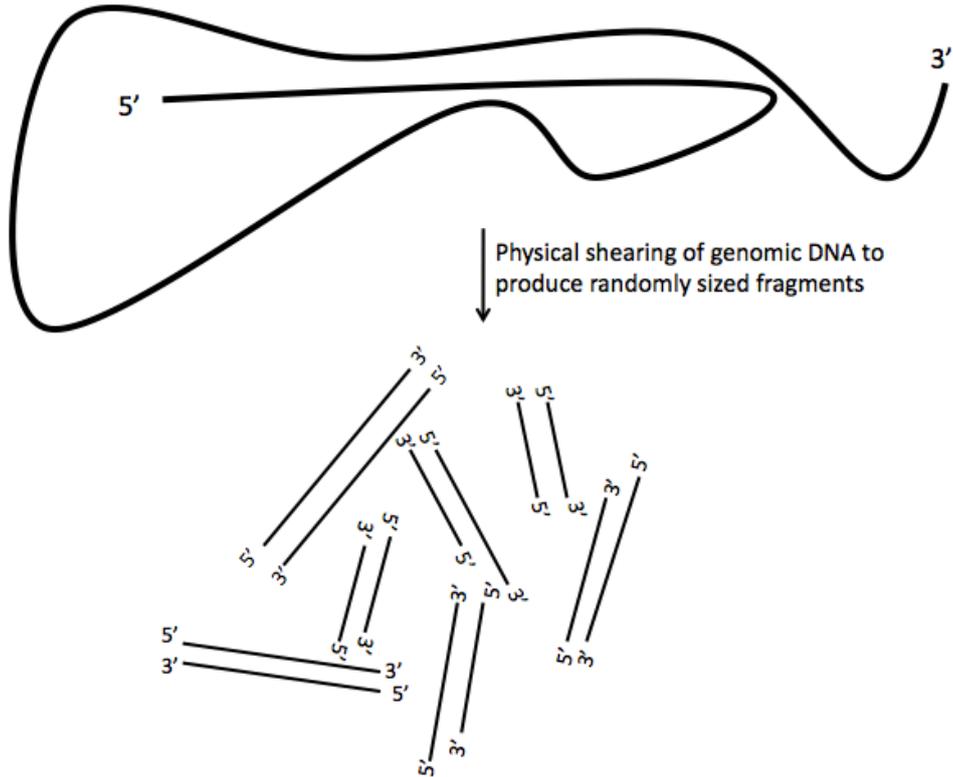
After quantification, both products were then pooled at equimolar concentrations to ensure even coverage of the mtDNA.

### **2.18.2. Fragmentation of the amplicons from long PCR**

Firstly, the pooled long PCR products were made up to a volume of 130 $\mu$ L in low TE Buffer and transferred to a Covaris microtube. Fragmentation was performed using the Covaris Adaptive Focused Acoustics (AFA<sup>TM</sup>) S220 system, within a full de-ionised water tank degassed at 4°C (**Figure 2.1**). The Covaris was operated via the Covaris Sonolab 7 software, and treatment parameters were as follows for the generation of 200bp fragments, as recommended by the manufacturer:

peak power (W) = 175.0, duty factor = 10% and cycles per burst = 100, treatment time (seconds) = 60 and number of cycles = 3.

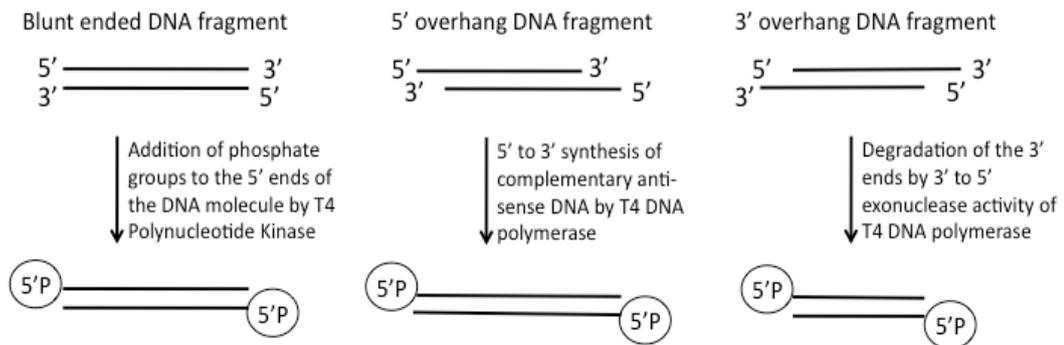
The duty factor (or duty cycle) refers to the percentage time the transducer transmits acoustic energy. Peak power (or intensity) refers to the amplitude of the waves and correlates with acoustic power. Cycles per burst represent the number of energy waves in a single burst.



**Figure 2.1.** Physical shearing of genomic DNA to produce random sized fragments.

### 2.18.3. End repair of the fragmented DNA and sample purification

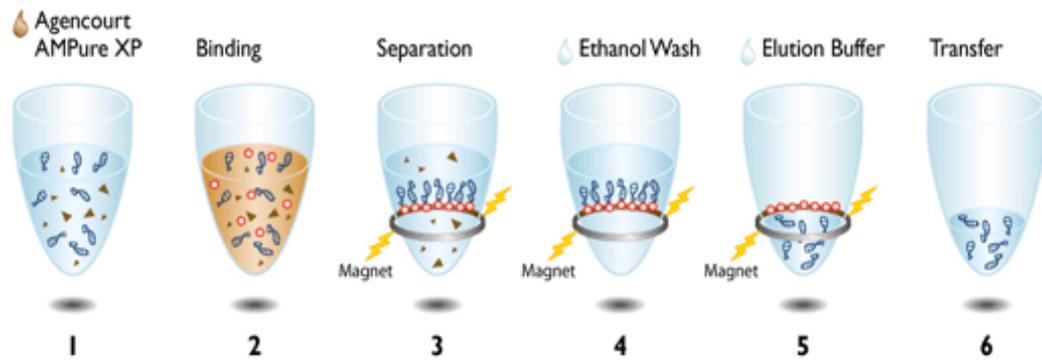
After fragmentation, the blunt-ended DNA is end-repaired to generate blunt-ended DNA possessing 5' phosphate and 3' hydroxyl groups (**Figure 2.2**).



**Figure 2.2.** End repair of blunt-ended DNA molecules

For the end repair procedure, 79 $\mu$ L (approximately 100ng) of the sheared DNA was transferred into a 1.5ml LoBind Tube, and 1X End Repair Buffer and 1 $\mu$ L of End Repair Enzyme was added and mixed with the sample, giving a total reaction volume of 100 $\mu$ L. This was incubated for 20 minutes at room temperature, and purification of the end-repaired DNA was performed using the Agencourt AMPure XP Kit in preparation for the ligation of Ion adapters to the DNA, which will be used to amplify the DNA.

Agencourt AMPure XP Reagent beads (Beckman Coulter, Inc, NSW, Australia) were added at a ratio of 1.8 to the sample volume of 100 $\mu$ L (total volume at this stage = 280 $\mu$ L) (**Figure 2.3**). The bead suspension and DNA were mixed by pipetting 5 times, followed by pulse centrifugation and incubation for 5 minutes at room temperature. After this, the sample was again pulse centrifuged and placed in a DynaMag-2 magnetic rack (Life Technologies) for 3 minutes until the magnetic bead suspension has been separated from its solution. The supernatant was removed and discarded, following which the magnetic reagent beads were resuspended in 500 $\mu$ L of fresh 70% ethanol for the first wash procedure, without removing the sample tube from the DynaMag-2 magnetic rack. The samples were incubated for 30 seconds, and then rotated twice in the rack to mobilise the reagent beads. Once the solution clears, the supernatant was again removed and the wash step was repeated a second time with 70% ethanol. For removal of residual ethanol, the sample tubes were pulse-centrifuged and loaded onto the DynaMag-2 magnetic rack before removing the supernatant.



**Figure 2.3.** DNA purification using the Agencourt AMPure XP kit. Image taken from: <https://www.beckmancoulter.com/wsrportal/wsr/research-and-discovery/products-and-services/nucleic-acid-sample-preparation/agencourt-ampure-xp-pcr-purification/index.htm> Accessed 19/09/2013.

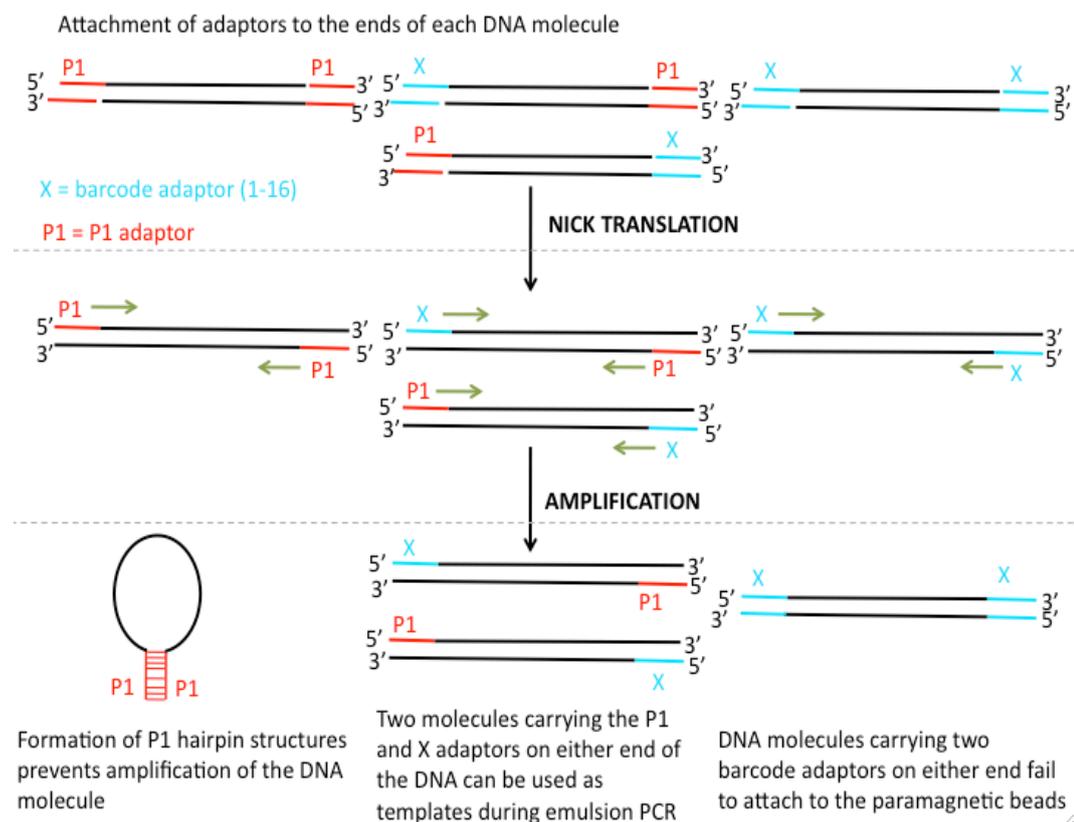
The reagent beads were air-dried at room temperature for up to 5 minutes, and then the sample tubes were removed from the rack in preparation for the elution step. 25µL of low TE was added to the samples, and mixed by pipetting 5 times. The sample was then vortexed for 10 seconds, pulse centrifuged and placed in the magnetic rack for 1 minute. The supernatant containing the eluted DNA was then collected into a new 0.2ml PCR Tube.

### 2.18.4. Ligation of adaptors, nick translation and DNA sample purification

Adaptor ligation is performed by nick translation, which allows for one adaptor (P1) to ligate to the 5' end of all DNA molecules and another adaptor to ligate to the 3' end (Figure 2.4, Table 2.1).

**Table 2.1.** Sequences of the Ion ligation adaptors

Adaptor	Terminal end	Sequence
P1	5'	CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGGTGAT
	3'	T*T*GGTGATGCGGAGGCGAAAGGAGAGATACCCGTCAGCCA CTA
A	5'	TTCCATCTCATCCCTGCGTGTCTCCGACTCAG
	3'	T*T*AAGGTAGAGTAGGGACGCACAGAGGCTGAGTC



**Figure 2.4.** Nick translation for the ligation of sequencing adaptors. Adaptors were ligated to either end of the DNA molecules before amplification of the DNA templates.

The reaction setup for ligation of the adapters used the following reagents added to the eluted (25 $\mu$ L) end-repaired and purified DNA in a 0.2ml PCR tube: 1X Ligase Buffer, 2 $\mu$ L of Ion P1 Adapter, 2 $\mu$ L Ion Xpress Barcode X( $\diamond$ ), 2 $\mu$ L dNTP mix, 2 $\mu$ L DNA Ligase and 8 $\mu$ L of Nick Repair Polymerase to a total reaction volume of 100 $\mu$ L by addition of nuclease-free water. The samples were mixed and loaded into a thermal cycler programmed to run using the following conditions: Hold for 15 minutes at 25°C (for nick translation), hold for 5 minutes at 72°C (heat-activation of the polymerase) and hold at 4°C. The reaction was immediately transferred into a 1.5ml LoBind Tube for purification of the adapter-ligated and nick repaired DNA.

( $\diamond$ ) X represents any of the 1-16 different barcodes chosen available with the kit.

For the purification procedure in preparation for creating a 200 base-read library, Agencourt AMPure XP Reagent was added at a volume 1.4x (140 $\mu$ L) to the sample volume (total volume at this stage = 240 $\mu$ L). The remainder of the protocol is the same as mentioned previously under section 2.18.3, with the exception that the elution step was performed in 20 $\mu$ L of low TE buffer. The supernatant containing the eluted DNA is collected and transferred to a 0.2 $\mu$ L PCR tube in preparation for the next step.

#### **2.18.5. Size-selection of the unamplified library**

Size-selection of the DNA library was used to select DNA fragments of the desired length for an unbiased amplification of the libraries. Size-selection was performed using a 2% E-Gel SizeSelect Agarose Gel system (Life

Technologies). 20 $\mu$ L of the purified samples were loaded onto alternate wells to avoid cross-contamination of DNA libraries. The 50bp reference DNA ladder (1 $\mu$ g/ $\mu$ L) provided with the kit was diluted 1:40 in low TE buffer. 250ng (10 $\mu$ L) of this was loaded onto the gel. Empty wells at the top of the gel were loaded with 25 $\mu$ L of deionized water. All collection wells at the bottom of the gel were loaded with 25 $\mu$ L of deionised water, with the exception of the middle lane, used for electrophoresis of the DNA reference ladder, which was loaded with 5-10 $\mu$ L of deionised water. Electrophoresis settings for the “Run Time to Reference Line” were set for 11 to 14.5 minutes, monitored regularly and stopped when the band of interest reached the reference line. Once this was achieved, the collection wells at the bottom of the gel were refilled to 25 $\mu$ L with sterile water. Electrophoresis time for the “Run Time from Reference Line to Collection Well” was set for 1 to 2 minutes. Towards the end of the run, occasionally the band of interest can be seen to migrate into the collection well, which can then be collected using a pipette. If the DNA band of interest has migrated past the collection well point, the ‘Reverse E-Gel’ program was run to reverse migrate the band into the collection well.

#### **2.18.6. Amplification of the 200 base-read DNA libraries and purification**

25 $\mu$ L of the size-selected unamplified library was used in combination with the following reagents to allow for amplification of the 200 base-read purified DNA libraries: 100 $\mu$ L of Platinum PCR SuperMix High Fidelity mastermix and 5 $\mu$ L of Library Amplification Primer Mix (All from Invitrogen), giving a total reaction volume of 130 $\mu$ L. The Platinum PCR SuperMix High Fidelity is composed of a mixture of recombinant Taq DNA polymerase, magnesium, dNTPs and salts.

The reaction mix (130 $\mu$ L) was transferred in equal volumes between two 0.2ml PCR tubes (65 $\mu$ L per 0.2ml PCR tube) and placed into a thermal cycler for amplification under the following conditions: Hold at 95°C for 5 minutes (a minimum time of 2 minutes is required for denaturation of the template and activation of the polymerase); 6 cycles of 95°C for 15 seconds, 58°C for 15 seconds and 70°C for 1 minute; products were then held at 4°C at the end of the reaction.

The DNA libraries were purified by addition of Agencourt AMPure XP Reagent at a volume of 1.5x the sample volume. The remainder of the protocol is the same as for section 2.18.3, with the exception that the elution step was performed in 20 $\mu$ L of low TE buffer. The supernatant containing the DNA was transferred to a new 0.2ml PCR Tube.

#### **2.18.7. Quantification of the DNA libraries**

A 1 $\mu$ L sample of the DNA library underwent analysis on the Agilent 2100 Bioanalyzer using the Agilent High Sensitivity DNA Kit (Agilent Technologies, Santa Clara, CA). First, the gel-dye mix was prepared. High Sensitivity DNA dye concentrate (stored in the dark) and High Sensitivity DNA gel matrix reagents were equilibrated for 30 minutes at room temperature. The dye concentrate was mixed by vortexing for 10 seconds followed by pulse centrifugation. 15 $\mu$ L of this was transferred into the gel matrix solution. The mixture was vortexed for 10 seconds, and the complete contents were transferred to a spin column. This was centrifuged at 6000rpm for 10 minutes at room

temperature. The spin filter was discarded and the gel-dye mix was kept in the dark (stored at 4°C for a maximum of 6 weeks).

A High Sensitivity DNA chip was set up on the chip priming station. 9µL of the room temperature gel-dye mix was loaded at the bottom of a single well marked with a circularised 'G' (markings are pre-labelled on the chip). The DNA chip was loaded onto the chip priming station and subjected to pressure by closing the chip priming station. After pressurisation, the chip priming station is opened and, 9µL of the gel-dye mix was then pipetted into the remaining three 'G' wells. 5µL of the High Sensitivity DNA marker was pipetted into the well marked with a ladder, in addition to the 11 sample wells. 1µL of High Sensitivity DNA ladder was loaded onto the chip, alongside 1µL of each sample (wells 1 to 11). In the case of empty wells, 1µL of marker was used instead. The chip was positioned horizontally in the holder of the IKA vortex mixer (Model MS3) to ensure minimal disruption to the chip during vortexing. The chip was vortexed for 60 seconds at 2400rpm, after which, the Agilent 2100 Bioanalyzer was set up and the chip placed into the instrument chip holder. The run was initiated using the Agilent 2100 expert software.

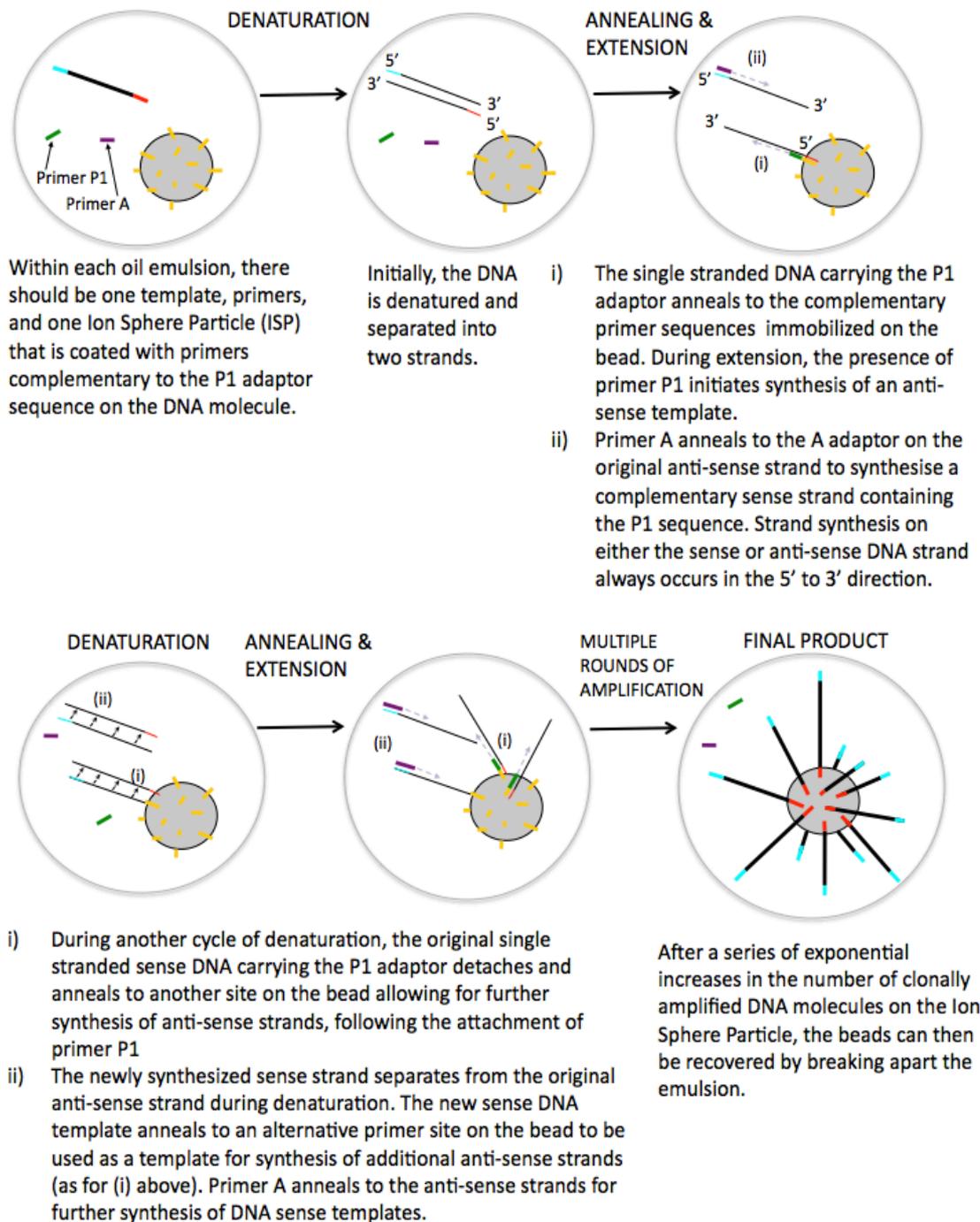
Electropherogram results were analysed, with fragment size indicated by the elution time and sample integrity indicated by the presence of well-defined peaks. Quantification of DNA concentration was possible through comparisons performed against the internal standards.

Once DNA libraries had passed this quality check process, the DNA libraries were diluted to 9pM, which was determined through optimisation to provide the highest amount of DNA per bead without generating polyclonal beads.

#### **2.18.8. Emulsion PCR on the Ion OneTouch instrument**

Before preparing the amplification solutions for the emulsion PCR step, the Ion OneTouch was set up by attaching one Ion OneTouch Reagent Tube containing Ion OneTouch Oil to the left side of the instrument, and another Reagent Tube containing Ion OneTouch Recovery solution to the right side. Two OneTouch Recovery Tubes were then installed into the front slots of the instrument to later collect the Ion Sphere Particles (ISPs) containing amplified DNA template.

To prepare the emulsion PCR (amplification) mix, the following reagents were added to a 1.5mL Eppendorf LoBind Tube: 280µL Nuclease-free water, 500µL Ion OneTouch 2X Reagent Mix, 100µL Ion OneTouch Enzyme Mix and 20µL of DNA library diluted to 9pM. This was mixed by pipetting before adding 100µL of Ion OneTouch ISPs (total volume of 1mL). The Ion OneTouch Reaction Filter was assembled by attaching a Reaction Tube and a 15mL conical tube to the Filter. 1mL of the amplification mix was loaded slowly into the sample port on the Filter. Following which, an additional 1.5mL of Ion OneTouch Reaction Oil was slowly added through the same sample port on the Reaction Filter. The filled Ion OneTouch Reaction Filter was fitted on to the Ion OneTouch instrument to begin the emulsion PCR reaction (**Figure 2.5**).



**Figure 2.5.** Emulsion PCR procedure illustrating the amplification of DNA template on the ISPs.

### **2.18.9. Recovery of the template-positive ISPs**

Upon completing the run, both Recovery Tubes loaded at the front of the instrument, which contain the ISPs with amplified DNA template were carefully removed. Without disturbing the pellet of template-positive ISPs, all but 50 $\mu$ L of the Ion OneTouch Recovery Solution was discarded from each Recovery Tube. The pellets from each Recovery Tube were resuspended in the remaining 50 $\mu$ L by pipetting, and both suspensions were transferred to a fresh 1.5mL LoBind Tube. 1mL of Ion OneTouch Wash Solution was added to the template-positive ISPs, and the tube centrifuged for 2.5 minutes at 15,500 x g. The supernatant was discarded, leaving behind a volume of 100 $\mu$ L, which was vortexed to resuspend the ISPs.

### **2.18.10. Enrich the template-positive ISPs**

'Melt-Off' Solution was prepared by combining 865 $\mu$ L of nuclease-free water, 125 $\mu$ L of 1M NaOH and 10 $\mu$ L of 10% Tween-20 (diluted in water) to produce a final concentration of 125mM NaOH with 0.1% Tween-20. 130 $\mu$ L of MyOne Dynabeads Wash Solution was pipetted into a fresh 1.5mL LoBind Tube. 13 $\mu$ L of Dynabeads MyOne Streptavidin C1 Beads was added to the Wash Solution and mixed by pipetting. The tube was loaded into a DynaMag-2 magnetic rack and incubated for 2 minutes to separate and discard the supernatant. Dynabeads were re-suspended with 130 $\mu$ L of Wash Solution and mixed by pipetting.

100 $\mu$ L of the re-suspended ISPs containing amplified DNA was transferred into the first well of an 8-well strip. 130 $\mu$ L of re-suspended Dynabeads was added to well 2, 300 $\mu$ L of Ion OneTouch Wash Solution was added to wells 3 to 5, wells

6 and 8 were left empty, and finally 300µL of previously prepared ‘Melt-Off’ Solution was loaded into well 7. The 8-well strip was fitted onto the Ion OneTouch Tray on the instrument. The Ion OneTouch Enrichment System was prepared by placing a new tip over the Tip Loader and placing an opened sterile 0.2mL PCR tube under the Tip Arm. To begin the automated enrichment process, ‘start’ was selected on the Ion OneTouch instrument. At the end of the run, the PCR tube containing the enriched ISPs was removed in preparation for the rinsing step.

#### **2.18.11. Rinsing of the enriched ISPs**

The enriched ISPs were centrifuged for 1.5 minutes at 15,500 x g. To ensure there is the absence of Dynabeads MyOne Streptavidin C1 Beads in the enriched ISP mix, the tube was loaded in to the DynaMag-2 magnetic rack for 4 minutes to separate the beads from the ISPs and to transfer the solution to a new tube before repeating the centrifugation step. ISPs were washed by discarding all but leaving 10µL of supernatant in the tube, followed by the addition of 200µL of Ion OneTouch Wash Solution, which was mixed by pipetting. The tube was centrifuged for 1.5 minutes at 15,500 x g and again the supernatant was discarded, leaving behind 10µL volume. 90µL of Wash Solution was added and mixed with the ISPs by pipetting.

#### **2.18.12. Preparation of the enriched ISPs for sequencing**

To prepare the enriched ISPs for sequencing, the sequencing primer is first annealed to the enriched beads. For barcoded samples to be multiplexed and run on a 316 chip, the total volume of the sample was transferred to a PCR tube.

Control ISPs are resuspended by vortexing and briefly centrifuged, before an aliquot of 5µL was transferred to the tube containing the enriched ISPs. A further 100µL of annealing buffer was added to the tube and mixed by pipetting. The tube was then subjected to centrifugation for 1.5 minutes at 15,500 x g, before removing the top supernatant and leaving behind a volume of 15µl containing the ISPs. 12µl of sequencing primer was then added and the pellets resuspended by pipetting. The tube was placed in a thermocycler at 95°C for 2 minutes followed by 37°C for 2 minutes. Immediately after, 3µl of sequencing polymerase was added, and the tube incubated at room temperature for 5 minutes.

The enriched ISPs were loaded onto the Ion Torrent sequencing chip, and the number of sequencing cycles was set to 125 cycles on the instrument for a 200bp sequencing run.

### **2.19. Detection of mtDNA variants**

Detection of variants, insertions and deletions of the genomic sequences was performed using CLC Genomics Workbench (v5.5.1). Reads were filtered to be above a 0.05 quality threshold level, (which accepts a base if its Phred quality score is above 15) and those below a read length of 15 nucleotides were excluded. The parameters applied when mapping the sequence to a reference genome (human mitochondrial Cambridge Reference Sequence; NC\_012920.1) included setting a mismatch cost of 2, and an insertion and deletion cost of 3. This meant that the individual score for a particular read, for example one that contained a single mismatch together with a single nucleotide deletion, would have a cost score of minus 5. If the read length in this scenario were 100bp, a

total cost of 5 would give the overall read a final percentage value of 95% identity. Applying these cost values then influenced whether the reads were considered acceptable for mapping, as only those reads containing at least 80% identity to the reference sequence were accepted into the final mapping.

For detection of single nucleotide polymorphisms (SNPs), further parameters were applied in addition to the above. The window length, which extends quality assessment to the neighbouring bases either side of the potential SNP position, was set to 7. The maximum number of gaps and mismatches allowed within this window length was set to 5, with the minimum quality of the potential SNP and its surrounding bases set to 15; and quality assessment of which is based on Phred score algorithms. The minimum coverage permitted for SNP detection was set to at least 5 reads, and the minimum variant frequency was set to 3%. All sequences were then filtered to remove duplicate reads that would otherwise over-estimate the variant frequencies. Finally, for a variant to be considered true, each was manually assessed to reside on both a forward and reverse read strand.

## **2.20. Predictive *in silico* analysis**

SNPs&GO (University of Bologna) is a web-based application used to predict whether a non-synonymous amino acid change causes disease using reliability index values. Results obtained are based on a combination of factors, inclusive of gene ontology, annotations and evolutionary information retrieved from the Swiss-Prot protein database <sup>174</sup>. MutPred (v.1.2; Buck Institute, Indiana University), on the other hand, is another *in silico* predictive tool, which calculates whether a non-synonymous amino acid substitution induces a

detrimental effect on protein function. It lists the top 5 features that are affected by the substitution and ranks these based on their probabilities for doing so <sup>175</sup>. Previous studies have used both SNPs&GO and MutPred as the preferred analytical tools for protein *in silico* analysis<sup>176,177</sup>.

### **2.21. Determining changes to the amino acid sequence**

The single letter abbreviated codes used to describe amino acid change in the coding regions of mtDNA are based on the International Union of Pure and Applied Chemistry (IUPAC) codes for mitochondrial-specific amino acids.

### **2.22. Determining susceptibility rates of the mtDNA regions to mutation**

To calculate the mutation rate for each of the mtDNA regions carrying variants, the number of mtDNA variants detected in each gene region was divided by the size of its corresponding mtDNA region (bp). This approach normalizes the different sizes of the mtDNA regions to allow for more effective comparison of base substitution rates across the mtDNA. To calculate the total mutation rate for the coding and non-coding mtDNA regions, the total number of mtDNA variants in each group was divided by, the total length of the corresponding coding or non-coding region (10,328bp and 6241bp, respectively). Throughout this thesis, the coding region is defined as a gene region on the mtDNA that encodes for a protein on the ETC. Likewise, the definition of a non-coding region in this thesis is a region that does not encode for a protein contributing directly to the ETC. This means that the ribosomal and transfer RNAs are categorized within the non-coding region alongside the regulatory D-loop in this thesis. Calculated values for mutation rates are all displayed to 3 significant figures.

### **2.23. Structural analysis of proteins**

SwissPDB viewer was used to visualize the locations of non-synonymous protein mutations. The protein structures used in this study were of bovine origin, and the PDB files for these are as follows: COX I: 1OCC\_A, COX II: 1OCC\_B, COX III: 1V54\_C, CYT B: 1BGY\_C, CYT C1: 1BGY\_D, Rieske protein: 1BGY\_E.

### **2.24. Phylogenetic analysis**

Alignment of multiple (>3) sequences was performed using the Clustal Omega tool (EMBL-EBI), with outputs saved in the PHYLIP format. Phylogenetic analysis was performed using sequential commands from the PHYLIP software (v.3.69). First, the program 'dnadist' was used to calculate pair-wise distances for all sequences, and to generate a distance matrix from this. Using these results, the program 'neighbor' was run to infer those distances calculated from dnadist, and to draw a phylogenetic tree via a neighbor-joining method that utilizes the F84 evolutionary model. Evolutionary models comprise a series of mathematical algorithms that attempt to take into account multiple substitutions of nucleotide bases at a single position of the sequence, and also bias from transition and transversion events. The F84 model uses two separate parameters to assess for differences in rates of change for both transition and transversion events, with rates exhibiting variability from one nucleotide base to another. 'Drawgram' was then used to produce a rooted tree for visual display of the outcomes.

### **2.25. Statistical analysis**

Graphical output and statistical analysis were performed using the Prism 5 software package version 5.0d (GraphPad Software, Inc., San Deigo, CA). For

analyses performed on more than 2 samples, a one-way ANOVA test was used followed by correction using the Bonferroni post-hoc test. Significance was considered when  $p < 0.05$ . All data are represented as mean  $\pm$  standard error of the mean (SEM).

**Chapter 3: Identification of MtDNA  
Variants in Patients Diagnosed with  
Multiple Myeloma**

### **3.1. Introduction**

The human genome was first sequenced over 10 years ago as part of the Human Genome Project, in an international collaboration involving 20 different research groups<sup>178</sup>. The main advantage of using next generation sequencing technology is primarily in the production of large volumes of data, with the genome being sequenced numerous times to generate a consensus template through assembly of short sequence reads. Coverage, defined as the number of times a particular nucleotide is sequenced by overlapping short fragments<sup>179</sup>, ensures accurate assembly of the genome. With the launch of various next generation sequencers, comparative studies have been performed on the different platforms<sup>180,181</sup>, with certain sequencers having characteristics that make them suitable for use in particular applications. Variability exists in terms of the length of the short sequence reads that could be used in the sequencing reactions, distribution of coverage for the genome and quality of the results obtained after the sequencing run.

In clinical practice, next generation sequencing technologies have already been implemented for a variety of solid cancers. Current projects include the Cancer Genome Atlas project, which anticipates genome-wide analysis of over 20 different cancer types. Sequencing data for several tumor types inclusive of glioblastoma multiforme<sup>182</sup> and melanoma<sup>183</sup> amongst others have been released in the last few years. In terms of cancers of haematological origin, acute myeloid leukemia<sup>184,185</sup> and chronic lymphocytic leukemia<sup>186</sup> have undergone whole genome sequencing, examining the effect of disease on the DNA sequence of nuclear genes. This same technique was also applied to multiple myeloma, in a

study involving 38 recruited individuals<sup>187</sup>. With support from the Multiple Myeloma Research Foundation (MMRF), the CoMMpass, is a large-scale study currently in progress, which aims to recruit 1,000 multiple myeloma patients within the next 5 years, in an effort to develop personalized treatments for patients [www.themmr.org].

To date, whole genome sequencing has been performed to determine sequence changes in the nuclear genome during the development of multiple myeloma at 4 stages, including diagnosis, first and second relapse phases and also during end-stage secondary plasma cell leukemia<sup>152</sup>. Findings of variants distinct to the different stages were able to identify potential variants that may be causative to the progression of tumor development at that particular stage. However, there are currently no studies that have explored the possibility of mtDNA in the initiation of multiple myeloma.

Over 100 multiple myeloma cell lines have so far been established<sup>188</sup>. One of the earliest and the best characterized is the U266 cell line, founded in 1968<sup>169</sup>. It was originally derived from the bone marrow of a patient diagnosed with IgE $\lambda$  secreting multiple myeloma. These U266 cells are reportedly characteristic of plasma cells not only in their morphology, but are also representative of these cells in their relapse, refractory or terminal phase of multiple myeloma<sup>169</sup>. Since their establishment, the cell line has undergone extensive cytogenetic and functional characterization<sup>189</sup>. Additionally, its immunoprofile has been well-defined<sup>188,190</sup>. For this reason, U266 cells were chosen as a model for multiple

myeloma in this Chapter, to determine whether there is a role for mtDNA in these tumor cells.

Previous publications have sequenced bacterial genomes using the Ion Torrent<sup>191</sup> and have demonstrated uniform coverage for reads across the genomes regardless of the proportion of GC content. MtDNA was evolutionarily derived from a bacterial genome, and due to their small genome size (16.6kb), the Ion Torrent was considered most suitable for the sequencing of mtDNA. Furthermore, the Ion Torrent benefits from high throughput (80-100Mb/h) and a short sequencing time of approximately 3 hours. However, it has been reported that the Ion Torrent is not sensitive enough in distinguishing sequence calls of any length within homopolymer regions<sup>181</sup>, and that a second validation technique should be adopted<sup>192</sup>.

Various approaches used by other studies to address this problem have included ignoring variants that appear within these homopolymer regions<sup>193,194</sup>. For studies that were unwilling to compromise with the lower sensitivity achieved from this approach, they resequenced these homopolymer regions numerous times and/or using different next generation sequencers to verify these variants<sup>195,196</sup>. However, for most research groups this is financially demanding. The approach chosen to overcome the possibility of reporting sequencing errors in this thesis, involved the use of high resolution melt (HRM) analysis to verify calls to support results retrieved from the Ion Torrent sequencer. This was because capillary sequencing was not sensitive enough to detect low-level variants below the 10% detection threshold<sup>197</sup>.

### **3.2. Aims and hypothesis**

The overall aim of this chapter was to determine whether changes to the mtDNA sequence could be responsible for initiating tumor development in patients characteristic of the relapse and/or refractory phase of multiple myeloma. Results were compared to an *in vitro* model of multiple myeloma to determine whether the outcomes were comparable.

It was hypothesized that the variants most likely to be responsible for the development of multiple myeloma would have high levels of mutation, and would likely alter critical mtDNA regions to promote mitochondrial respiration via glycolysis. Different multiple myeloma tumors were hypothesized to have mutations that are variable to one another, which would reflect differences between individual patients and their severity for multiple myeloma progression. However, these tumors were also anticipated to share a common set of cancer mutations that define multiple myeloma within their mtDNA sequence. Furthermore, it was probable that the majority of variants detected using an *in vitro* system could differ from those identified in patient tumor samples, primarily because of the presence of additional factors within an *in vivo* environment that could increase the tumorigenicity of these cancer cells. Nonetheless, it was anticipated that there may also be a small subset of common cancer mutations shared between cultured multiple myeloma cells and those obtained from patient tumors.

### **3.3. Preliminary experiments**

To prepare templates that could be used for sequencing of the whole mitochondrial genome using the Ion Torrent sequencing platform, long PCR was used. This ensures that the sequence amplified is purely representative of mtDNA. Moreover, other studies have used long PCR for whole genome sequencing of mtDNA<sup>198-201</sup>. Primers in this study were published previously<sup>202</sup>. The first of the two overlapping primer sets spanned from tRNA phenylalanine to the ATP6 region (for fragment A), and the second from the ATP6 to the 12S rRNA region (for fragment B). These were of product sizes 8382bp and 8704bp, respectively. However, optimisation of the PCR conditions was required.

#### **3.3.1. Optimisation of the long PCR**

Several enzyme systems were tested. These included the Platinum *Pfx* DNA Polymerase (Invitrogen), the *Pfx50* DNA Polymerase (Invitrogen) and the Platinum *Taq* DNA Polymerase (Invitrogen). Each of these enzymes were tested with or without PCR<sub>X</sub> Enhancer Solution (Invitrogen), which facilitates amplification of difficult templates that are rich in GC content by lowering the melting temperature of the DNA (**Figure 3.1**). Conditions for each of the reactions are listed below, with all made to a total volume of 50µL with sterile water:

(1) Platinum *Pfx* DNA Polymerase system:

1X *Pfx* Amplification Buffer, (1X PCR<sub>X</sub> Enhancer Solution), 15mM dNTP Mix, 1mM MgSO<sub>4</sub>, 2.5 units of Platinum *Pfx* DNA Polymerase enzyme and 15µM each of the forward and reverse primers.

(2) *Pfx50* DNA Polymerase system:

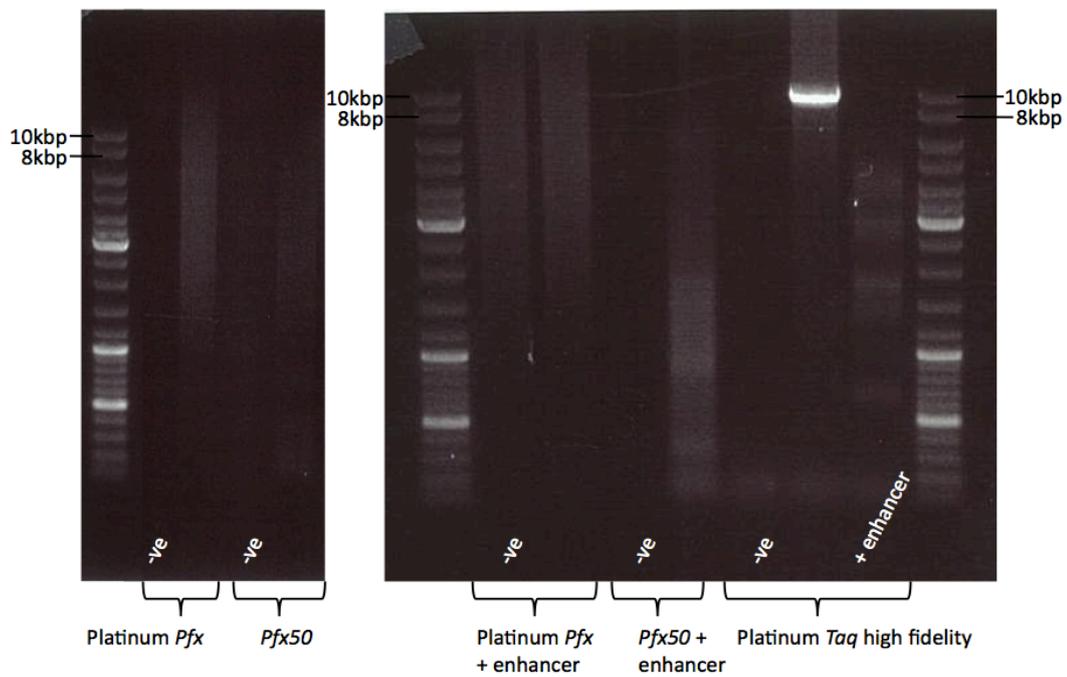
1X *Pfx50* PCR Mix (containing 1.2mM MgSO<sub>4</sub>), (1X PCR<sub>X</sub> Enhancer Solution), 15mM dNTP Mix, 5 units of *Pfx50* DNA Polymerase and 15μM each of the forward and reverse primers.

(3) Platinum *Taq* DNA Polymerase High Fidelity system:

1X High Fidelity PCR Buffer, (1X PCR<sub>X</sub> Enhancer Solution), 10mM dNTP Mix, 2mM MgSO<sub>4</sub>, 2.5 units of Platinum *Taq* High Fidelity enzyme and 15μM each of the forward and reverse primers.

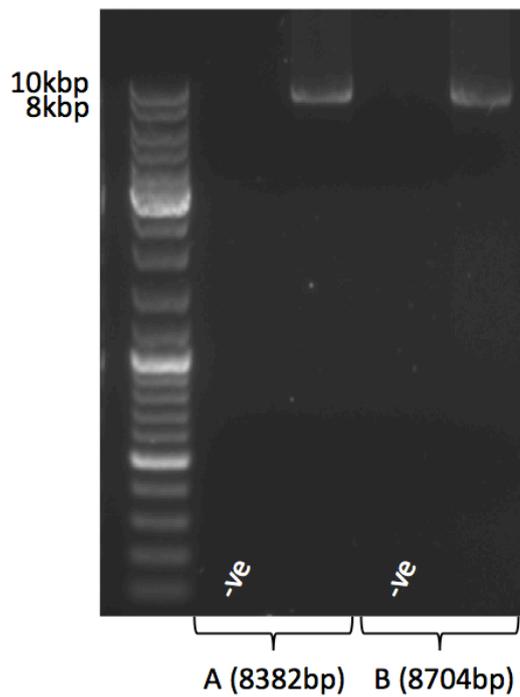
Cycling conditions for all enzyme systems were as follows: 94°C for 2 minutes, 35 cycles of 94°C 15 seconds, 63°C for 30 seconds<sup>#</sup> and 68°C for 8 minutes 45 seconds, with final extension of the products performed at 68°C for 10 minutes. Products were held at 4°C at the end of the reaction.

<sup>#</sup> If 1X PCR<sub>X</sub> Enhancer Solution was added, the annealing temperature was reduced to 61°C.



**Figure 3.1.** Long PCR optimisation. Amplification of one long PCR fragment was attempted using the Platinum Pfx DNA Polymerase, Pfx50 DNA Polymerase and Platinum Taq DNA Polymerase High Fidelity enzyme systems on human fibroblast (BJ) DNA. Each enzyme was tested in the presence and absence of 1X PCR<sub>X</sub> Enhancer Solution, which lowers the primer annealing temperature for low quality and/or GC-rich templates. Products were resolved by electrophoresis on a 0.7% agarose gel.

Results demonstrated that only the Platinum Taq DNA Polymerase High Fidelity enzyme system, without the presence of 1X PCR<sub>X</sub> Enhancer Solution, produced the amplicon of interest. To verify this, the reaction was repeated alongside the primers used to produce the second long PCR fragment (**Figure 3.2**).



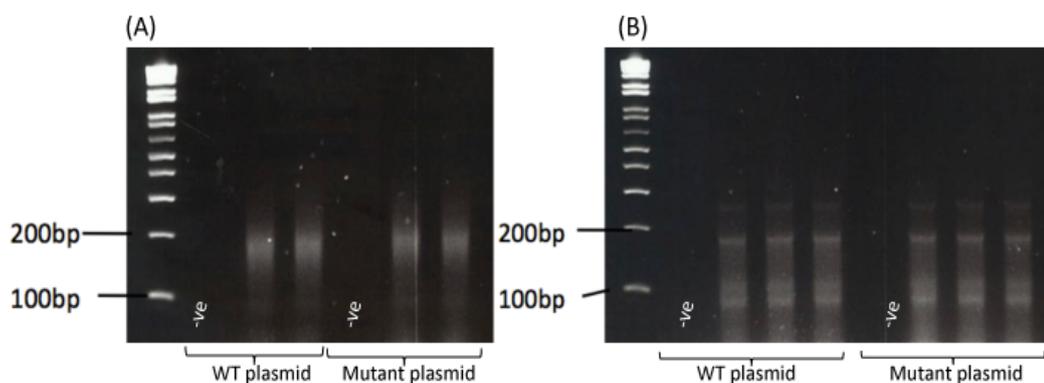
**Figure 3.2.** Long PCR amplification of human mtDNA. The Platinum Taq DNA Polymerase High Fidelity enzyme amplified two overlapping fragments (A and B) each of sizes 8382bp and 8704bp, respectively. PCR products were resolved on a 0.7% agarose gel.

Results confirmed that the Platinum *Taq* DNA Polymerase High Fidelity enzyme system is suitable for the generation of long PCR fragments for sequencing on the Ion Torrent platform.

### 3.3.2. Optimisation of the Ion Torrent sequencing workflow

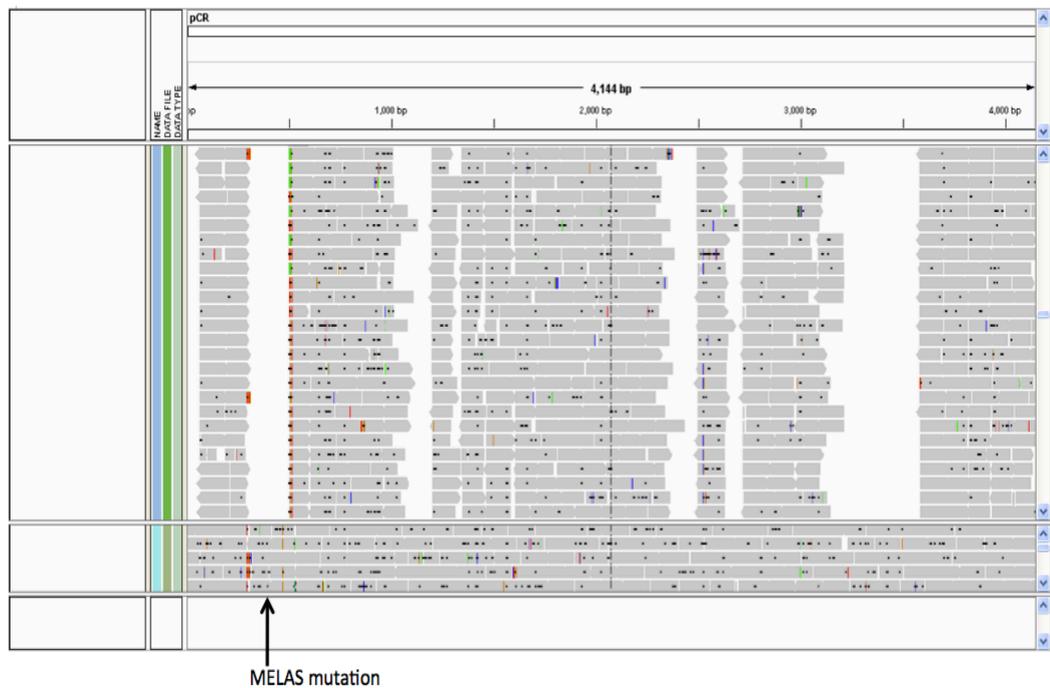
As part of the library preparation for next generation sequencing, DNA is sheared so that the randomly sized short fragments can be assembled to generate a consensus sequence. This can be done either mechanically or enzymatically. Both these methods were investigated to determine the optimal workflow for generating DNA libraries suitable for next generation sequencing analyses.

Recombinant plasmids containing either the wild-type or mutant insert for the Mitochondrial Encephalomyopathy, Lactic Acidosis, Stroke-like episodes (MELAS) mutation (A3243G), were sheared mechanically using the Covaris S220 system, or enzymatically using the Ion Xpress Plus Fragment Library Kit. Results were analysed on an E-gel to investigate the size distribution of fragments after both types of shearing (**Figure 3.3**).



**Figure 3.3.** Mechanical and enzymatic shearing of plasmid DNA libraries. (A) Covaris S220 system, (B) Ion Xpress Plus Fragment Library Kit. Plasmids containing the wild-type or mutant insert for the MELAS A3243G mutation were loaded in duplicate onto an E-gel system for electrophoresis. Sheared DNA of approximately 200bp is ideal for further processing and for sequencing on the Ion Torrent.

Results revealed that enzymatic shearing produced distinct DNA bands on the E-gel, which suggested that this method of shearing introduced an element of bias (**Figure 3.3B**). It is likely that the restriction enzymes cleaved DNA at specific sequence sites, which would not be ideal for sequencing as this would result in uneven coverage of the genome (**Figure 3.4**). On the other hand, the DNA smear produced from mechanical shearing indicated the presence of randomly sized fragments, the majority of which were of the ideal size (200bp) and therefore suitable for further processing and sequencing on the Ion Torrent (**Figure 3.3A**).



**Figure 3.4.** Sequence coverage of enzymatically sheared MELAS plasmid DNA. The image was taken from the Integrative Genomics Viewer (IGV) software (Broad Institute).

Automated and manual protocols for library preparation were also compared to identify the method that would minimise variability between sequencing runs on the Ion Torrent sequencer (**Table 3.1**). The Ion OneTouch system requires only

the manual loading of the DNA library and amplification master mix onto the instrument, as clonal amplification of template on the Ion Sphere Particles and sample recovery is automatically performed.

To perform this comparative test, mtDNA amplified from HSR-GBM1 cells was mechanically sheared and separated into six portions, three of which would be examined for reproducibility using the Ion OneTouch system, and the other three for the same using the manual preparation method.

The percentage level of mutation identified from libraries prepared using the manual method were more variable amongst the triplicate data sets, and had higher SEM values compared to those obtained from the Ion OneTouch system for each of the variants identified (**Table 3.1**). Greater variability in the manual preparation of sequencing libraries meant that the Ion OneTouch workflow was more consistent in between experimental sequence runs during the triplicate test.

Based on these preliminary experiments, it was determined that a combination of using the Covaris and the Ion OneTouch system, would provide the most appropriate workflow for generating high quality results from the Ion Torrent.

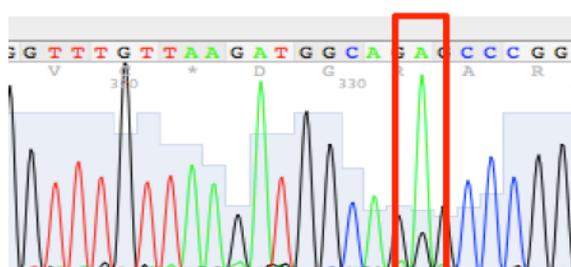
**Table 3.1.** Comparison of the automated and manual protocols for minimizing experimental variability. Labels 1, 2 and 3 indicate a triplicate data set.

		Percentage change in variant (%)										
		Ion OneTouch (automated)					Manual preparation					
	Reference Position	Variant change	HSR-GBM1 [1]	HSR-GBM1 [2]	HSR-GBM1 [3]	SEM	HSR-GBM1 [1]	HSR-GBM1 [2]	HSR-GBM1 [3]	SEM	Gene region	
Non-coding region	16218	C→T	15.7	21.2	17.6	<b>1.61</b>	15.9	26	19.2	<b>2.97</b>	D loop - Hypervariable segment 1, 7S DNA, membrane attachment site	
	1386	T→C	20.5	20.2	16.4	<b>1.32</b>	9.7	18.2	18.6	<b>2.9</b>	12s ribosomal RNA (rRNA)	
	2702	G→A	4.4	4.2	3.5	<b>0.27</b>	5.9	5	4	<b>0.54</b>	16s ribosomal RNA (rRNA)	
Coding region	6999	G→A	11	9.9	9.4	<b>0.47</b>	3	10.6	10.9	<b>2.58</b>	Cytochrome C Oxidase I: COX I	

### 3.3.3. Ion Torrent sequencing on control DNA samples.

Using the optimized sequencing workflow, mtDNA from two control DNA samples were sequenced, which included two human fibroblast cell lines. One is known to contain the MELAS mutation and the other is a commercially available (BJ) cell line. By sequencing the MELAS fibroblast cell line it becomes possible to confirm the sensitivity of the Ion Torrent in being able to detect a known (A3243G) mutation that is characteristic of mitochondrial disease. On the other hand, sequencing the BJ cell line would reveal the baseline levels of mutations in a normal cell type that could be used as a reference sample for High-Resolution Melt (HRM) analysis, during comparisons to cancer DNA samples.

The MELAS mutation is classified pathogenic due to misincorporation of amino acids in mutant tRNA leucine structures that can impair proper protein translation<sup>203</sup>. Sequencing results positively detected the characteristic A3243G MELAS mutation at a level of 20.9% in the tRNA leucine 1 region (**Table 3.2**), which was an observation consistent with other studies<sup>204</sup>. The variant was confirmed at low levels using capillary sequencing (**Figure 3.5**), which is believed to have a minimum threshold level of 10% for mutation detection<sup>197</sup>. Since the additional variants listed in **Table 3.2** have not been reported to be associated with MELAS disease, it may be that the mutations have arisen as natural variation during cell culture.



**Figure 3.5.** Capillary sequencing of the fibroblast cell line containing the MELAS A3243G mutation.

**Table 3.2.** Ion Torrent sequencing on the human fibroblast cell line containing the MELAS A3243G mutation.

	Reference Position	Variant change	Percentage change in variant (%)	Gene region	Amino Acid Change	SNPs&GO	MutPred	
							Probability of deleterious mutation	Top 5 predicted features caused by amino acid mutation
Non-coding region	16186	C→T	5.9	D loop - Hypervariable segment 1, 7S DNA, membrane attachment site				
	1661	A→G	57.4	tRNA valine				
	1722	A→T	8.8	16s ribosomal RNA (rRNA)				
	2224	C→T	30.8					
	2228	A→T	25					
	2960	T→G	98.6					
	3243	A→G	20.9	tRNA leucine 1				
Coding region	9568	T→C	42.3	Cytochrome C Oxidase III: COXIII	I121T	Neutral RI 1, Uniprot P00414	0.749	<b>Gain of glycosylation at I121 (P = 0.0049)</b> <b>Gain of disorder (P = 0.0059)</b> <b>Loss of catalytic residue at P123 (P = 0.0396)</b> Loss of stability (P = 0.1428) Gain of phosphorylation at T119 (P = 0.1623)
	10550	A→G	3.3	NADH Dehydrogenase 4L: ND4L	Syn (M)			
	11467	A→G	30.4	NADH Dehydrogenase 4: ND4	Syn (L)			
	15264	C→T	7.1	Cytochrome B: CYT B	P173L	Disease, RI 4, Uniprot P00156	0.361	Loss of relative solvent accessibility (P = 0.0793) Loss of solvent accessibility (P = 0.089) Gain of methylation at R177 (P = 0.1226) Loss of glycosylation at S172 (P = 0.1763) Loss of disorder (P = 0.2084)
	15267	C→G	6.4		T174S	Disease, RI 3, Uniprot P00156	0.866	Gain of glycosylation at T174 (P = 0.0587) Gain of disorder (P = 0.0665) Loss of catalytic residue at T174 (P = 0.1513) Loss of methylation at R177 (P = 0.2045) Loss of sheet (P = 0.3635)
	15812	G→A	16.4		V356M	Neutral RI 6, Uniprot P00156	0.502	<b>Loss of catalytic residue at V356 (P = 0.0473)</b> Loss of stability (P = 0.4396) Loss of glycosylation at T361 (P = 0.4462) Loss of helix (P = 0.4763) Loss of sheet (P = 0.6423)

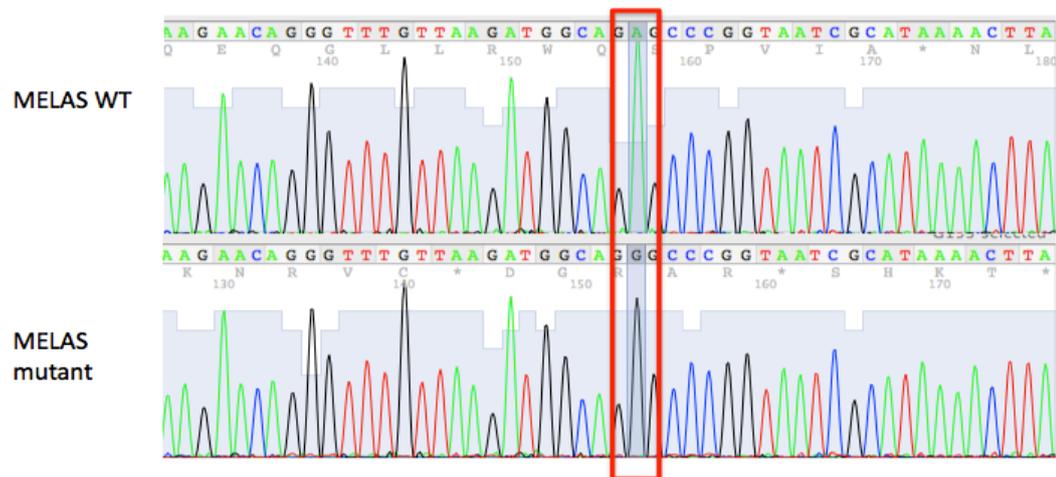
Sequencing of the BJ fibroblast mtDNA revealed the presence of only two mutations within the cytochrome B region (**Table 3.3**). These two variants appeared at low levels, and were also identified present in the other MELAS fibroblast cell line (**Table 3.3**). From this, it is possible to speculate that perhaps this region is especially susceptible to the development of mutations in the case of non-tumorigenic cells.

***Table 3.3.*** Ion Torrent sequencing of the BJ skin fibroblast cell line.

Reference Position	Reference	Variant	Percentage change in variant (%)	Gene region	Amino Acid Change	SNPs&GO	Mutpred	
							Probability of deleterious mutation	Top 5 predicted features caused by amino acid mutation
15264	C	C/T	9.8	Cytochrome B: CYTB	P 173 L	Disease, RI 4, Uniprot P00156	0.361	Loss of relative solvent accessibility (P = 0.0793) Loss of solvent accessibility (P = 0.089) Gain of methylation at R177 (P = 0.1226) Loss of glycosylation at S172 (P = 0.1763) Loss of disorder (P = 0.2084)
15267	C	C/G	9.6	Cytochrome B: CYTB	T 174 S	Disease, RI 3, Uniprot P00156	0.866	Gain of glycosylation at T174 (P = 0.0587) Gain of disorder (P = 0.0665) Loss of catalytic residue at T174 (P = 0.1513) Loss of methylation at R177 (P = 0.2045) Loss of sheet (P = 0.3635)

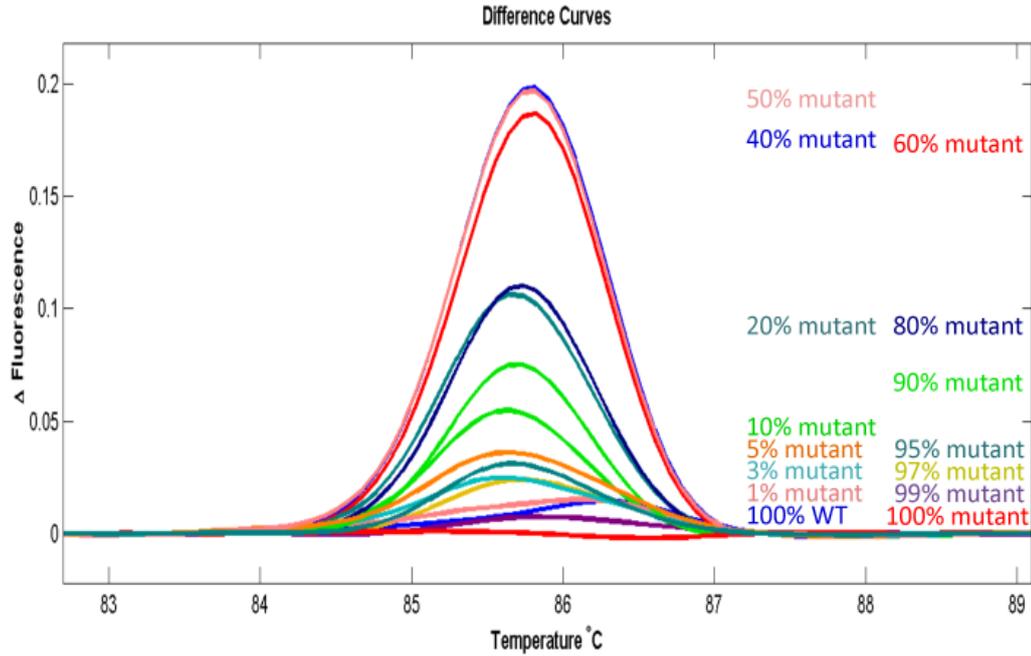
### 3.3.4. Determining the sensitivity of variant detection on the Ion Torrent sequencer and HRM system

The sensitivity limits on both the Ion Torrent and HRM system were tested to determine the lowest possible mutation level that could be accurately detected in both systems. To do this, a PCR amplicon of 223bp containing the MELAS mutation was cloned into the pCR2.1 bacterial cloning vector, as outlined in sections 2.15 and 2.15.1. Bacterial clones containing plasmids with the wild-type and mutant mtDNA sequence were identified and confirmed by capillary sequencing (**Figure 3.6**).



**Figure 3.6** Confirmation of MELAS wild-type and mutant inserts in recombinant plasmids.

Purified wild-type and mutant plasmids were mixed at specific ratios to generate dilutions of plasmid DNA at varying mutational loads. These wild-type/mutant plasmid dilutions were analysed by Ion Torrent and HRM.



**Figure 3.7.** HRM analysis of plasmid DNA mixed at specific mutational loads.

HRM analysis revealed that the change in fluorescence detection was proportional to the sequence difference between the two heteroduplex strands of the MELAS wild-type/mutant DNA (**Figure 3.7**). The 100% wild type and 100% mutant DNA did not result in a noticeable change in fluorescence levels unlike the 40%, 50% and 60% mutational loads, which would have required less energy input to separate the double-stranded heteroduplex DNA. It was possible to determine from this analysis that even mutations at the 1% level could be seen from the HRM results.

Similar tests on the Ion Torrent were performed. Results demonstrated little difference between the expected mutation load and actual levels of mutation detected (**Table 3.4**). From these results, the 3% mutation level was chosen as

the lowest threshold level, above which variants could be accurately identified without detection of background sequencing errors.

***Table 3.4. Determining the sensitivity levels on the Ion Torrent.***

Anticipated percentage mutation (%)	Percentage mutation detected on the Ion Torrent (%)	Detectable on the HRM
0	0.04	Yes
1	0.5	Yes
3	1	Yes
10	7.4	Yes
100	100	Yes

HRM is therefore more sensitive than Ion Torrent in the detection of mutations, which supports the use of HRM as a validation tool to confirm the outcomes identified from the Ion Torrent sequencer.

### **3.4. Methods**

Methods used are described in the following sections: 2.1.5, 2.1.9, 2.2 to 2.4.1, 2.6, 2.8.1 to 2.9.1, 2.10, and 2.16 to 2.22. Additional methods are provided below:

#### **3.4.1. HRM primers for the multiple myeloma mtDNA variants**

Primers were designed for each of the mtDNA variants detected from the Ion Torrent sequencing of multiple myeloma cells and tumors. The primers used for HRM all generated a product of between 100-250bp (**Table 3.5**), and all had an optimal annealing temperature of 55°C.

#### **3.4.2. Ethics statement**

Plasma cells were obtained from patients who had provided full written consent. All patients were diagnosed with multiple myeloma in the relapse and/or refractory stage. Patient consent and the clinical procedures for obtaining the cells were approved by The Alfred Health Human Ethics Committee (Melbourne, Australia).

**Table 3.5:** Primers used for multiple myeloma variant screening on the HRM

Variant Position (bp)	Forward primer (5'→3')	Reverse primer (5'→3')	Product size (bp)
557	CATCCTACCCAGCACACACA	TGATGTGAGCCCGTCTAAAC	134
659	TGTAGCTTACCTCCTCAAAG	TCGTGGTGATTTAGAGGGTG	160
976	CACACGATTAACCCAAGTCA	CAGTTTGGGTCTTAGCTATTG	168
1082	CTCCAGTTGACACAAAATAG	TCTGGCGAGCAGTTTTGTTG	161
1152	CACTATGCTTAGCCTAAACC	GAGGTTGATCGGGGTTTATC	151
1389	CAAGGTGTAGCCCATGAGGT	CTGTTCAACTAAGCACTCTAC	132
1589	CCTCAAGTATACTTCAAAGGAC	AGAGCGGTCAAGTTAAGTTG	177
1717	CTTAACTTGACCGCTCTGAG	TTCCCTTGCGGTACTATATC	144
2553	GAACTCGGCAAATCTTACCC	ATTATGCTACCTTGCACCGG	131
3208	GAGAAATAAGGCCTACTTCAC	GTTGGCCATGGGTATGTTGTT	184
3391, 3417	CTACTCCTCATTGTACCCAT	TATGGCGTCAGCGAAGGGTT	140
5432, 5438	CTACCTACGCCTAATCTACT	TAGGAGTAGCGTGGTAAGGG	135
5752	CAGCTAAGCACCCTAATCAAC	TCTAAAGACAGGGGTTAGGC	161
6352	CAGGTTGAACAGTCTACCCT	AAGAGGGGCGTTTGGTATTG	180
6774	CATTGGATACATAGGTATGGTC	AGTGTGGCGAGTCAGCTAAA	180
7718	CATGATCACGCCCTCATAAT	AGGATAGTTCAGACGGTTTC	133
8010, 8018	GGACTAATCTTCAACTCCTAC	TAAGCCTAATGTGGGGACAG	167
8252	CTCTGAAATCTGTGGAGCAAAC	CTAAGTTAGCTTTACAGTGGG	131
8759	CACCACCCAACAATGACTAA	AGATAGTTGGGTGGTTGGTG	168
8821	TTGCCACAACCTCCTC	GTGCCTTGTTGTAAGAAGTG	162
9111	CTCATGCACCTAATTGGAAG	TCGTGCAGGTAGAGGCTTA	163
9216, 9226	CCAAGCCTACGTTTTCACAC	AGCGTTATGGAGTGGAAAGTG	179
10657	TCGCTGTTTATTATAGCTACTCTC	GTACGTAGTCTAGGCCATATG	156
11220	CTTGGCTATCATCACCCGAT	CCTAGGGTGTGTGAGTGTA	128
11445	TACCTCTTACGGACTCCAC	AGGCTATGTGTTTTGTCAGG	155
11674, 11675	CTGCCTACGACAAACAGAC	GGCAGAATAGTAATGAGGAT G	160
12847	CCTTCTTGCTCATCAGTTGA	GTTGTGGGTCTCATGAGTTG	146
14772	CAACCACGACCAATGATATG	GAGTGAGCCGAAGTTTCATC	163
15264, 15267	GAGGCCAAATATCATTCTGAG	GTTGTTTGATCCCGTTTCGTG	218
16186	CCACCATGAATATTGTACGGTA	TGGCTTTGGAGTTCAGTTG	147

### **3.5. Results**

#### **3.5.1. Detection of mtDNA variants in multiple myeloma tumors**

Five independent tumor samples were analysed by next generation sequencing. A 3% threshold criteria was applied for the detection of mutations, and from this, 16 heteroplasmic variants were identified (**Table 3.6**). Half of these identified variants have not previously been reported before in MitoMap (557bp, 659bp, 1082bp, 1152bp, 2553bp, 12847bp, 14772bp and 15264bp). 8 variants were present in each of the non-coding and coding regions of mtDNA. All variants identified, except for those at positions 15264bp and 15267bp, were unique to each of the tumor samples. Where there were similarities in the location of these variants, there were also differences in their levels of mutation. Levels ranged from 3.1% (D loop; 16186bp) to 43.9% (12s rRNA; 1152bp) in the non-coding region, and for the coding region, values were between 3.1% (CYT B 14772bp) to 28.7% (ND5 12847bp). Variants within the coding regions of mtDNA induced both synonymous and non-synonymous amino acid changes in the corresponding protein sequence. CYT B nucleotide positions at 15264bp and 15267bp appeared as a potential hotspot region for variants. At least 3 of the 5 tumor samples analysed possessed a variant at a frequency above 3% at each of these positions.

#### **3.5.2. Verification of variants and screening of additional tumors using HRM**

The same tumor samples from Table 3.6 underwent validation using the HRM, which confirmed the presence of these variants (**Table 3.7**). For these 5 tumor samples, HRM analysis also revealed the presence of additional variants. These are likely to represent variants between levels of 1% and 3%, which were

**Table 3.6.** MtDNA variants detected from multiple myeloma tumors. Results were filtered to exclude presence of duplicate reads. All variants listed represent those detected above the 3% mutation threshold level.

Reference Position	Variant change	Percentage change in variant (%)					Gene region	Amino Acid Change	SNPs&GO	Mutpred	
		S228	S219	S140	S139	S241				Probability of deleterious mutation	Top 5 predicted features caused by amino acid mutation #
16186	C→T					3.1	D-Loop - Hypervariable segment 1, 7S DNA, membrane attachment site				
557*	C→T					3.3	D-Loop - Major H-strand promoter				
659*	T→C		20.4				12s rRNA				
1082*	A→G					36.5	12s rRNA				
1152*	A→G		43.9				12s rRNA				
1389	G→A					8	12s rRNA				
1717	T→C					5.3	16s rRNA				
2553*	G→A	25.2					16s rRNA				
5432	A→G					4.7	NADH Dehydrogenase: ND2	Syn (K)	-	-	
5438	C→T			5.4			NADH Dehydrogenase: ND2	Syn (T)	-	-	
7718	A→C		3.9				Cytochrome C Oxidase: COX II	T 45 P	Disease, RI 0, Uniprot P00403	0.668	<b>Loss of stability (P = 0.0296)</b> <b>Gain of glycosylation at T45 (P = 0.0341)</b> <b>Gain of helix (P = 0.0496)</b>
9111	T→C		3.6				ATP Synthase 6: ATP6	Syn (I)	-	-	
12847*	G→A	28.7					NADH Dehydrogenase: ND5	A 171 T	Neutral, RI 1, Uniprot P03915	0.713	
14772*	C→T					3.1	Cytochrome B: CYT B	P 9 L	Disease, RI 1, Uniprot P00156	0.643	<b>Loss of disorder (P = 0.0255)</b> <b>Loss of glycosylation at P9 (P = 0.0326)</b> <b>Loss of methylation at K12 (P = 0.0427)</b>
15264*	C→T	5.2	6.3			5.4	Cytochrome B: CYT B	P 173 L	Disease, RI 4, Uniprot P00156	0.361	
15267	C→G	7.3	8.7	3.5		6.8	Cytochrome B: CYT B	T 174 S	Disease, RI 3, Uniprot P00156	0.866	

\* Novel variants not reported in MitoMap # Only predictions of P<0.05 are shown.

**Table 3.7.** Screening additional multiple myeloma tumors using HRM.

Gene region		Non-coding region							Coding region								
		D-loop		12s rRNA			16s rRNA		ND2		COX II	ATP 6	ND5	CYT B			
Variant change		C→T	C→T	T→C	A→G	A→G	G→A	T→C	G→A	A→G	C→T	A→C	T→C	G→A	C→T	C→T	C→G
Reference position		16186	557	659	1082	1152	1389	1717	2553	5432	5438	7718	9111	12847	14772	15264	15267
tumor samples (n = 25)	S238																
	S171																
	S162																
	S156																
	S237																
	S235																
	S341																
	S254																
	S184																
	S335																
	S325																
	S228	####	####					####	####				####	####		####	####
	S390																
	S327																
	S298																
	S334																
	S252																
	S354																
	S219											####	####	####		####	####
	S140			####	####		####	####		####		####					####
S139		####	####	####		####	####				####		####	####			
S241	####						####		####		####				####	####	
S307																	
S226																	
S216																	
Total		94.44	61.90	69.57	82.61	100.00	72.22	90.91	53.33	96.00	80.00	96.00	84.00	95.65	96.00		
(%)		5.56	38.10	30.43	17.39	-	27.87	9.09	46.67	4.00	20.00	4.00	16.00	4.35	4.00		
Significance		**			*	**		**		**	*	**	*	**	**		

Key:	<span style="background-color: green; border: 1px solid black; display: inline-block; width: 15px; height: 10px;"></span> Variant present	<span style="background-color: lightgreen; border: 1px solid black; display: inline-block; width: 15px; height: 10px;"></span> Variant present - analysed jointly (either 1 present or both)	*	Variant present in >80% of samples
	<span style="background-color: red; border: 1px solid black; display: inline-block; width: 15px; height: 10px;"></span> Variant absent	<span style="background-color: pink; border: 1px solid black; display: inline-block; width: 15px; height: 10px;"></span> Variant absent - analysed jointly (either 1 absent or both)	**	Variant present in >90% of samples
	<span style="background-color: black; border: 1px solid black; display: inline-block; width: 15px; height: 10px;"></span> Sample not available	####	Variant presence/absence confirmed with both Ion Torrent and HRM	

determined as the threshold levels of variant detection for the HRM and Ion Torrent, respectively. In addition to the variants reported from **Table 3.6**, a further 20 tumor samples were screened with HRM to identify the patterns of variants across a range of tumors representative of multiple myeloma (**Table 3.7**). Based on the available material, seven tumor samples were screened to be positive for all variants (S237, S341, S254, S325, S327, S354, S226).

Generally, more variants were detected in the non-coding than in the coding region. The nucleotide position most variable in its presence or absence of mutation was at 557bp, within the D-loop major H-strand promoter region. For the collection of samples screened, the total proportion of samples containing the presence of the screened variant was calculated (**Table 3.7**). Values were adjusted based on the actual number of samples screened, since not all tumor samples could be analysed across the complete range of variants, due to limited availability of material.

Over 90% of the tumor samples shared the presence of particular variants located at positions 1152bp (12s rRNA), 1717bp (16s rRNA), 5432/5438bp (*ND2*), 9111bp (*ATP6*), 14772bp (*CYT B*), 15264/15267bp (*CYT B*) and 16186bp (D-loop). The D-loop variant at position 16186bp resided within the hypervariable segment 1 region of mtDNA, which is associated with natural sequence variation in diverse populations. The variant at position 2553bp within 16s rRNA displayed the most variability, with only 8/15 samples found to have the presence of this variant. This could be attributed to the reduced number of samples available for examining this particular variant during the screening procedure.

### 3.5.3. Susceptibility of mtDNA regions to the development of mutations

The sizes of mtDNA gene regions can impact on the number of variants detected within these regions. As such, by normalizing the number of variants to the size of the affected mtDNA regions, it is possible to determine the susceptibility of different mtDNA genes to the development of mutations.

The 12s rRNA region had the greatest number of variants above the 3% threshold criteria, compared to any other region investigated, based on the calculated susceptibility score of  $4.20 \times 10^{-3}$  (**Table 3.8**). 12s rRNA was discovered to be 7.60x more likely to acquire a mutation than the *ND5* region (4 variants occurring across the 954bp region of the 12s rRNA gene, compared to 1 variant in the 1812bp region of *ND5*) (**Table 3.9**).

*CYT B* was the second-highest scoring region that was susceptible to nucleotide change relative to the other mtDNA regions. *CYT B* was 4.77 fold more susceptible compared to *ND5* to developing a variant above the 3% threshold level of mutation, which is also the second highest factor difference, below that of the 7.60 fold difference between the 12s rRNA and *ND5* regions (**Table 3.9**). *COX II* and *ATP6* were two regions that had an equal chance of undergoing a mutational change, which perhaps reflected the close proximity of these two genes in the mitochondrial genome.

**Table 3.8.** Susceptibility of mtDNA regions to the development of mutations.

Reference Position	Gene region	Nucleotide positions for gene regions (bp)	Size of gene region (bp)	Number of variants detected in gene region	Normalised data		
					Mutation rate in each gene region	Mutation rate grouped for coding/non-coding regions	
557, 16186	Non-coding region	D-Loop	1..576, 16024..16569	1122	2	0.00178	0.00128
659							
1082		12s rRNA	648..1601	954	4	0.00420	
1152							
1389							
1717		16s rRNA	1671..3229	1558	2	0.00128	
2553							
5432	Coding region	NADH Dehydrogenase: ND2	4470..5511	1042	2	0.00192	0.000775
5438							
7718		Cytochrome C Oxidase: COXII	7586..8269	684	1	0.00146	
9111							
12847		ATP Synthase 6: ATP6	8527..9207	681	1	0.00147	
14772							
15264		NADH Dehydrogenase: ND5	12337..14148	1812	1	0.000552	
15267							
15267	Cytochrome B: CYTB	14747..15887	1141	3	0.00263		

**Table 3.9.** Fold change differences on the susceptibility of mtDNA regions to mutation.

Gene region	D loop	12s rRNA	16s rRNA	ND2	COX II	ATP6	ND5	CYT B
D loop		0.43	1.39	0.93	1.22	1.21	3.23	0.68
12s rRNA	2.35		3.27	2.18	2.87	2.85	7.60	1.59
16s rRNA	0.72	0.31		0.67	0.88	0.87	2.32	0.49
ND2	1.08	0.46	1.50		1.31	1.31	3.48	0.73
COX II	0.82	0.35	1.14	0.76		1.00	2.65	0.56
ATP6	0.82	0.35	1.15	0.77	1.00		2.66	0.56
ND5	0.31	0.13	0.43	0.29	0.38	0.38		0.21
CYT B	1.48	0.63	2.05	1.37	1.80	1.79	4.77	

**Key**

if value >1 susceptibility of the region to variant changes is greater for the gene listed on the left  
if value <1 susceptibility of the region to variant changes is less for the gene listed on the left

**3.5.4. Identification of variants during mtDNA replenishment in U266 cells**

Sequencing of the U266 cells (Table 3.10) identified variants that were different to the tumor sample variants (Table 3.6). This confirmed that myeloma plasma cells cultured *in vitro* exhibit different biological properties than those grown *in vivo*, most likely due to myeloma cells in the body being able to interact with other stromal cells in the bone marrow microenvironment<sup>205</sup>. 4 novel variants were identified that had not previously been reported in MitoMap (9529bp, 13058bp, 13985bp and 14159bp). Since there were no common variants detected between the tumors and the U266 cells, it was of interest to determine whether U266 variants remain the same or differ after a period of 7 and 14 days of mtDNA depletion, followed by recovery. This was examined using the Ion Torrent sequencer.

Variants at 14155bp and 14160bp within the *ND6* region displayed random change to their levels of mutation (Table 3.10), which would argue against the

**Table 3.10.** Identification of variants during mtDNA replenishment in U266 cells. Cells were depleted of mtDNA for 7 or 14 days followed directly by recovery for 3 (R3), 7 (R7) or 14 days (R14).

	Reference Position	Variant	Percentage change in variant (%)							Gene region	Amino acid change	SNPs&GO	Mutpred	
			U266 control	U266 day 7 R3	U266 day 7 R7	U266 day 7 R14	U266 day 14 R3	U266 day 14 R7	U266 day 14 R14				Probability of deleterious mutation	Top 5 predicted features caused by amino acid mutation <sup>#</sup>
Non-coding	16519	T→C	51.7	55.0	53.5	52.4	58.1	54.2	52.6	D loop - Membrane attachment site, 7S DNA				
	568	C→T	5.3							D loop - Hypervariable segment 3				
	1243	T→C	40.5	45.9	42.3	43.0	43.6	39.5	40.2	12s ribosomal RNA				
Coding region	8994	G→A	53.6	64.4	66.1	58.7	67.6	69.7	54.5	ATPase 6	Syn (L)	-		
	9529*	C→T				4.8				Cytochrome c oxidase III	P108L	Disease, RI 2, Uniprot P00414	0.676	<b>Loss of disorder (P = 0.01)</b> <b>Loss of glycosylation at P108 (P = 0.05)</b>
	12705	C→T	80.6	78.1	76.6	76.0	79.4	78.0	79.4	NADH Dehydrogenase 5	Syn (I)			
	13058*	C→A							3.1		T241N	Neutral, RI 5, Uniprot P03915	0.684	<b>Loss of glycosylation at T241 (P = 0.02)</b>
	13981	C→G							3.5		P549A	Neutral, RI 9, Uniprot P03915	0.447	<b>Gain of helix (P = 0.03)</b>
	13984	C→T							3.5		Syn (L)	-		
	13985*	T→C							3.5		L550P	Neutral, RI 1, Uniprot P03915	0.457	<b>Gain of loop (P = 0.001)</b> <b>Loss of helix (P = 0.004)</b> <b>Gain of catalytic residue at P549 (P = 0.02)</b> <b>Gain of glycosylation at L550 (P = 0.04)</b> <b>Gain of relative solvent accessibility (P = 0.05)</b>
	14155	C→T	3.7		4.3	4.9			3.1	NADH Dehydrogenase 6	Syn (G)	-		
	14159*	C→G			5.4						R172P	Disease, RI 3, Uniprot P03923	0.423	<b>Loss of methylation at R172 (P = 0.03)</b>
	14160	G→C	9.0	3.7	5.5	4.1		5.1			R172G	Neutral, RI 2, Uniprot P03923	0.442	<b>Loss of methylation at R172 (P = 0.03)</b>

\* Novel variants not reported in MitoMap<sup>#</sup> Only predictions of P<0.05 are shown.

presence of mechanisms that would either favor or selectively eliminate certain mtDNA variants during mtDNA depletion. Four out of the 13 variants detected were shared between all samples, which included two in the non-coding region of mtDNA (16519bp within D-loop and 1243bp within 12S rRNA) and two in the coding region (8994bp within *ATP6* and 12705bp within *ND5*).

For the D-loop variant at position 16519bp, prolonged depletion increased its percentage frequency relative to day 0 cells. The 7-day mtDNA depleted cells recovered for 3 days had a lower level of mutation (55%) than U266 cells that had been depleted for 14 days, and allowed to recover for the same length of time (58.1%). Comparing these different depletion/replenishment stages, prolonged depletion increased the levels of mutation for some the mtDNA variants, such as for the 12705bp C→T variant within the *ND5* region. However, this is not always the case, as the variant at position 1243bp within 7-day depleted cells recovered for 3 days had higher levels of the variant than the 14-day depleted cells also recovered for 3 days. Decreasing levels of mutation for the 1243bp variant was evident across the other depletion/replenishment stages. Noticeably, the variant at 8994bp had variable changes during the course of depletion and replenishment, but was the only variant that displayed the greatest changes to its levels of mutation. Changes to the other variants were only marginal in comparison.

Loss of the 568bp C→T variant in U266 cells during mtDNA depletion emphasized that apart from changes in the levels of mutation during the progressive decline of mtDNA copy number, it is possible for some variants to

be lost entirely. Although this could occur via a selection mechanism, whereby the mtDNA copies that do not possess this variant could be preferentially replicated to decrease the overall mutation levels of the variant within the mitochondrial population. On the contrary, replenishment of mtDNA copy number for 14 days after 7 days of depletion, revealed the acquisition of an additional low level variant (4.8%) at position 9529bp, not previously reported for the control cells. As such, there are no selective forces to either favour or eliminate certain mtDNA variants during mtDNA depletion.

Overall, the majority of variants were discovered in the 14-day depleted and 14-day recovered cells, which acquired low levels of mutation during the course of extended mtDNA depletion.

### **3.5.5. Susceptibility of depleted and recovered U266 cells to mutations**

*ND6* was identified as the region most susceptible to mutation within the coding region (**Table 3.11**). For the non-coding regions, this was the D-loop, calculated to have a fold change 1.70 times greater than the 12s rRNA region (**Table 3.12**). Comparing *ND6* to the D-loop region, *ND6* was found to have a 3.21 fold greater probability of developing a mutation. Likewise, this fold change difference was observed to be greatest between *ND6* and the 12s rRNA region, which reported a 5.46 fold difference.

**Table 3.11.** Susceptibility of depleted and recovered U266 cells to mutation.

Reference Position	Gene region		Nucleotide positions for gene regions (bp)	Size of gene region (bp)	Number of variants detected in gene region	Normalised data	
						Mutation rate in each gene region	Mutation rate grouped for coding/non-coding regions
16519...568	Non-coding	D loop	1..576, 16024..16569	1121	2	0.00178	0.000481
1243		12s rRNA	648..1601	953	1	0.00105	
8994	Coding region	ATPase 6	8527..9207	680	1	0.00147	0.000968
9529		Cytochrome C Oxidase III	9207..9990	783	1	0.00128	
12705, 13058, 13981, 13984, 13985		NADH dehydrogenase 5	12337..14148	1811	5	0.00276	
14155, 14159, 14160		NADH dehydrogenase 6	14149..14673	524	3	0.00573	

**Table 3.12.** *Fold change differences for the susceptibility of depleted and recovered U266 cells to the development of mutations.*

Gene region	D-loop	12s rRNA	ATP 6	COX III	ND5	ND6
D-loop		1.70	1.21	1.40	0.65	0.31
12s rRNA	0.59		0.71	0.82	0.38	0.18
ATP 6	0.82	1.40		1.15	0.53	0.26
COX III	0.72	1.22	0.87		0.46	0.22
ND5	1.55	2.63	1.88	2.16		0.48
ND6	3.21	5.46	3.89	4.48	2.07	

**Key**

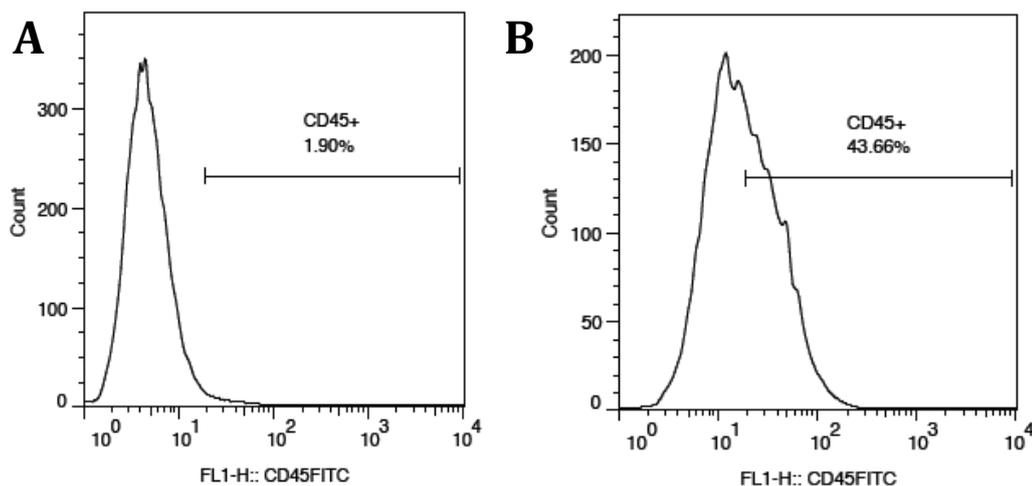
if value >1 susceptibility of the region to variant changes is greater for the gene listed on the left  
if value <1 susceptibility of the region to variant changes is less for the gene listed on the left

**3.5.6. MtDNA variants in U266 sub-populations of CD45+ and CD45- cells**

U266 cells are heterogenous for the expression of CD45, a marker of malignancy. FACS sorted CD45+ and CD45- U266 cells were sequenced by Ion Torrent to determine whether variants found in these sub-populations were similar or different to those in the total population of U266 cells. From these results, it would be possible to determine whether the two CD45 sub-populations harbour characteristic mutations that could distinguish each based on sequence alone. Sequence variations could also be responsible for differences between the biological properties and cellular phenotypes of the two sub-populations<sup>206</sup>. The hypothesis was that the CD45- population would harbour more detrimental variants, as these cells have been reported as important indicators of tumor malignancy in multiple myeloma<sup>207</sup>.

It was estimated that 43.66% of U266 cells after FACS were positive for the CD45 marker and 56.34% were negative (**Figure 3.8**). Whilst it is known that the proportion of CD45+/- cells undergo constant conversion into one another

dependent on the exposure of stimulants to the U266 cells<sup>208</sup>, the majority of the heterogeneous U266 cell population was identified negative for CD45, which supported the observation of others<sup>208</sup>.



**Figure 3.8.** Fluorescent-activated cell sorting (FACS) analysis on U266 cells. U266 cells are known to be heterogenous for expression of the surface marker CD45. (A) Before enrichment for CD45 (B) After enrichment for CD45, 43.66% of cells were detected positive of the surface marker.

Sequencing identified 23 variants, 13 of which had not been previously identified in MitoMap (976bp, 6352bp, 6774bp, 8010bp, 8018bp, 8759bp, 8928bp, 8934bp, 9216bp, 9226bp, 11220bp, 11445bp and 11675bp). 20 variants were located within the coding regions of mtDNA (**Table 3.13**). The 3 variants within the non-coding regions resided in the 12s rRNA (976bp), 16s rRNA (3208bp) and 5752bp (origin of light strand replication). The mutations levels ranged from 12.3% to 53.7% (both in 12s rRNA) (**Table 3.13**). For the coding-region variants, 14 non-synonymous amino acids mutations were identified in total, 9 of

**Table 3.13.** MtDNA variants in U266 sub-populations of CD45+ and CD45- cells.

	Reference Position	Variant	Percentage change in variant (%)		Gene region	Amino Acid Change	SNPs&GO	MutPred	
			CD45+	CD45-				Probability of deleterious mutation	Top 5 predicted features caused by amino acid mutation <sup>#</sup>
Non-coding	976*	A→C	53.7	12.3	12s rRNA				
	3208	C→A	33.3		16s rRNA				
	5752	A→G	30		Orign of L-strand replication				
Coding region	3391	G→A	37.5	21.4	NADH Dehydrogenase: ND1, Uniprot P03886	G 29 S	Neutral, RI 0	0.895	
	3417	C→A	29			Syn (P)	-		
	6352*	T→G	23.6		Cytochrome C Oxidase: COX I, Uniprot P00395	L 150 W	Disease, RI 2	0.797	Loss of stability (P = 0.0384) Gain of catalytic residue at I147 (P = 0.257) Loss of helix (P = 0.4763)
	6774*	C→G	22.7			H 291 D	Disease, RI 8	0.844	
	8010*	T→G		33	Cytochrome C Oxidase: COX II, Uniprot P00403	V 142 G	Disease, RI 1	0.531	Loss of stability (P = 0.002)
	8018*	C→T		28.6		P 145 S	Disease, RI 8	0.871	
	8252	C→A	27.3			P 223 T	Neutral, RI 9	0.327	
	8759*	T→G		33.3	ATP synthase 6: ATP6, Uniprot P00846	F 78 C	Disease, RI 4	0.837	
	8821	T→C	20.2	12.7		S 99 P	Disease, RI 6	0.819	Gain of glycosylation at T96 (P = 0.0461)
	8927	C→A	34.2	9.7		P 134 H	Disease, RI 6	0.749	Loss of glycosylation at P134 (P = 0.029)
	8928*	T→C	34.2	10		Syn (P)	-		
	8934*	C→G	22.9	6.4		Syn (P)	-		
	8939	T→A	26.5	9.3	I 138 N	Disease, RI 3	0.793	Gain of disorder (P = 0.0401)	
	9216*	C→A		28.6	Cytochrome C Oxidase: COX III, Uniprot P00414	Q 4 K	Neutral, RI 8	0.493	Gain of methylation at Q4 (P = 0.0048)
	9226*	C→G	44	56.5		A 7 G	Neutral, RI 8	0.529	
	10657	T→G	33.3		NADH Dehydrogenase: ND4L	M 63 X(stop)	Uniprot P03901		
	11220*	T→G		25.6	NADH Dehydrogenase: ND4, Uniprot P03905	L 154 R	Disease, RI 3	0.738	Gain of methylation at L154 (P = 0.0051) Loss of stability (P = 0.0279)
11445*	T→G	23.5	4.9	M 229 X(stop)					
11674	C→T		55.6	Syn (T)		-			
11675*	C→T	21.7		P 306 S		Neutral, RI 6	0.579		

\* Novel variants not reported in MitoMap<sup>#</sup> Only predictions of P<0.05 are shown.

which were predicted to have the potential to impair protein function, with the remaining 5 predicted to induce neutral effects reported at mtDNA sites 3391bp (*ND1*), 8252bp (*COX II*), 11675bp (*ND4*), 9216bp and 9226bp (*COX III*).

The *COX II* variant at 8018bp, predicted to cause protein dysfunction, occurred at a frequency almost three times higher in the CD45<sup>-</sup> cells than in its positive counterpart (**Table 3.13**). However, a selection of variants including the 8821bp mutation within the *ATP6* region was detected at two-fold higher levels in the CD45<sup>+</sup> than in the CD45<sup>-</sup> cells. Likewise, the 8927bp mutation also within *ATP6*, presented with approximately a four-fold increase in the CD45<sup>+</sup> cells than in the CD45<sup>-</sup> cells. The variant at 8939bp exemplifies this further containing a percentage mutation at three-fold higher levels in the CD45<sup>+</sup> compared to the CD45<sup>-</sup> cells, even though CD45<sup>-</sup> cells are responsible for maintaining the tumorigenic characteristics of multiple myeloma.

Overall, 78.26% of variants were common to both sub-populations, with a few extra variants within the CD45<sup>-</sup> cells. At several sites, variants appeared exclusive for either the CD45<sup>+</sup> or CD45<sup>-</sup> cells (**Table 3.13**). Those present only in the CD45<sup>+</sup> sub-population included variants at sites 3208bp, 3417bp, 5752bp, 6352bp, 6774bp, 8252bp, 10657bp and 11675bp. Those exclusive to the CD45<sup>-</sup> cells were found at positions 8010bp, 8018bp, 8759bp, 9216bp, 11220bp and 11674bp. The variant located at position 8821bp within *ATP6*, was the only variant sharing less than a 10% difference in variant level between the CD45<sup>+</sup> and CD45<sup>-</sup> samples.

Data retrieved from the Ion Torrent were subsequently validated using the HRM, which corresponded with the variants identified from next generation sequencing data (**Table 3.14**). Additional variants were also detected, which were likely to have mutation levels between the threshold levels of variant detection for the HRM (1%) and Ion Torrent (3%).

**Table 3.14.** Confirmation of CD45+ and CD45- variants using HRM analysis.

	Reference Position	Variant	Cells		Gene region
			CD45+	CD45-	
Non-coding region	976	A→C	####	####	12s rRNA
	3208	C→A	####	####	16s rRNA
	5752	A→G	####	####	Origin of light strand replication
Coding region	3391	G→A	####	####	NADH Dehydrogenase: ND1
	3417	C→A	####	####	
	6352	T→G	####	####	Cytochrome C Oxidase: COX I
	6774	C→G	####	####	
	8010	T→G	####	####	Cytochrome C Oxidase: COX II
	8018	C→T	####	####	
	8252	C→A	####	####	ATP synthase 6: ATP6
	8759	T→G	####	####	
	8821	T→C	####	####	
	8927	C→A	####	####	
	8928	T→C	####	####	
	8934	C→G	####	####	Cytochrome C Oxidase: COX III
	8939	T→A	####	####	
	9216	C→A	####	####	NADH Dehydrogenase: ND4L
	9226	C→G	####	####	
	10657	T→G	####	####	NADH Dehydrogenase: ND4
	11220	T→G	####	####	
11445	T→G	####	####		
11674	C→T	####	####		
	11675	C→T	####	####	

<b>Key:</b>		Variant present
		Variant absent
		Variant present - analysed jointly (either 1 present or both)
		Variant absent - analysed jointly (either 1 absent or both)
	####	Variant presence/absence confirmed in both Ion Torrent and HRM

### 3.5.7. Susceptibility of CD45+ and CD45- cells to mutations

After normalising the number of variants to the sizes of the gene regions for the CD45 cells, the origin of light strand replication was most susceptible to the development of a mutation (1 variant present within a 77bp region of mtDNA) (Table 3.15). However, since the susceptibility score for the origin of light strand replication may be over-estimated due to the presence of a single variant within a very short (77bp) region of mtDNA, it would be necessary to consider the importance of the second-most susceptible gene to the development of mutations.

**Table 3.15.** Determining the susceptibility of CD45+ and CD45- cells to mutations.

Reference Position (bp)	Gene region	Nucleotide positions for gene regions (bp)	Size of gene region (bp)	Number of variants detected	Normalised data			
					Mutation rate in each gene region	Mutation rate grouped for coding/non-coding regions		
976	Non-coding	12s rRNA	648..1601	954	1	0.00105	0.000481	
3208		16s rRNA	1671..3229	1558	1	0.000642		
5752		Origin of L-strand replication	5721..5798	77	1	0.0130		
3391	Coding region	NADH Dehydrogenase: ND1	3307..4262	955	2	0.00209	0.00194	
3417			5904..7445	1541	2	0.00130		
6352		Cytochrome C Oxidase: COX I	7586..8269	684	3	0.00439		
6774			8527..9207	681	6	0.00882		
8010		ATP synthase 6: ATP6	9207..9990	783	2	0.00255		
8018			10470..10766	296	1	0.00338		
8252		Cytochrome C Oxidase: COX III	10760..12137	1377	4	0.00290		
8759								
8821								
8927								
8928								
8934								
8939								
9216								
9226								
10657								
11220								
11445								
11674								
11675								

The region that was second most likely to acquire variants was *ATP6*, calculated to have a variant susceptibility value of  $8.82 \times 10^{-3}$ . 16s rRNA was the least susceptible to the development of mutations in the CD45 cells (susceptibility value:  $6.42 \times 10^{-5}$ ). Comparatively, the origin of light strand replication was 1.47 fold greater than the *ATP6* region to the development of a mutation. However, the greatest fold change difference occurred between the origin of light strand replication and the 16s rRNA, which was calculated to be 20.23 (**Table 3.16**).

**Table 3.16.** Fold change differences for the of CD45+ and CD45- cells to mutation.

Gene region	12s rRNA	16s rRNA	OriL	ND1	COX I	COX II	ATP6	COX III	ND4L	ND4
12s rRNA		1.63	0.08	0.50	0.81	0.24	0.12	0.41	0.31	0.36
16s rRNA	0.61		0.05	0.31	0.49	0.15	0.07	0.25	0.19	0.22
OriL	12.38	20.23		6.20	10.01	2.96	1.47	5.08	3.84	4.47
ND1	2.00	3.26	0.16		1.61	0.48	0.24	0.82	0.62	0.72
COX I	1.24	2.02	0.10	0.62		0.30	0.15	0.51	0.38	0.45
COX II	4.19	6.84	0.34	2.10	3.38		0.50	1.72	1.30	1.51
ATP6	8.41	13.75	0.68	4.21	6.80	2.01		3.45	2.61	3.04
COX III	2.43	3.98	0.20	1.22	1.97	0.58	0.29		0.76	0.88
ND4L	3.22	5.26	0.26	1.61	2.60	0.77	0.38	1.32		1.16
ND4	2.77	4.53	0.22	1.39	2.24	0.66	0.33	1.14	0.86	

**Key**

if value >1 susceptibility of the region to variant changes is greater for the gene listed on the left  
if value <1 susceptibility of the region to variant changes is less for the gene listed on the left

**3.5.8. The influence of GC content on the development of mutations**

To determine whether the GC content of an mtDNA region influences its susceptibility to the development of mutations, each mtDNA region from the sequencing outcomes of the multiple myeloma tumors (**Table 3.6**), was

examined. For few variants, a correlation existed between the percentage GC and probability of mutation acquisition (**Table 3.17**).

**Table 3.17.** Calculating the GC content of each mtDNA variant region.

Gene region		Nucleotide positions for gene regions (bp)	Size of gene region (bp)	Nucleotide count				% G/C content
				Adenine (A)	Thymine (T)	Guanine (G)	Cytosine (C)	
Non-coding	D-Loop	1..576, 16024..16569	1122	339	260	152	371	46.6
	12s rRNA	648..1601	954	315	205	181	253	45.5
	16s rRNA	1671..3229	1558	547	344	270	397	42.8
Coding region	NADH Dehydrogenase: ND2	4470..5511	1042	326	268	99	349	43
	Cytochrome C Oxidase: COXII	7586..8269	684	196	172	102	214	46.2
	ATP Synthase 6: ATP6	8527..9207	681	206	174	71	230	44.2
	NADH Dehydrogenase: ND5	12337..14148	1812	551	447	192	622	44.9
	Cytochrome B: CYTB	14747..15887	1141	326	287	137	291	46.3

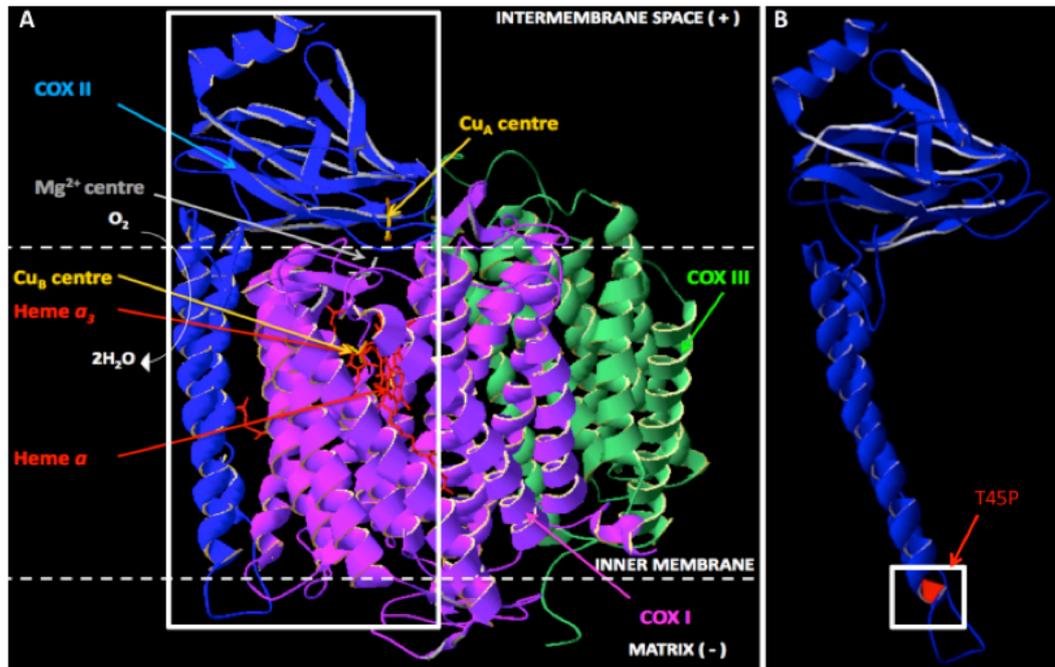
*CYT B* was calculated to have a probability score of  $2.632 \times 10^{-3}$ , which was the second-highest most susceptible region for mutation. It was also ranked second in terms of its GC content (46.3%), below the D-loop (46.6%). Similarly, the 16s rRNA region was ranked 7<sup>th</sup> of all genes analysed for its susceptibility to mutation ( $1.284 \times 10^{-3}$ ), wherein it also possessed the lowest GC content amongst all mtDNA regions assessed (42.8% GC). Finally, *ATP6* was determined the 5<sup>th</sup> most susceptible region for mutation based on normalized data ( $1.4706 \times 10^{-3}$ ), which almost corresponded with it possessing the 6<sup>th</sup> most GC rich region of all others examined (44.2%).

From **Table 3.6**, 50% of mutational events involved a pyrimidine-pyrimidine substitution, closely followed by purine-purine (37.5%) and lastly by purine-pyrimidine transitions (12.5%). However, Table 3.9 did not appear to follow suit, as the majority (69.6%) of variants involved a purine (A or G) and a pyrimidine (T or C) substitution. In 26.1% of cases these variants were pyrimidine-pyrimidine in nature, very rarely (8.7%) did these involve purine-purine substitutions. This suggests that there was no bias towards the variant introduced in place of a natural nucleotide base in mtDNA, thereby giving all residues an equal (33.3%) chance of being substituted during a mutational event.

### **3.5.9. *In silico* and protein structure analysis**

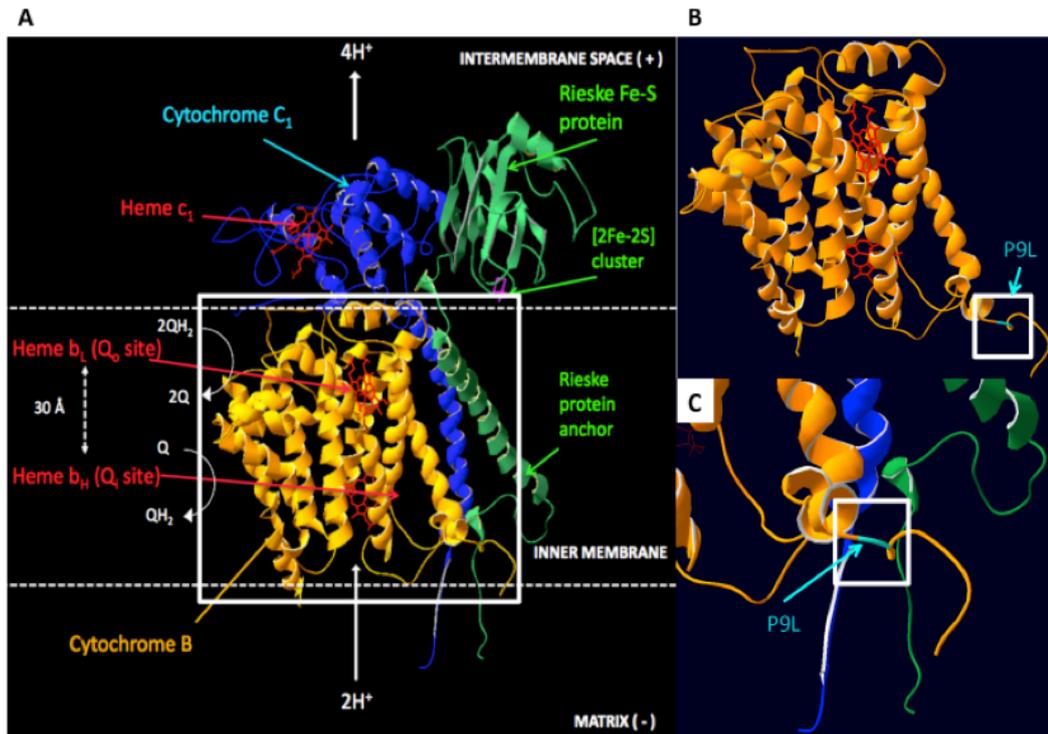
#### **3.5.9.1. Multiple myeloma tumor sample variants**

The amino acid substitution caused by the variant at position 7718bp induced a change from threonine to proline (T45P) in the COX II region, which was predicted to impair protein function (**Table 3.6**). MutPred predicted a high probability ( $P < 0.05$ ) for the T45P mutation to disrupt the structure of COX II, in addition to the potential gains of glycosylation and local gain of helix structure (**Table 3.6**). The mutation resided on the end of a transmembrane-spanning helix (**Figure 3.9**) on the bovine COX II model, which shared 72% sequence identity with the human COX II structure. Although modeling was performed using the bovine protein structure, results from SNPs&GO and MutPred were based on its human counterpart.



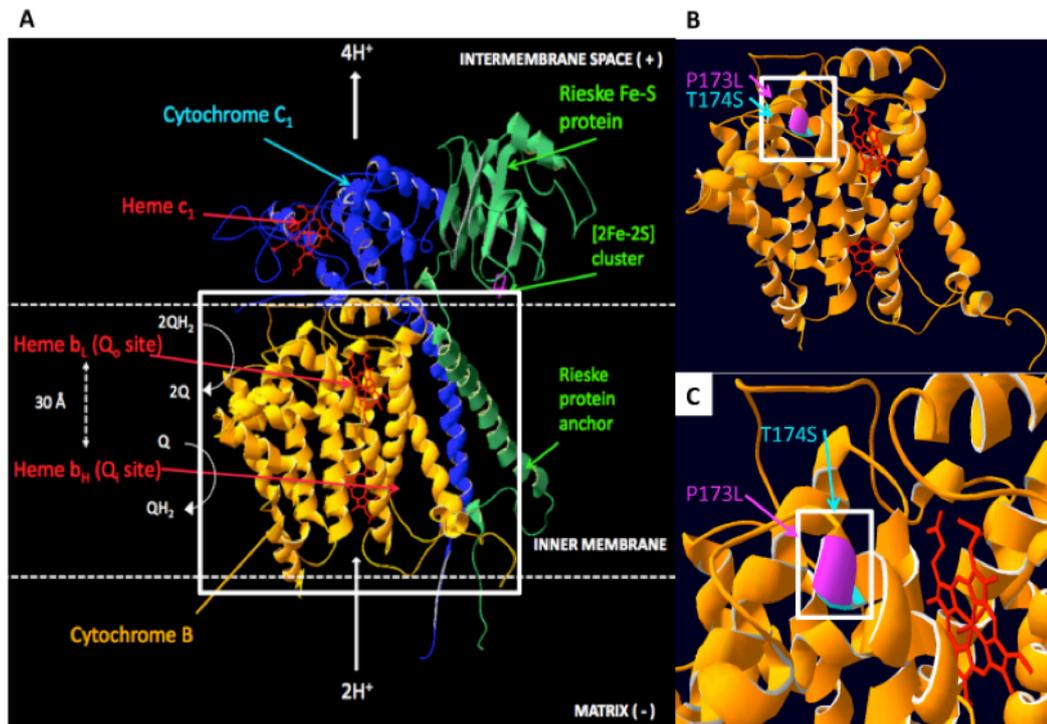
**Figure 3.9.** Location of the T45P mutation using the bovine model of Complex IV. (A) Protein structure of the catalytic centre of Complex IV, which comprises of COX I, II and III in their entirety. (B) Inset of the COX II subunit containing the T45P mutation on the transmembrane helical region. In the bovine model, a methionine exists in place of the threonine that is found at position 45 of the human COX II protein.

Although the variant at 9111bp of mtDNA resulted in a synonymous protein mutation, a base change at nucleotide position 14772bp appeared to induce a significant change (RI = 1) for sample S139 within the CYT B region (**Table 3.6**). This proline to leucine mutation at position 9 of the protein sequence (P9L) was predicted at the 95% confidence level to cause loss of glycosylation, and consequently loss of methylation on lysine downstream at position 12 of the protein sequence (**Table 3.6**). The P9L mutation was modeled on the bovine protein structure, which shares 78% protein sequence identity to its human counterpart, and identified to reside within the loop region (**Figure 3.10**).



**Figure 3.10.** Location of the CYT B P9L mutation using the bovine model. (A) The catalytic core of the cytochrome BC<sub>1</sub> complex. Events of the Q cycle are highlighted, which contribute towards the generation of an electrochemical gradient across the mitochondrial inner membrane. (B) The P9L mutation resides on a loop region of the bovine cytochrome B protein. (C) Magnified inset of the P9L mutation.

Mutations at positions 15264bp and 15267bp also resulted in amino acid changes (P173L and T174S), which were predicted to impair protein function (RI = 4 and 3 respectively) (**Table 3.6**). For these mutations, MutPred predictions on protein features were not considered significant ( $P > 0.05$ ). Both the P173L and T174S mutations were modeled on the bovine CYT B structure to reside on a helix structure within the mitochondrial inner membrane space (**Figure 3.11**). The presence of two adjacent mutations could disrupt helix-helix interactions that may be important in maintaining structure of the CYT B subunit.

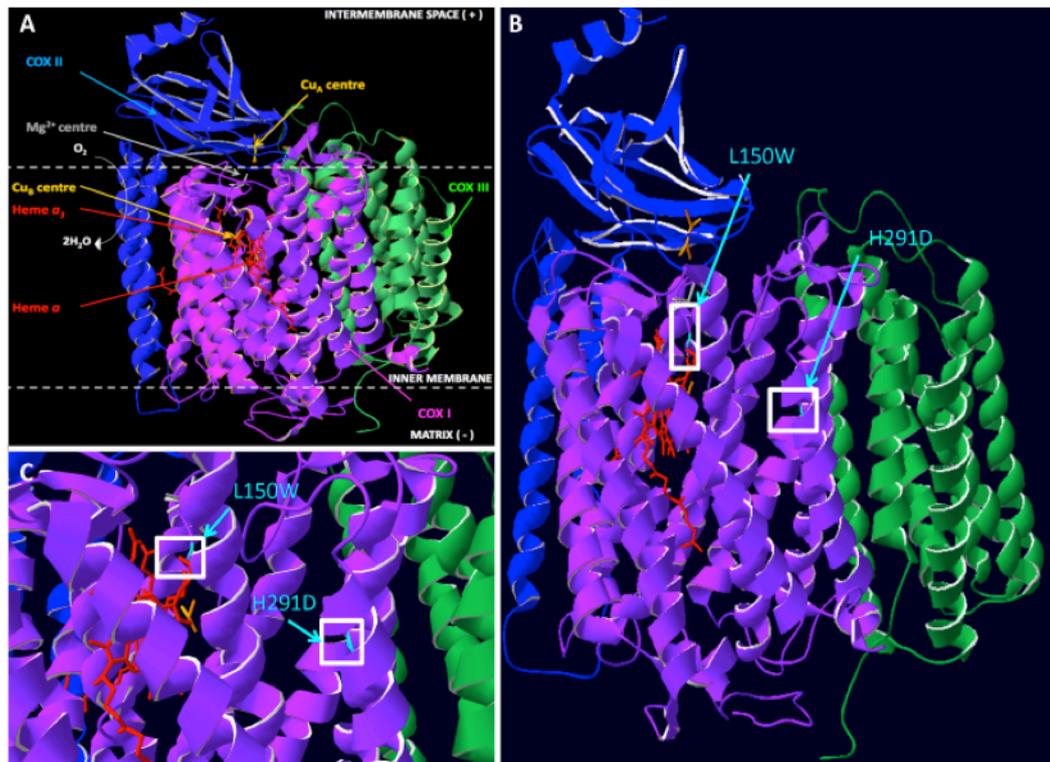


**Figure 3.11.** Mapping of the CYT B mutations P173L and T174S. (A) The catalytic core of Complex III. (B) The P173L and T174S mutations arise within a membrane-spanning helix region of cytochrome B. The T174S mutation represents a loss of a conserved threonine amino acid, whereas for the P173L mutation, the proline residue was replaced by alanine in the bovine model of CYT B subunit. (C) Magnified image of the P173L and T174S mutations.

### 3.5.9.2. CD45+ and CD45- sub-population variants

For the variants in the CD45+ and CD45- cells, the leucine to tryptophan (L150W) mutation present at position 6352bp was located within the loop region of COX I (Table 3.13 and Figure 3.12). Substitution of histidine to aspartate (H291D), another COX I mutation, was located on a transmembrane helical region (Figure 3.12). These observations corresponded to the *in silico* predictions, since the reliability index for L150W shows a value of 2, which supports the hypothesis that mutations in the loop regions do not exert much

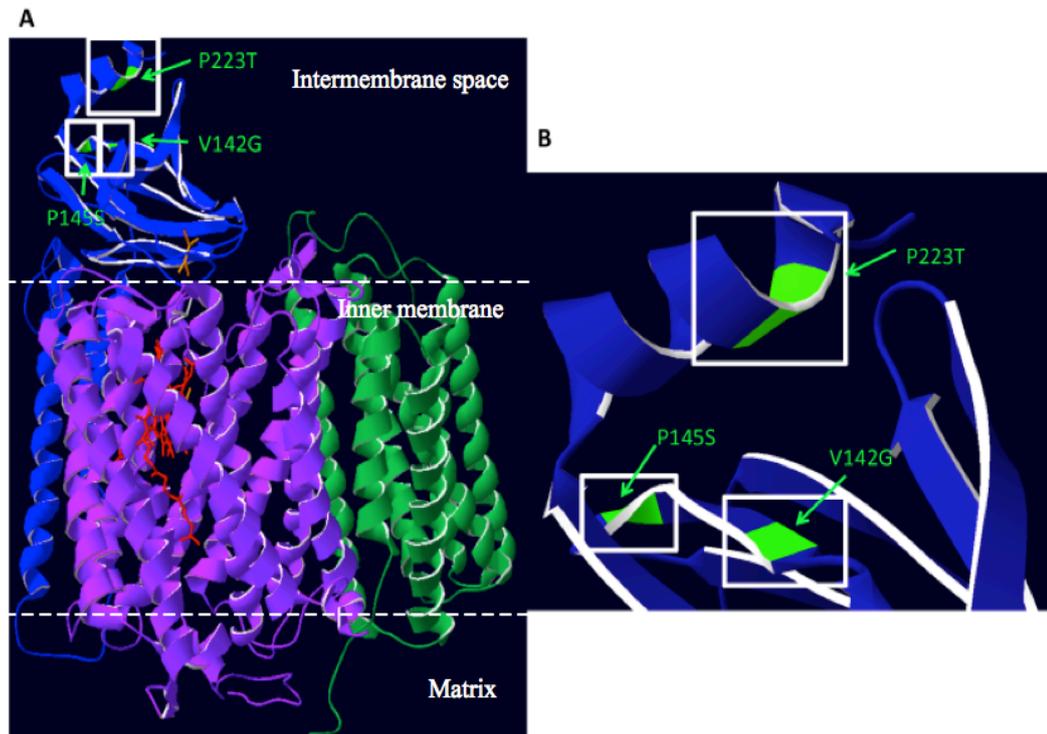
influence on the overall protein structure. Likewise, for H291D the reliability index was predicted to be 8, which matches the hypothesis that mutations in structural regions of the protein are more likely to be detrimental.



**Figure 3.12.** Modelling of the COX I mutations L150W and H291D in a bovine model. (A) Representation of the whole catalytic core of Complex IV. (B) Locations of the L150W and H291D mutations on COX I, both of which involve loss of a conserved residue between the human and bovine species. (C) Magnified inset of the L150W and H291D mutations.

Other mutations within the COX II and III regions were also mapped using the bovine protein structure. The V142G mutation lies within the 10-stranded beta barrel structure located at the C-terminal end of COX II, in the mitochondrial intermembrane space. Valine is often found within beta sheets, as it is a C-beta branched molecule (**Figure 3.13**). The V142G amino acid substitution was

predicted by MutPred to impair protein function and cause loss of stability to the protein structure (**Table 3.13**).

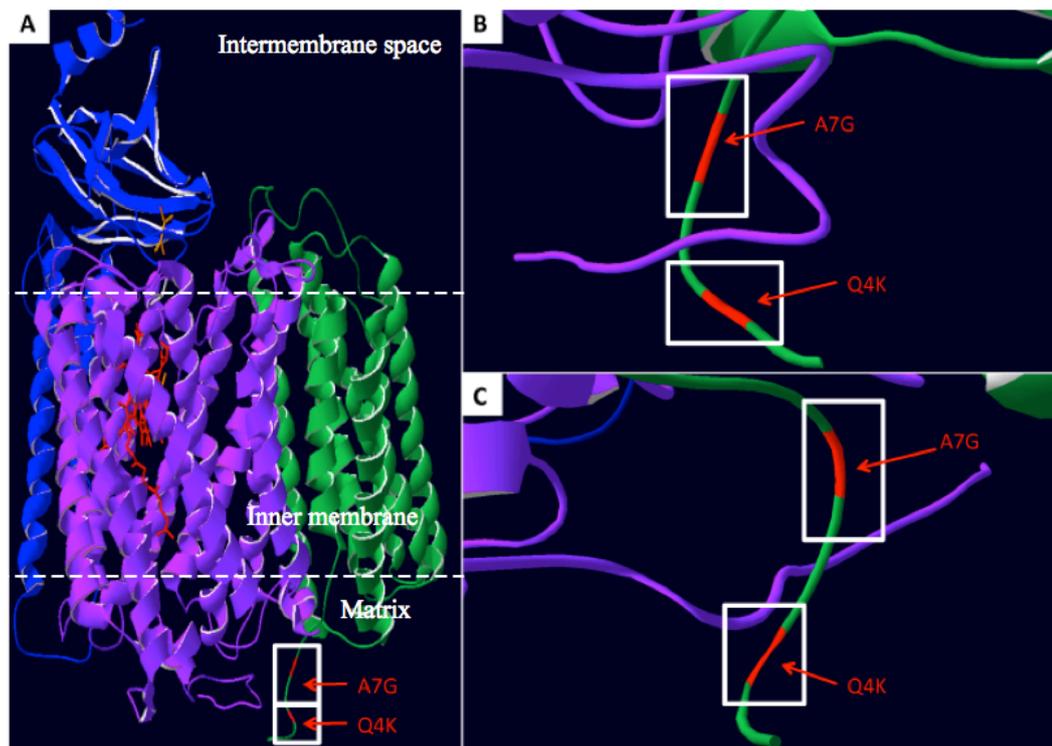


**Figure 3.13.** Mapping of the COX II variants within CD45+ and CD45- cells. (A) The P223T mutation found within the CD45+ cells, and the CD45- mutations V142G and P145S were located on the bovine COX II structure. The P223T mutation does not involve the loss of a conserved amino acid residue, as serine replaces the presence of a proline in the bovine model. However, both the V142G and P145S mutations involve the loss of a conserved residue. (B) Magnified view of the three COX II mutations.

Also within COX II, the P145S mutation located downstream of V142G resided on the same beta sheet structure (**Figure 3.13**). The P145S mutation was predicted to induce protein dysfunction, with a reliability index of 8, greater than that for V142G (RI = 1) (**Table 3.13**). Finally, the P223T mutation in COX II was modeled to reside within the intermembrane space region on a helix

structure (**Figure 3.13**). Due to its location further from the core region of Complex IV, it was predicted with confidence that the P223T mutation exerted a neutral effect (RI value = 9) (**Table 3.13**).

Two COX III mutations at positions 9216bp and 9226bp, causing amino acid mutations glutamine to lysine (Q4K) and alanine to glycine (A7G), were modeled on the bovine structure. Both were located on the loop regions close to the mitochondrial matrix (**Figure 3.14**), and were predicted with high confidence to be neutral mutations (RI values = 8) (**Table 3.13**).



**Figure 3.14.** Mapping of COX III variants within CD45<sup>+</sup> and CD45<sup>-</sup> cells. (A) Location of the A7G and Q4K mutations relative to the catalytic core of Complex IV. (B) Magnified inset of the two COX III mutations. (C) Alternative view.

Synonymous amino acid changes were reported at several sites, 3417bp (*ND1*), 8928bp and 8934bp (*ATP6*) and 11674bp (*ND4*) (**Table 3.13**). Where amino acid substitutions involved a transition to a stop codon (10657bp; *ND4L* and 11445bp; *ND4*), it was not possible to perform *in silico* predictions on protein function. However, premature stop codons would terminate mitochondrial protein synthesis.

The 10657bp variant that induces a substitution from methionine to a stop codon at position 63 of the *ND4L* polypeptide chain will produce a premature protein chain that is truncated by 35 amino acids. The resultant polypeptide will consist of only 64.3% of the complete protein sequence. For the 11445bp variant in *ND4*, only the first 50% of the polypeptide chain will be synthesised, due to the mutation generating a stop codon that arises exactly half-way in the protein wild-type sequence for *ND4*. Shortening of the polypeptide chain will prevent proper assembly of the protein complex, thereby disrupting its overall ability to function correctly. The earlier the termination of protein synthesis, the more deleterious the effect imposed upon the protein structure.

### **3.6. Discussion**

#### **3.6.1. Multiple myeloma variants**

Variants at positions 16186bp<sup>209</sup> (D-loop), 1389bp (12S rRNA)<sup>210</sup>, 3391bp (*ND1*)<sup>211</sup>, 5432bp (*ND2*)<sup>212</sup>, 8928bp<sup>213</sup>, 9111bp<sup>214</sup> (both within *ATP6*) and 11674bp<sup>215</sup> (*ND4*) have been previously reported, based on the human mitochondrial genome database, MitoMap. However, none of these have been associated specifically with multiple myeloma. This suggests that the variants in this Chapter represent novel discoveries, and could potentially be specific disease-causing variants.

CpG dinucleotide regions have been thought to be more susceptible than other sequence regions to mutation<sup>216</sup>. However, only 12.5% of variants identified from the tumor samples in **Table 3.6** were located within these regions (at 1389bp and 12847bp), more than that observed within the U266 cells in **Table 3.10** (8.7% of variants at 8252bp and 9226bp). These relatively low values either suggest that the incorporation of mtDNA variants occurs randomly, or highlights the presence of other factors that could influence these variants to arise in certain regions of the mtDNA over others.

#### **3.6.2. Impact of variants located in the non-coding regions of mtDNA**

While the impact of amino acid substitutions can be predicted using a combination approach of *in silico* and protein modeling analysis, the impact of mutations affecting the non-coding regions of the mtDNA can also be predicted based on our understanding of the biological functions for these genes.

MtDNA replication is only possible following initiation at the origin of heavy (H) strand replication<sup>217</sup>. The H-strand initiates transcription of mtDNA from two promoters, HSP1 and 2, which differ in their sites of initiation and termination of transcription, with the latter located 90-100bp downstream of HSP1<sup>218</sup>. In basal conditions, HSP1 demonstrates higher activity than at the light strand promoter (LSP) region<sup>219</sup>. Under the control of the heavy strand promoter, transcription results in generation of the 12s and 16s rRNA, as well as 12 mtDNA-encoded genes and 14 tRNAs<sup>220</sup>. Whereas, under the influence of the light strand promoter, transcription results in the production of 1 mtDNA-encoded gene (*ND6*) in addition to 8 tRNAs<sup>219</sup>. Nevertheless, the light strand promoter, which functions independently of the H-strand promoter<sup>220</sup>, remains important for the synthesis of RNA primers. Once processed via cleavage reactions, these RNA primers are required for mtDNA replication<sup>217</sup>.

Variants were identified within the H-strand promoter region (HSP1) at 557bp, (**Table 3.6**) and also within the origin of light strand replication (at 5752bp) (**Table 3.13**). The absence of a functional HSP1 site could severely reduce transcription for most of the mtDNA-encoded genes. On the other hand, disruption to the integrity of the origin of light strand replication, is likely to impair mtDNA replication, as *in vivo* mutagenesis studies suggest that this region is essential for maintaining mtDNA replication<sup>221</sup>.

Variants have also been identified in the 12s and 16s rRNA regions (**Tables 3.6 and 3.13**). Depending on the severity of these mutations, it is possible for mitochondrial transcription to be defective, as both 12s and 16s rRNA encode

mitochondrial ribosome subunits. Since there is absence of variant detection in any of the mtDNA tRNA gene regions, it is likely that the mitochondrial translational machinery remains functional in multiple myeloma.

### **3.6.3. Replenishment of mtDNA and mtDNA variants**

Variants that arose as a result of replenishment of mtDNA copy number for 14 days in U266 cells, after 14 days of mtDNA depletion, noticeably occurred within *ND5*. Previously, variants in the *ND5* region have been reported to enhance tumorigenicity of cells<sup>222</sup>. Re-populated mtDNA copies in the 14-day recovered U266 cells, could exhibit more tumorigenic properties compared to control U266 cells. Although predictions from SNPs&GO have determined these variants individually exert a neutral effect on the *ND5* protein, perhaps together the novel variants will be detrimental to Complex I function. This is likely to be the case, as mathematical modeling supports the concept that the accumulation of mutations promotes tumor establishment, demonstrating a positive correlation between the two variables<sup>223</sup>.

Prolonged culture of the 14-day recovered U266 cells is likely to increase the frequency of mutation for the low level variants that accumulate within mtDNA. This is because, with each mtDNA replication cycle, the presence of a heteroplasmic variant is likely to be maintained and depending on the fidelity of the mitochondrial replication machinery, further variants could be introduced. However, one study reported that the average frequency of heteroplasmy remains the same even across different stages of colorectal tumors<sup>224</sup>. Since not all tumors behave in the same manner, it remains possible for mutation levels to increase

alongside the duration of culture. For example, variants in the recovered sample may have been originally present in U266 control cells, but did not reach above the 3% threshold detection level until it underwent induced stress in the presence of ddC.

#### **3.6.4. Additional considerations**

According to clinical studies, multiple myeloma patients possessing a higher ratio of CD45<sup>-</sup> than CD45<sup>+</sup> plasma cells have lower survival rates, higher possibility of developing extra-medullary disease and display higher resistance to chemotherapy treatment<sup>207,225</sup>. Moreover, in the 5T2MM murine model of systemic myelomatosis, *in vivo* transition of predominantly CD45<sup>+</sup> 5T2MM cells into the CD45<sup>-</sup> counterparts, correlated with the development of more aggressive disease kinetics, neo-angiogenesis and the elaboration of angiogenic growth factors<sup>226,227</sup>. Furthermore, fractionated CD45<sup>-</sup> patient myeloma cells have greater engraftment capacity than CD45<sup>+</sup> cells in a xenotransplantation mouse model<sup>228</sup>. In support of this, results from this Chapter have revealed that mtDNA variants are more abundant in the CD45<sup>-</sup> population than in CD45<sup>+</sup> cells.

However, separation of the U266 cells into CD45<sup>+</sup> and CD45<sup>-</sup> cells may induce phenotypic and biological changes<sup>229</sup>. This could explain why variants in a heterogenous population of U266 cells were different to those detected in the sub-populations enriched for CD45<sup>+</sup> and CD45<sup>-</sup>. Furthermore, there may be inter-dependence between the two CD45 sub-populations that preserve a defined set of variants.

### 3.6.5. Protein modeling analysis

The results from the protein modeling analyses were examined in the biological context to confirm the biological reliability of these *in silico* results. The amino acid change from threonine to proline, caused by the variant at 7718bp within COX II, is likely to be biologically reliable in terms of its prediction on protein dysfunction (**Figure 3.9**). Threonine is a hydrophilic molecule due to possession of a hydroxyl group, and structurally can reside on either the interior or surface sides of a protein. Any substitution involving an amino acid change into proline, will impact on protein structure. Since the side chain in proline connects twice to the protein backbone, introduction of this imino acid induces rigidity to the local protein region. As a result, proline preferably resides on the protein surface.

While variants present in loop regions are not likely to impact significantly on the overall protein conformation, flexible loop regions remain important for protein-ligand and protein-protein interactions essential for mediating biochemical signal transduction<sup>230</sup>. For the substitution involving proline to leucine (P9L) at 14772bp within CYT B (**Figure 3.10**), it is possible that further flexibility of the local region will be induced, as proline is lost in this case. In addition, the L150W protein mutation reported in COX I, assigned a RI value of 2, is likely to be biologically reliable as the variant residing in the loop region is unlikely to initiate disease (**Figure 3.12**).

The RI value of 8, given to the second COX I mutation H291D, appears to be acceptable as the location of the variant appears buried within the interior of COX I (**Figure 3.12**). Furthermore, the H291D mutation is biochemically

unfavourable. Although histidine and aspartate are both polar molecules, the latter is also negatively charged. While histidine can reside within the interior or the surface of a protein, aspartate displays preference to reside on the protein surface. If however, aspartate is present on the interior side, it often participates in the formation of salt-bridges with other positively charged residues to create hydrogen bonds that assist in maintaining protein stability<sup>231</sup>.

Histidine at position 291 is also a copper B metal binding site (Uniprot P00395). Histidine is commonly found in enzyme reactive centers, as the imidazole ring present in histidine allows it to accept or donate protons near to physiological pH, making these amino acids ideal for enzyme catalysis reactions<sup>232</sup>. Substitution of histidine is likely to interfere with the iron present in heme a<sub>3</sub>, which functions to form a heterobimetallic site with the Cu<sub>B</sub> centre<sup>233</sup>, since histidine 291 represents one of three histidine residues participating in the stabilization of the Cu<sub>B</sub> centre<sup>234</sup>. Altogether, these observations suggest that loss of the conserved histidine at position 291 can disrupt functionality of the catalytic cycle. This will interfere with the transfer of electrons from the hydrophobic carrier ferrocyanochrome C to the heme a<sub>3</sub>/Cu<sub>B</sub> centre.

The amino acid substitution V142G in COX II represents another mutation that is likely to biologically induce disease (**Figure 3.13**). Valine is limited in the conformations it can form, since there is the presence of two non-hydrogen side chains. While valine is unreactive, its hydrophobic property allows it to participate in interactions with other hydrophobic molecules, for example, lipids<sup>231</sup>. Glycine on the other hand, carries only hydrogen on its side chain. Its

small size means that presence of glycine will confer flexibility of the protein structure where it resides. The loss of rigidity in protein structure from a substitution involving valine to glycine, could interfere with stability of the beta barrel structure of COX II, a consequence which was reflected in the *in silico* predictions made in **Table 3.13** (P = 0.002).

With reference to the P145S mutation also in the COX II protein (**Figure 3.13**), the substitution event may be unfavorable due to serine primarily being found in functional centers, as its hydroxyl group is capable of hydrogen bonding with a range of polar molecules<sup>231</sup>. Both V142G and P145S were found to be present in both the CD45+ and CD45- cells. The combination of both is likely to severely reduce stability of the beta barrel structure. The final COX II mutation, P223T, was predicted to induce a neutral effect (**Table 3.13**), most likely due to its location being relatively distant from the main catalytic core center (**Figure 3.13**).

It is probable that the two mutations residing adjacent to each other on the helical segment of CYT B (**Figure 3.11**) exert a greater effect on protein function than the presence of a single mutation. The biochemical effect of a transition from proline to leucine has been discussed previously. On the other hand, a change from threonine to serine at position 174 on the protein sequence is considered a favourable substitution. Both residues are polar in nature, with the chemical difference between them being a hydrogen group in serine, whereby a methyl group is found at the same location in threonine. However, together these mutations may disrupt overall structure to CYT B. Given that variants 15264

and 15267bp appear common across the myeloma tumors screened, their joint presence is likely detrimental to protein stability.

Not surprisingly, all non-synonymous ATP6 mutations (F78C, P134H, I138N and S99P) were predicted to impair protein function. Due to its essential role in the production of ATP, loss of ATP6 function is likely to disrupt respiratory capacity of a cell, thereby promoting aerobic glycolysis. According to the Warburg theory, under these conditions, proliferation of the tumor cells will increase, therefore enhancing malignancy of the tumor.

A study performed in the *Chlamydomonas reinhardtii* model have presented results to support the theory that suppression of ND4L polypeptide production impairs assembly of the entire Complex I protein structure<sup>235</sup>. Moreover, ND4 participates in ubiquinone-binding<sup>236</sup> as well as in the translocation of protons across the mitochondrial inner membrane<sup>237</sup>. Therefore, incomplete translation of both these gene regions, caused by the introduction of pre-mature stop codons in the mtDNA sequence, are likely to impose adverse implications.

### **3.6.6. Contribution of mtDNA variants to tumor biology**

The development of multiple myeloma requires the collective influence of variants, which has been the case for disorders such as MELAS, whereby the G13513A mutation was found to be associated with the disease<sup>238</sup> alongside the characteristic 3243bp mutation<sup>239</sup>. Acquisition of these cumulative mtDNA variants, leads to dysfunction of several of the protein complexes that are responsible for respiration via OXPHOS.

Inefficient energy production through OXPHOS could be a stimulant for the cells to switch metabolism towards glycolysis, even in the presence of sufficient oxygen concentration (the Warburg effect). However, if mtDNA variants only impair Complex I function, electrons are still able to enter the electron transfer chain through Complex II, which receives electrons from succinate and transfers these to FADH<sub>2</sub> and the hydrophobic carrier coenzyme Q, for shuttling to Complex III<sup>240</sup>. Electron transfer initiating from Complex II yields fewer protons overall for use in the synthesis of ATP through Complex V, as Complex I would not be contributing to the generation of a proton gradient in this situation.

### **3.7. Conclusion**

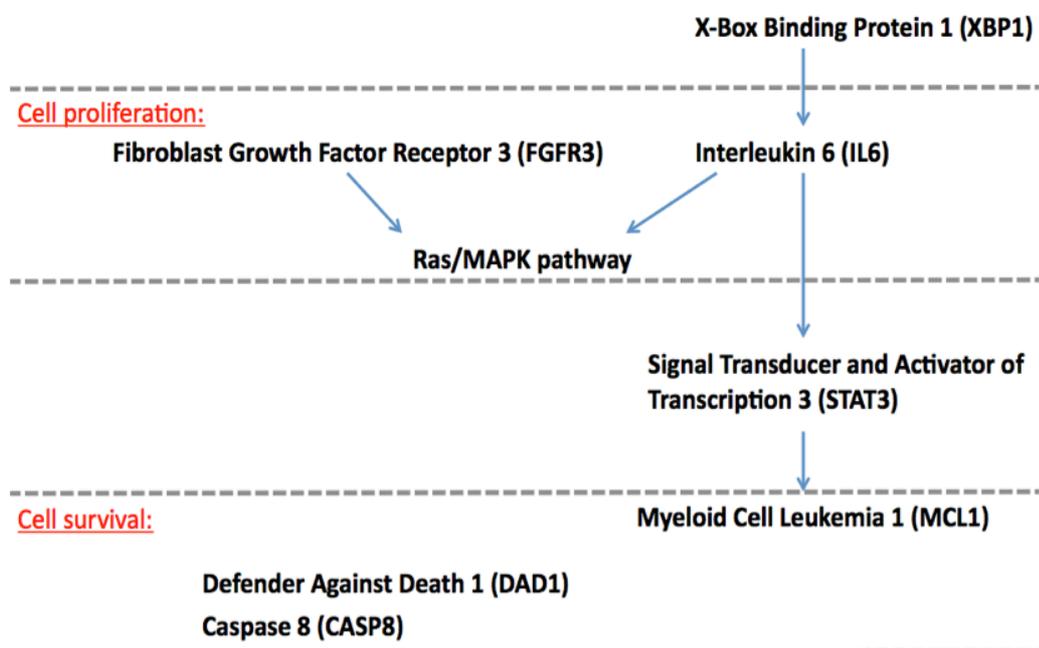
Different multiple myeloma tumors characteristic of patients in their relapse and/or refractory phase, are capable of exhibiting unique mutations that could enable each to behave in a different manner and to progress at variable rates during the course of tumor establishment. Cytochrome B was determined as a potential hotspot region for mutations, especially at nucleotide positions 15264bp and 15267bp. However, the 12S rRNA region was the most susceptible to development of mutations.

Accumulation of mtDNA variants could induce cells to respire via aerobic glycolysis, therefore promoting the progression of multiple myeloma. It is possible that the accumulation of these variants could simultaneously alter the ratio of CD45+ and CD45- cells to increase the CD45- malignant sub-population, thereby further encouraging malignancy.

**Chapter 4: The Role of MtDNA in the  
Maintenance of Mitochondria and  
Multiple Myeloma**

#### **4.1. Introduction**

Multiple myeloma cells have the ability to migrate to the bone marrow, where they are able to interact with cells within the bone microenvironment, such as the stromal cells<sup>241</sup>. The interactions formed are able to stimulate signal transduction pathways that promote the proliferation and metastatic potential of these tumor cells. Some of the key factors involved with the tumor biology of multiple myeloma are represented in **Figure 4.1**.



***Figure 4.1.*** Key factors in the tumor biology of multiple myeloma. CASP8 and DAD1 are additional proteins that are important to the development of this cancer, but do not form interactions with the other tumorigenic proteins mentioned.

Interleukin 6 (IL6) is an *in vivo* growth factor produced mainly by stromal cells in the bone marrow to promote the growth of multiple myeloma. It is capable of stimulating the Ras/MAPK cell proliferation pathway<sup>242</sup>, and is considered

critical, since proliferation is inhibited in the presence of anti-IL6 antibodies<sup>243,244</sup>. However, IL6 is not consistently found to correlate with multiple myeloma<sup>147</sup>. Homodimerisation of gp130 is induced following binding of IL6 to its receptor, which initiates signalling via the RAS/MAPK and Janus-activated kinase (JAK)/signal transducers and activators of transcription (STAT) pathway<sup>245</sup>. Fibroblast growth factor receptor 3 (FGFR3) is another molecule that utilises the RAS/MAPK pathway, and is responsible for the regulation of cellular proliferation<sup>246</sup>. Downstream of the IL6 signalling cascade, activation of the transcription factor STAT3, upon its phosphorylation, enhances the expression of survival-related genes<sup>241,247</sup>.

X-box binding protein 1 (XBP1) functions upstream of IL6 and is capable of increasing gene expression levels and secretion of IL6, by interacting and binding to the IL6 promoter<sup>248</sup> at the NFkB site to control its gene expression levels<sup>249</sup>. IL6 does not influence expression of the survival factor MCL1<sup>250</sup>, despite a previous report suggesting otherwise<sup>251</sup>. Instead, MCL1 expression is thought to be regulated at the post-translational level, with one study suggesting that it is the interaction between deubiquitinase USP9X and MCL1 that allows for stabilisation of the survival protein<sup>252</sup>. This provides evidence that increases in MCL1 protein levels occur independently of the regulatory mechanisms during transcription<sup>252</sup>. However, others have suggested that STAT3 influences the expression of MCL1<sup>253,254</sup>, therefore indicating the presence of multiple regulatory mechanisms that control MCL1 expression.

Apart from ATP generation, mitochondria also have a role in the control of apoptosis, which is mediated via the intrinsic or extrinsic pathways. The intrinsic apoptosis pathway releases Cytochrome c from damaged mitochondria, which forms a complex with APAF-1 to generate an apoptosome<sup>255</sup>. The extrinsic pathway on the other hand, functions independently of mitochondria following the stimulation of a receptor-mediated signalling response. Different caspases have distinct roles during apoptosis; caspase 8 and 10 function as initiators of the extrinsic pathway, whereas caspase 9 is associated with the intrinsic mitochondrial-dependent pathway<sup>256</sup>. Caspase 8 is unique in that it can also mediate the effects of cell survival<sup>257,258</sup>. In addition, defender against death 1 (DAD1) represents another factor participating in the control of cell survival<sup>259</sup>, which is believed to be highly conserved across a range of plant and animal species<sup>260</sup>.

2'3'-dideoxycytidine (ddC) is a nucleoside analog that was previously used as an anti-retroviral drug for the treatment of human immunodeficiency virus (HIV)<sup>261</sup>. ddC prevents mtDNA replication by inhibiting activity of the mitochondrial POLG complex<sup>262</sup>. This is because the lack of a hydroxyl (-OH) group at the 3' position of the carbon ring within a nucleotide base, which is normally present in deoxynucleotides prevents further addition of other nucleotides during DNA synthesis<sup>263</sup>. ddC is selective towards inhibiting replication of mtDNA, since it has been reported to exert no effect on the DNA polymerase  $\alpha$  complex, which is the enzyme responsible for chromosomal DNA synthesis<sup>264</sup>. Moreover, it has been reported that in the presence of ddC, mtDNA replicates at a rate 50,000 fold slower than in its absence<sup>265</sup>.

Since the Warburg effect is characteristic to many cancer types, the use of 2-deoxyglucose (2-DG) as a potential anti-cancer drug has gained interest in recent years<sup>266</sup>. 2-DG is a competitive analog of glucose and functions to inhibit glycolysis in its phosphorylated state, since 2-DG-6-phosphate is unable to be further metabolized by the body<sup>267</sup>. Clinical trials have so far determined that oral administration of 2-DG at doses of 250-300mg/kg of body weight is safe for GBM patients<sup>268</sup>. However, pre-clinical results from mouse models have determined that higher doses can induce cardiac toxicity, which has limited its use as a successful drug in the treatment of GBM, and potentially other cancers. Recently, there has been growing interest to suggest that combination therapy to include the use of 2-DG, together with mitochondrial inhibitors, such as antimycin A or oligomycin, could be a possible approach to induce apoptosis, which was tested in several breast cancer cell lines<sup>269</sup>. However, to date, this remains a controversial research area. Nonetheless, 2-DG is an effective inhibitor of glycolysis for *in vitro* studies.

In this Chapter, multiple myeloma cells were depleted of their mtDNA using ddC, and allowed to re-populate in the presence and absence of 2-DG. Metabolic measurements were also taken to identify any changes to the production of ATP or lactate in these cancer cells during the process of mtDNA depletion.

#### **4.2 Aims and hypothesis**

The aim was to assess the importance of mtDNA in the maintenance of cell proliferation in multiple myeloma cells, by examining mtDNA copy number

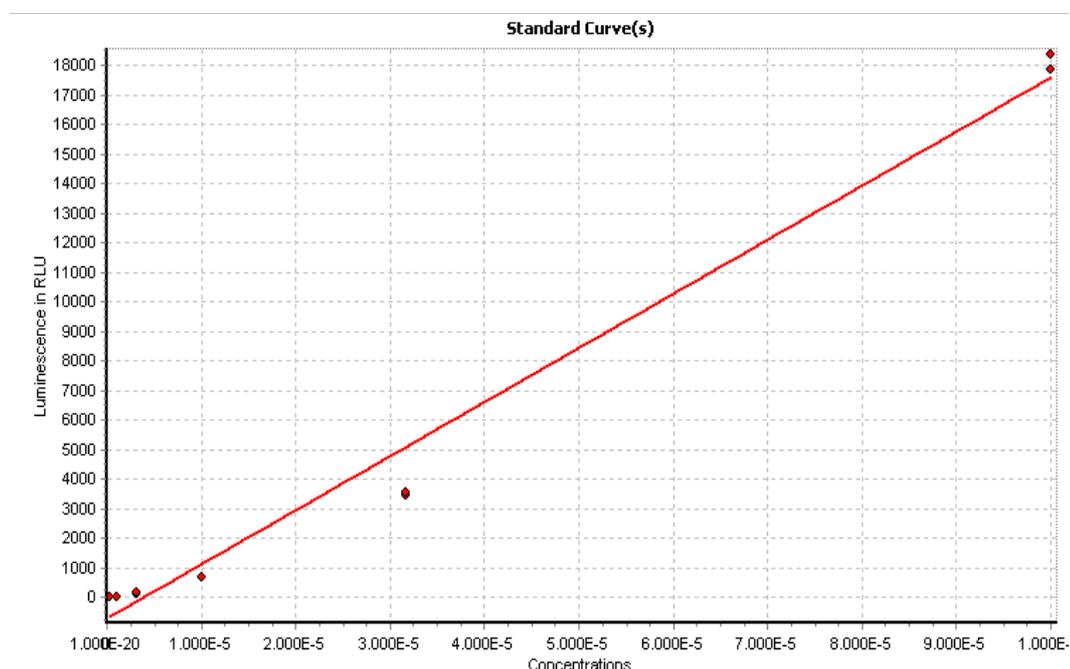
alongside expression analyses of mtDNA and nuclear-encoded genes that are involved in the processes of mitochondrial biogenesis, transcription and replication. Several key tumor biology genes associated with multiple myeloma were examined to determine the influence of mtDNA copy number on the malignant phenotype of these cells. The contribution of the glycolysis pathway towards maintaining the tumorigenic phenotype of these cancer cells was also investigated.

It was anticipated that the expression of the mtDNA-encoded genes would decrease during mtDNA depletion, as other studies have reported this using ddC<sup>270</sup>. Although the use of other mtDNA depletion agents, including ethidium bromide, result in same outcome<sup>271</sup>, these have been reported to exert greater non-specific effects by altering cell metabolism, for example, by increasing superoxide and lactate production<sup>272</sup>. Here, it was hypothesised that the reduction of mtDNA copy number in multiple myeloma cells using ddC would increase the expression of genes involved with tumor development, and perhaps also genes involved with the maintenance of mtDNA. This was because other studies have reported mtDNA depletion to be capable of enhancing the malignant phenotype of prostate carcinoma<sup>273,274</sup>. Furthermore, it was hypothesized that by culturing cells under conditions that inhibit glycolytic activity, the tumorigenic properties of these cancer cells would be suppressed, as cells could be forced to utilize OXPHOS in the absence of glycolysis.

### **4.3. Preliminary experiments**

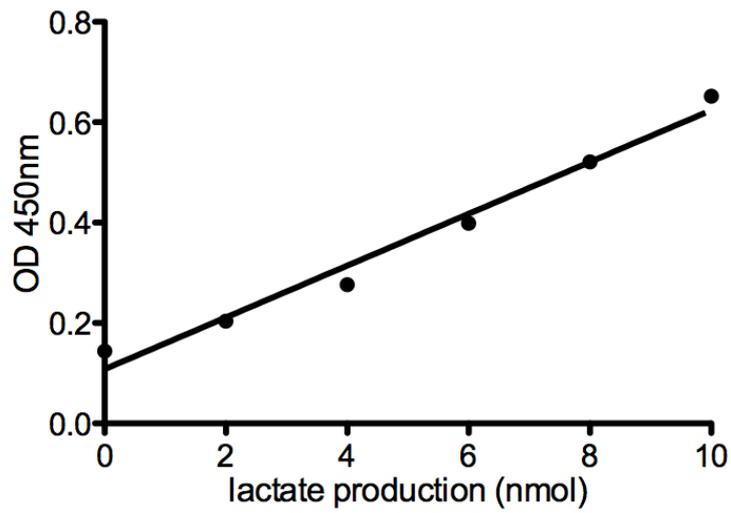
#### **4.3.1. Standard curves for analyzing ATP and lactate production**

For quantification of ATP and lactate production in cells, dilutions of a known stock concentration (10mM) was performed for the generation of a standard curve from which it would be possible to extrapolate data from. For the ATP assay, the following standards were used:  $10^{-4}$ ,  $10^{-4.5}$ ,  $10^{-5}$ ,  $10^{-5.5}$ ,  $10^{-6}$ ,  $10^{-6.5}$ . Results produced a standard curve with a linear regression fit of  $r^2 = 0.98$  (Figure 4.2).



**Figure 4.2.** Standard curve for measuring ATP production. ( $r^2 = 0.98$ ).

A standard curve dilution was also performed for use with the lactate colorimetric assay. The following concentrations were used: 0nmol, 2nmol, 4nmol, 6nmol, 8nmol and 10nmol. The linear regression fit gave an  $r^2$  value of 0.97 (Figure 4.3).

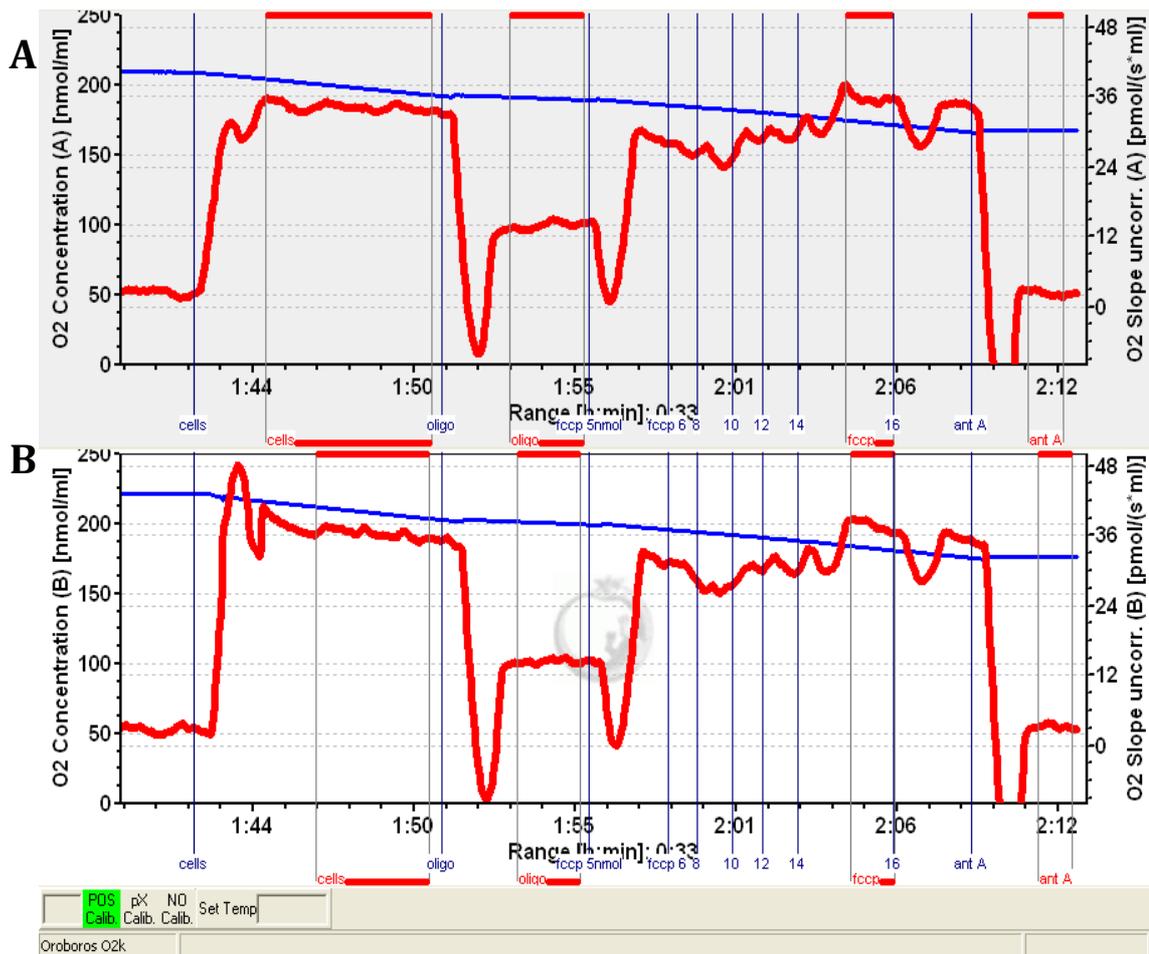


***Figure 4.3.*** Standard curve for measuring lactate production. ( $r^2 = 0.97$ ).

Standard curves were performed with each experimental run to ensure consistency of the dilutions and to obtain an accurate value for the production of lactate and/or ATP from cell cultures.

### 4.3.2. Measuring oxygen consumption rates using respirometry

Respirometry measurements were taken to identify the preferential method through which glucose was metabolized in cells. The outcomes of this experiment were informative of how dependent the cells are on mitochondrial function, and therefore on mtDNA copy number. However, prior to the experiments, optimization had to be performed on various concentrations of the uncoupler drug Trifluorocarbonylcyaniide Phenylhydrazone (FCCP) on the cell type of interest. The drug FCCP determines the maximal respiratory capacity in cells, and so the highest concentration that could be used, without inhibiting further respiration, had to be optimized (**Figure 4.4**).



**Figure 4.4.** Determining the optimal concentration of Trifluorocarbonylcyaniide Phenylhydrazone (FCCP) using respirometry.

Optimisation was performed using U266 cells. The red line represents oxygen flux, and the blue represents total oxygen concentration within the closed chambers; both were measured as a function of time and detected by the polarographic oxygen sensors. Blue vertical lines represent the addition of events throughout the duration of the experiment, with labels signifying the following events: “Cells” – the addition of U266 cells into the closed chambers; “Oligo” – the addition of oligomycin into the closed system to determine the amount of proton leakage during respiration; “FCCP (nmol)” – the total concentration of FCCP within the cell chamber, ranging from 5nmol to 16nmol; “Ant A” – the addition of antimycin A to terminate mitochondrial oxygen consumption. Panels (A) and (B) are duplicate experiments to illustrate the reproducibility of the assay between the two independent cell chambers on the Oxygraph O2K instrument.

Titration of the various FCCP concentrations ranged from 5nmol (lowest) and was increased at increments of 1 or 2nmol, up to the point at which maximal respiration can no longer be sustained after addition of further FCCP. A concentration of 14nmol FCCP was identified to be optimal for U266 cells, as 16nmol was inhibitory for stimulating further respiration. The optimization was repeated in a similar manner for additional cell lines.

## 4.4. Methods

The methods used in this Chapter are described in the following sections: 2.1.4 to 2.1.6, 2.1.9 to 2.2, 2.4.1, 2.5.1, 2.6 to 2.8.2, 2.9.2 and 2.11 to 2.14.

### 4.4.1. Real-time PCR primers for gene expression and mtDNA copy number analysis

Primers used for determining gene expression changes in U266 cells during, and after, mtDNA depletion are displayed in **Table 4.1**. The primer sets used for determining mtDNA copy number are also shown.

**Table 4.1.** Primers used to examine the expression of genes and mtDNA copy number.

Gene	Abbreviation	Forward/ Reverse	Sequence (5'→3')	Annealing temperature (°C)	Product size (bp)
ATP synthase F0 subunit 6	ATP6	F	CAGTGATTATAGGCTTTCGCTC	56	343
		R	GTGTTGTCGTGCAGGTAGAG		
Beta actin	β-actin	F	CAAAACCTAACTTGCGCAGA	53	261
		R	TTTtaggatggcaagggact		
Beta globin	β-globin	F	CAACTTCATCCACGTTCCACC	56	268
		R	GAAGAGCCAAGGACAGGTAC		
Caspase 8	CASP8	F	CTGCTGGGGATGGCCACTGTG	63	113
		R	TCGCCTCGAGGACATCGCTCTC		
Cytochrome B	CYT B	F	ATGACCCCAATACGCAAACT	54	401
		R	GGGAGGACATAGCCTATGAA		
Cytochrome C oxidase subunit II	COXII	F	CTGCTCCTAGTCCTGTATG	56	235
		R	GTCGGTGTACTCGTAGGTTT		
Defender against cell death 1	DAD1	F	CATCCTAGCGTTTGCCTGA	56	282
		R	AGCAGAGCTAGCAGTAAGTG		
Fibroblast growth factor receptor 3	FGFR3	F	TGCTGAATGCCTCCCACG	57	123
		R	CGTCTTCGTCATCTCCCGAG		
Interleukin 6	IL6	F	GCACTGGCAGAAAACAACCT	55	253
		R	CAGGGGTGGTTATTGCATCT		
Mitogen- activated kinase 2	MAPK2	F	GAGATCAGCATTTGCATGGAACAC	59	323
		R	TCCGACTGCACCGAGTAATGTG		
Myeloid cell leukemia 1	MCL1	F	CTGGAGTAGGAGCTGGTTTG	55	439
		R	GTGCTTCTCTTAACACTACAG		
Mitochondrial copy number	MTDNA	F	CGAAAGGACAAGAGAAATAAGG	53	151
		R	CTGTAAGTTTAAAGTTTATGCG		
Mitochondrial single-stranded DNA binding	MTSSB	F	GAAGCCATGTTTCGAAGACCTG	57	271
		R	CTGATATTCTGTGCCATGTTGTC		

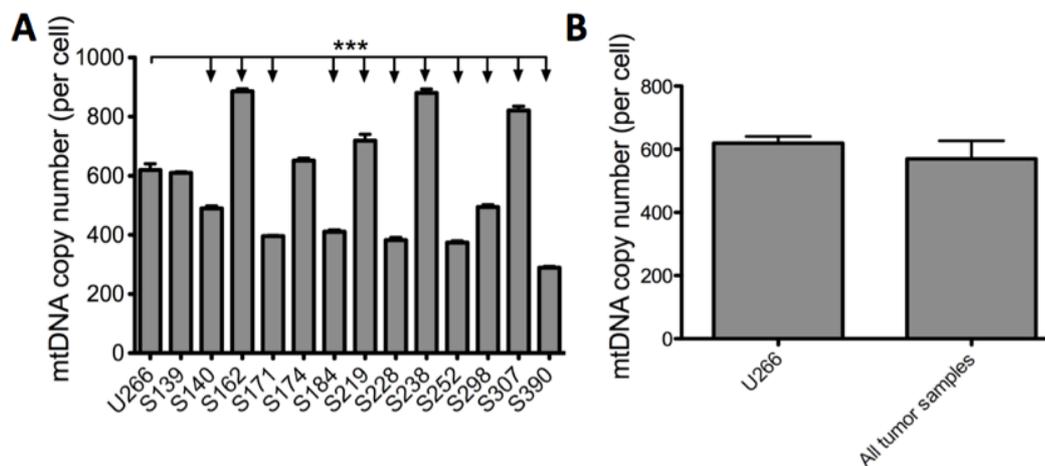
**Table 4.1: Continued**

Gene	Abbreviation	Forward/ Reverse	Sequence (5'→3')	Annealing temperature (°C)	Product size (bp)
Mitochondrial transcription termination factor	MTERF	F	CGGTTTATTGCAGGGAAGTAC	56	243
		R	TGGAGGTGATCGCATCTTTC		
NADH dehydrogenase subunit 1	ND1	F	TCTACCATCGCTCTTCTAC	56	349
		R	GGTTGGTCTCTGCTAGTGTG		
NADH dehydrogenase subunit 6	ND6	F	CCGCACCAATAGGATCCTCCCGA	63	187
		R	GCATGGGGGTGAGGGGTTGAG		
Nuclear respiratory factor 1	NRF1	F	GCCACAGCCACACATAGTAT	55	348
		R	AAGCGCCATAGTGACTGTAG		
Nuclear respiratory factor 2β	NRF2β	F	ATCCAAGGCAACAGATGAAACGGG	60	215
		R	TGATGACTTGCTGACCCCTGA		
Peroxisome proliferative activated receptor, gamma, coactivator 1 alpha	PGC1α	F	TTGCGCAGGTCAAACGAAAC	56	241
		R	GCGCTGTCCCATGAGGTATT		
DNA polymerase gamma subunit 1	POLGA	F	CACACCTAAACTCATGGCAC	56	436
		R	GTCCACGTCGTTGTAAGGTC		
DNA polymerase subunit gamma-2	POLGB	F	GTTTGCCATGAGTCCATCTAAC	56	280
		R	CTCTGTCAGCTGGAAAGAATC		
Peroxisome proliferator- activated receptor delta	PPARδ	F	GCGGGAGAATTCTGCGGAGCC	62	154
		R	CCCATCAGCCTCGTTGGTGCAT		
NAD- dependent deacetylase sirtuin-3	SIRT3	F	GAAGTGGAGGCAGCAGTGACAAG	62	432
		R	TGCAGGTGGCAGAGGCAAAGGTTC		
Signal transducer and activator of transcription 3	STAT3	F	GTTCTCTATCAGCACAATCTAC	55	435
		R	CAGCTCACTCAGATGCTTC		
Mitochondrial transcription factor A	TFAM	F	ATTGGGGTCGGGTCCTGCCTCA	63	361
		R	TACCTGCCACTCCGCCCTATAAGC		
Mitochondrial transcription factor B1	TFB1M	F	GCCTCCGGAAAAGTCTCAGCACTTG	61	249
		R	TTCAACCACCAGAAGTTCAGCGAC		
Mitochondrial transcription factor B2	TFB2M	F	CGAAAGCATTGCGGGGAGG	61	273
		R	GAGCGCAACCACTTTGGCACCA		
Progressive external ophthalmoplegia 1	TWINKLE	F	GCACAAGTCCATCGTATCTTTC	56	197
		R	CATACTCACTGATGAATGTCGTC		
X-box binding protein 1	XBP1	F	GTAGTCTGGAGCTATGGTGGT	57	283
		R	TCGATCTCTGGCAGTCTGAG		

## 4.5. Results

### 4.5.1. MtDNA copy number in multiple myeloma cells and patient tumor samples

MtDNA copy numbers in clinical samples were obtained from patients with the confirmed diagnosis of multiple myeloma. These were compared to mtDNA copy number in U266 cells (**Figure 4.5A**). The average mtDNA copy number amongst all tumor samples ( $566.8 \pm 28$ ) revealed that there was no significant difference between the U266 mtDNA copy number ( $619.3 \pm 21.5$ ) and that from the tumor samples ( $P = 0.60$ ) (**Figure 4.5B**).



**Figure 4.5.** MtDNA copy number for U266 cells and patient tumor samples. (A) MtDNA copy numbers for all tumor samples (B) The mean mtDNA copy number for all tumor samples, presented against U266 cells. All data are represented by mean  $\pm$  SEM. \*\*\*  $P = \leq 0.001$ .

### 4.5.2. Oxygen consumption rates in the U266 cell line.

The respiratory capacity of U266 cells was compared with the oxygen consumption readings from other cancer cell lines. These included the HSR-

GBM1 and hepatocarcinoma (HepG2) cell lines (**Table 4.2**). HepG2 cells were included as they represent a highly metabolically active cell type, which means that these cells are highly dependent on their mitochondrial function. The HSR-GBM1 cells were included as it is the best characterized GBM cell line, expressing neural stem cell markers, such as CD133, NESTIN and MUSASHI1<sup>275</sup>. The HSR-GBM1 cell line would also be used in the following two Chapters for determining the role of mtDNA in GBM tumor biology, which makes it an appropriate cell line for comparison against the U266 cells.

**Table 4.2.** *Respirometry measurements on U266, HSR-GBM1 and HepG2 cells.*

	Oxygen Consumption (pmol/s*Mil)			ETC coupling	ETC reserve capacity
	Basal	Non-phosphorylating	Maximal	(Basal / Non-phosphorylating)	(Maximal / Basal)
U266	17.01 ± 1.72	9.00 ± 0.912	29.44 ± 2.02	1.89 ± 0.06	1.73 ± 0.04
HSR-GBM1	14.86 ± 0.57	11.45 ± 0.56	36.17 ± 0.10	1.30 ± 0.05	2.43 ± 0.003
HepG2	27.22 ± 1.99	9.89 ± 0.55	33.12 ± 0.32	2.75 ± 0.32	1.22 ± 0.02

U266 cells have relatively low rates of oxygen consumption (approximately 17pmol/(s\*Mil), which is almost half of that detected for hepatocarcinoma (HepG2) cells but not as low as for HSR-GBM1 cells (**Table 4.2**). HSR-GBM1 cells were identified to have basal oxygen consumption rates of 14.86 ± 0.57pmol/(s\*Mil), which is within range of the values reported from other studies, namely between 10-17pmol/s\*Mil<sup>166</sup>.

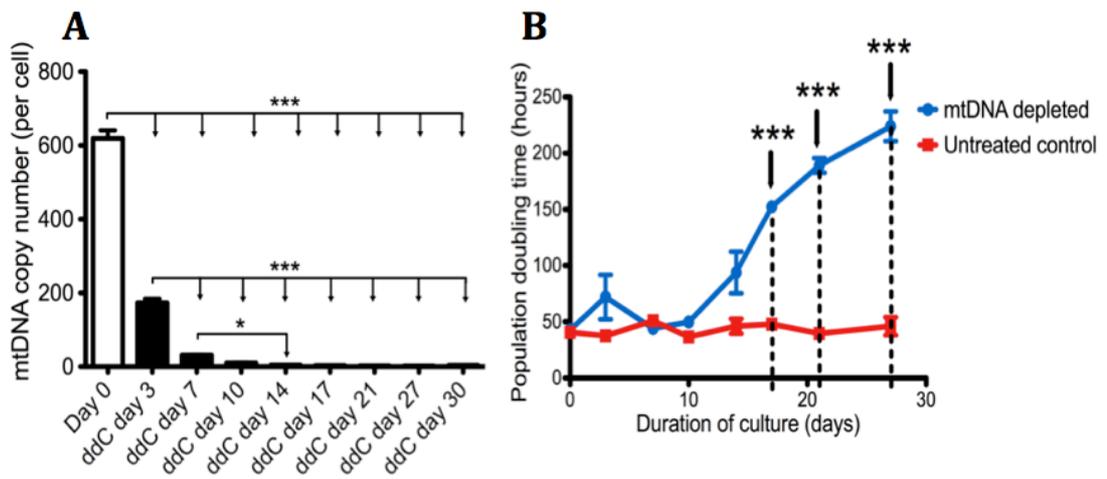
Comparing U266 cells to other tumor types, ovarian cancer cells have even higher basal rates for oxygen consumption around 42pmol/s\*Mil<sup>276</sup>. U266 cells

also have spare oxygen reserve capacity, since values above 1.0 for the ETC reserve capacity, indicate that these cells can increase their basal oxygen consumption rates to match increases in energy demands. Additionally, activity of the U266 ETC is tightly coupled to OXPHOS, but not as tightly coupled as HepG2 cells, since values close to 2 are representative of this feature. Altogether, based on the differences between the basal and maximal oxygen consumption rates, these respirometry results show that HepG2 cancer cells respire predominately via OXPHOS, whereas both the U266 and HSR-GBM1 cells respire primarily via glycolysis. However, there is still active OXPHOS activity within both U266 and HSR-GBM1 cells.

#### **4.5.3. Depletion of mtDNA in U266 cells.**

U266 cells were then depleted of their mtDNA for over 30 days (**Figure 4.6A**). From day 10 of depletion, there was a significant decline in mtDNA copy number, suggesting that the maximal effect of ddC was exerted by this stage of culture.

Analysis of cell population doubling times, defined as the duration of time required for the complete division of a single cell<sup>277</sup>, revealed that the rate of proliferation in U266 cells decreased with prolonged exposure of ddC (**Figure 4.6B**), whilst non-treated cells continued to proliferate at a consistent rate with a population doubling time of  $43.16 \pm 1.73$  hours.



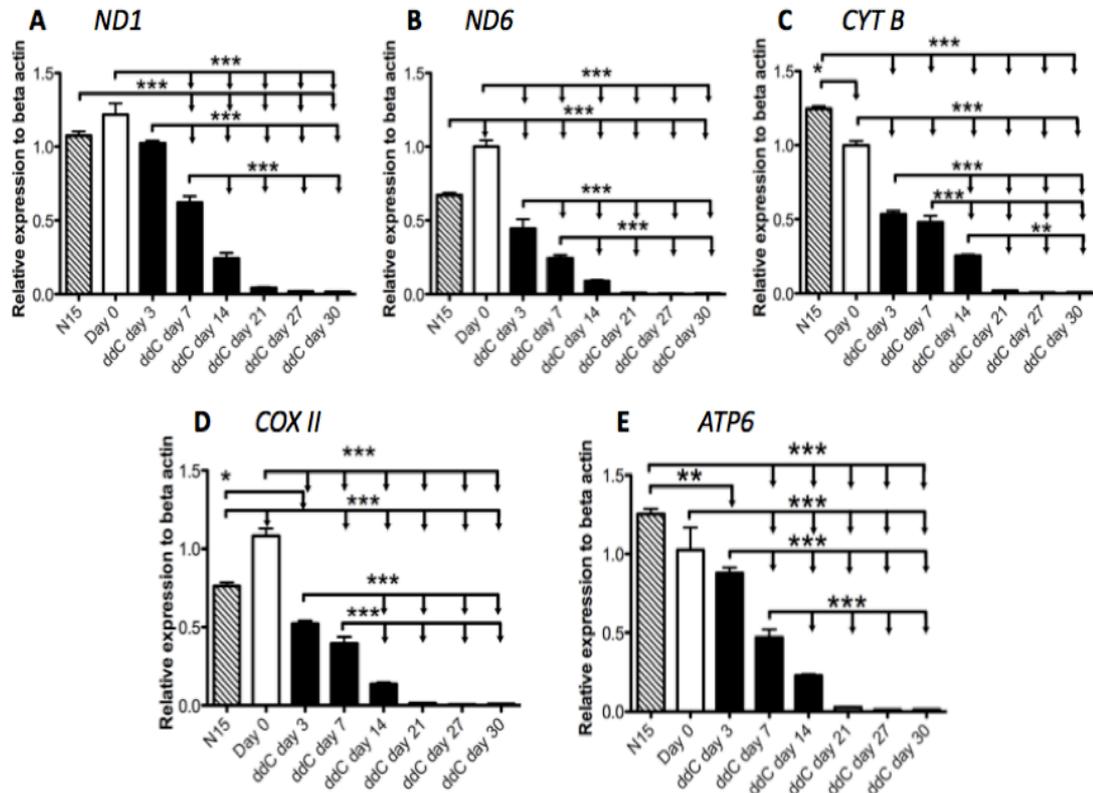
**Figure 4.6.** *MtDNA depletion of U266 cells for 30 days using ddC (A) Changes in the pattern of mtDNA copy numbers per cell are measured during mtDNA depletion (black). (B) Cell population doubling times during mtDNA depletion (blue) of the U266 cells relative to the untreated control (red). \* $P = \leq 0.05$ , \*\*\* $P = \leq 0.001$ .*

#### 4.5.4. Expression of the mtDNA genes during depletion of U266 cells

After observing changes to mtDNA copy number during mtDNA depletion, the expression of the mtDNA-encoded genes was analysed. These gene products are responsible for maintaining the function of the ETC.

The progressive loss of mtDNA copy number induced by ddC correlated with the gradual decline of expression for the mtDNA-encoded genes. Amongst the genes assessed, there were differences in the rates at which the expression levels decreased. Both *ND1* (Figure 4.7A) and *ATP6* (Figure 4.7E) decreased their relative expression levels at a slower rate compared to *ND6* (Figure 4.7B), *CYTB* (Figure 4.7C) and *COX II* (Figure 4.7D), which showed a more rapid decrease in their relative expression levels during the early stages of mtDNA depletion

(between days 0 and 3). For all genes, relative expression levels were almost negligible from day 21 of mtDNA depletion onwards.



**Figure 4.7.** Changes to the relative expression of the mtDNA-encoded genes during depletion. Complex I (A) ND1 and (B) ND6, Complex III (C) CYT B, Complex IV (D) COX II and Complex V (E) ATP6. The N15 sample (grey) represents a blood plasma non-tumor control for U266 cells (white). \* $P = \leq 0.05$ ; \*\*  $P = \leq 0.01$ ; \*\*\*  $P = \leq 0.001$ .

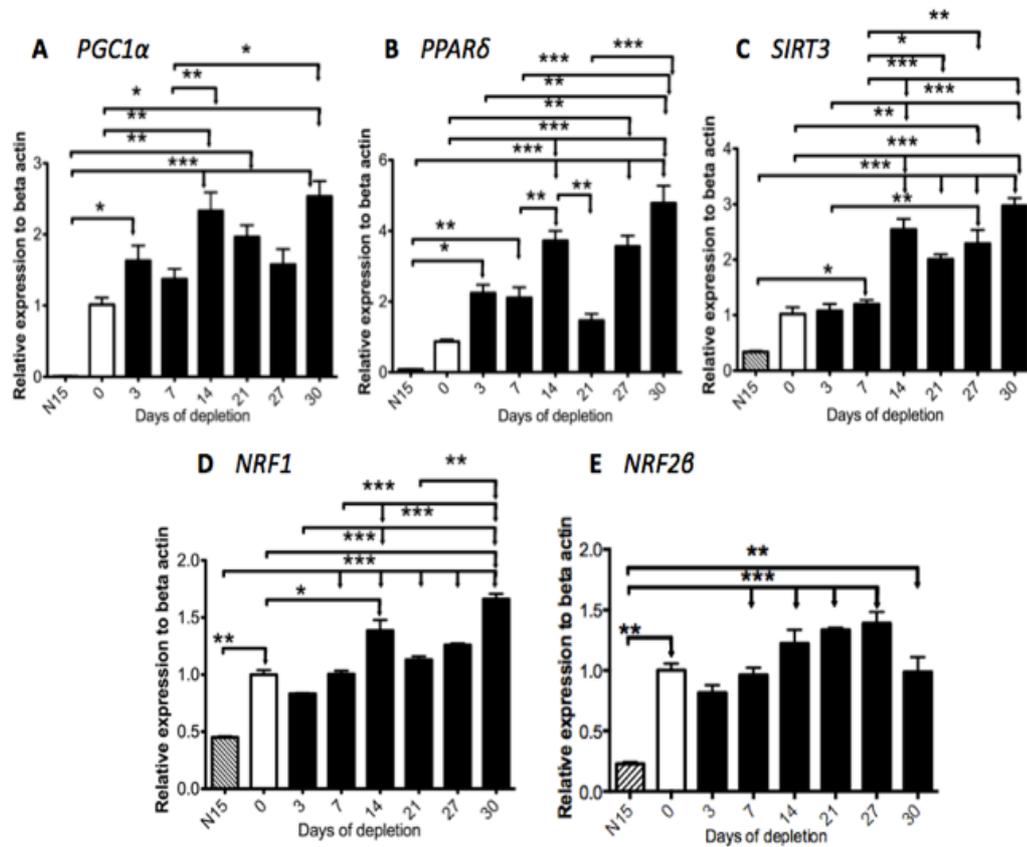
Comparisons between non-treated U266 cells and the blood plasma N15 sample, which represents the non-tumorigenic counterpart of U266 cells, showed variable patterns of expression for all mtDNA-encoded genes examined. Normally, one would expect the U266 cells to have decreased expression of OXPHOS genes. This is because the activity of the ETC protein complexes have been shown to be

down-regulated in other cancer types, such as in renal carcinoma cells when compared to their normal counterparts<sup>278</sup>. In the case of renal carcinoma, a correlation has even been determined between low OXPHOS activity and the stage of tumor progression<sup>279</sup>. However, gene expression results showed that levels of *CYT B* and *ATP6* were both higher in the blood plasma sample than in U266 day 0 cells, which emphasizes differences between cancers of different origin.

#### **4.5.5. Expression changes to the mitochondrial transcription and replication genes**

The expression of the nuclear genes involved with the transcription and replication of mtDNA were examined during mtDNA depletion. Throughout the course of depletion, relative expression of the genes involved with mitochondrial biogenesis (*PGC1 $\alpha$* , *PPAR $\delta$*  and *SIRT3*, **Figures 4.8A-C**) were maintained above levels of day 0 cells at all times. For each of these genes of mitochondrial biogenesis, their relative expression levels increased by at least 2 fold by day 30 of mtDNA depletion compared to day 0 cells. *PPAR $\delta$* , *SIRT3* and *NRF1* (**Figure 4.8D**) expression levels displayed a two-stage increase in relative levels during mtDNA depletion, which occurred from day 3 to day 14, and from day 21 to day 30.

The fall in relative *PPAR $\delta$*  expression at day 21 of mtDNA depletion corresponded with the gradual decline in *PGC1 $\alpha$*  expression from day 14 to 27 of mtDNA depletion. This change in expression pattern supports the role of *PPAR $\delta$*  regulating the expression of *PGC1 $\alpha$* . The gene expression profile for



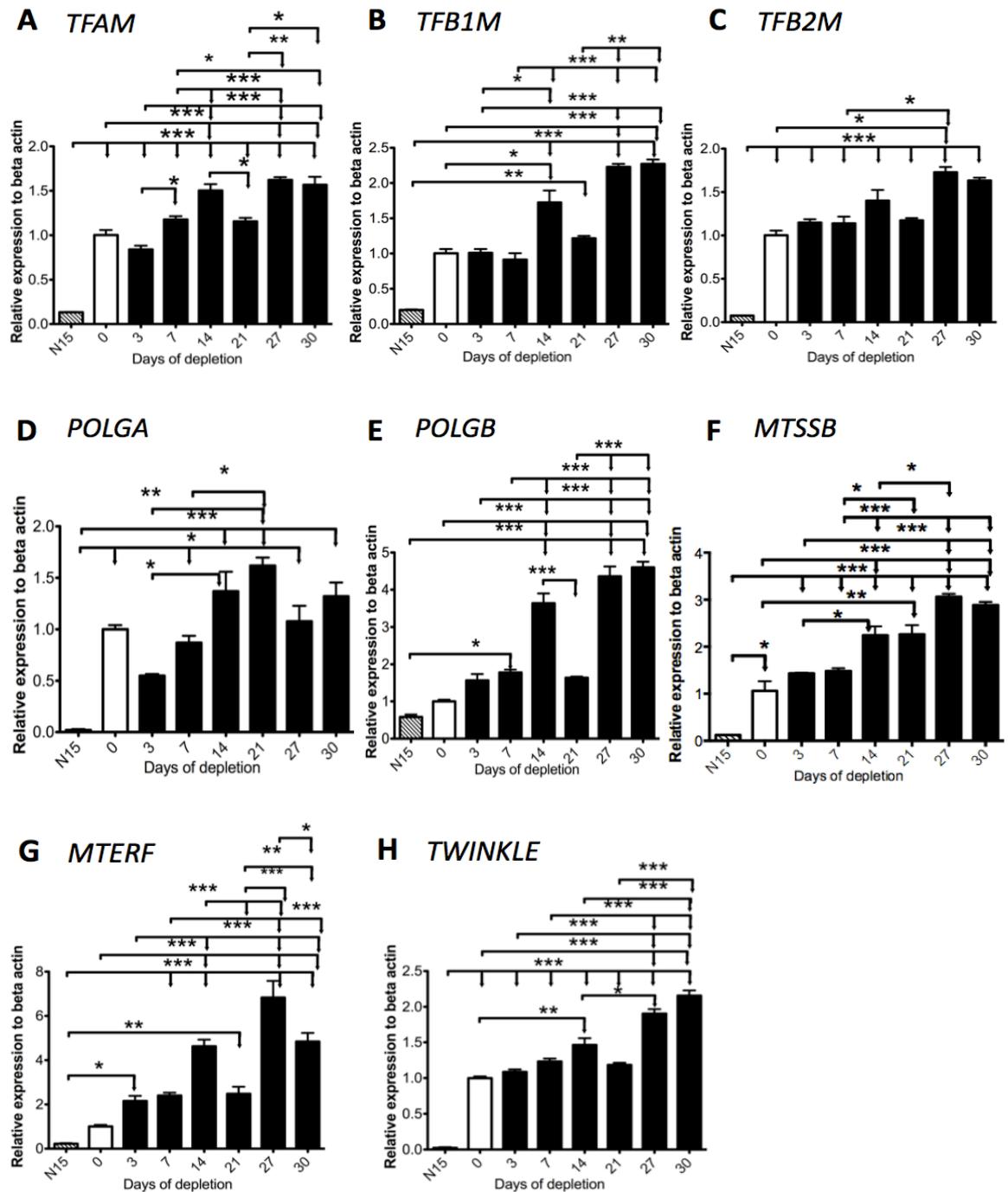
**Figure 4.8.** Gene expression levels for the upstream mitochondrial regulators of mitochondrial biogenesis factors (A) *PGC1α* (B) *PPARδ* and (C) *SIRT3*; and those that control mitochondrial transcription (D) *NRF1* and (E) *NRF2β*. \* $P = \leq 0.05$ ; \*\*  $P = \leq 0.01$ ; \*\*\*  $P = \leq 0.001$ .

*SIRT3* from day 21 of mtDNA depletion revealed the opposite trend to *PGC1α* levels. It is possible that *SIRT3*, as a downstream target of *PGC1α*, could have up-regulated its gene expression levels to compensate for the down-regulation of *PGC1α* levels and to maintain mitochondrial biogenesis by this late stage of mtDNA depletion. Both *NRF1* and *NRF2β* decreased their relative gene expression levels during the early stages of mtDNA depletion, which gradually increased above levels of day 0 cells after prolonged mtDNA depletion. All the mitochondrial biogenesis and nuclear respiratory factor genes demonstrated

higher expression levels in U266 day 0 cells compared to the N15 blood plasma sample.

All mitochondrial transcription and replication genes experienced an up-regulation in their expression levels by the late stages of mtDNA depletion (**Figure 4.9**). Although, the expression patterns during mtDNA depletion varied between genes, several displayed common trends. For *TFAM*, *TFB1M* (**Figure 4.9B**) and *TFB2M* (**Figure 4.9C**) expression levels were similar from day 14 to day 30 of mtDNA depletion. The similarity in pattern could be due to TFAM directing regulating TFB1M and TFB2M during the initiation of mtDNA replication<sup>280</sup>.

Perhaps the most important amongst all the genes analysed are the transcription factors *TFAM*, *TFB1M* and *TFB2M* (**Figure 4.9C**), since the expression of these could influence the expression of other genes examined. For example, the decrease in the expression levels of *TFAM*, *TFB1M* and *TFB2M* at day 21 of mtDNA depletion, could be responsible for the decrease in expression levels at day 21 of depletion for *PGC1 $\alpha$*  (**Figure 4.8A**), *PPAR $\delta$*  (**Figure 4.8B**), *SIRT3* (**Figure 4.8C**), *NRF1* (**Figure 4.8D**), *POLGB* (**Figure 4.9E**), *MTERF* (**Figure 4.9G**) and *TWINKLE* (**Figure 4.9H**).



**Figure 4.9.** Expression of the mitochondrial transcription and replication factors. Mitochondrial regulators of transcription (A) TFAM, (B) TFB1M and (C) TFB2M, and the mitochondrial replication genes (D) POLGA, (E) POLGB, (F) MTSSB, (G) MTERF and (H) TWINKLE were examined during mtDNA depletion in U266 cells. \* $P = \leq 0.05$ ; \*\*  $P = \leq 0.01$ ; \*\*\*  $P = \leq 0.001$ .

The highest fold-change increase in gene expression levels was reported for *MTERF*, which exhibited over a 6-fold increase in expression levels at day 27 of depletion compared to day 0 cells. *POLGB* had the second highest fold-change difference detected at day 30 of mtDNA depletion, which was 4.5 fold-higher than day 0 cells. Both *POLGA* and *TFAM* experienced an initial decline in relative gene expression at day 3 of depletion, which was followed by up-regulation in expression levels until days 14 or 21, respectively.

For a few genes including *TFB1M*, *TFB2M*, and *MTSSB* (**Figure 4.9F**), gene expression levels were similar at days 3 and 7 of mtDNA depletion. This could suggest a delay in the response of the transcriptional machinery to respond to the loss of mtDNA copies during depletion, since expression levels at day 3 and 7 were comparable to day 0 cells. Despite minor changes to expression levels of *TFB1M*, *POLGB* and *MTERF* during the early stages of mtDNA depletion, by day 14, relative expression of these genes almost doubled to compensate for the minimal changes to gene expression that occurred earlier during mtDNA depletion.

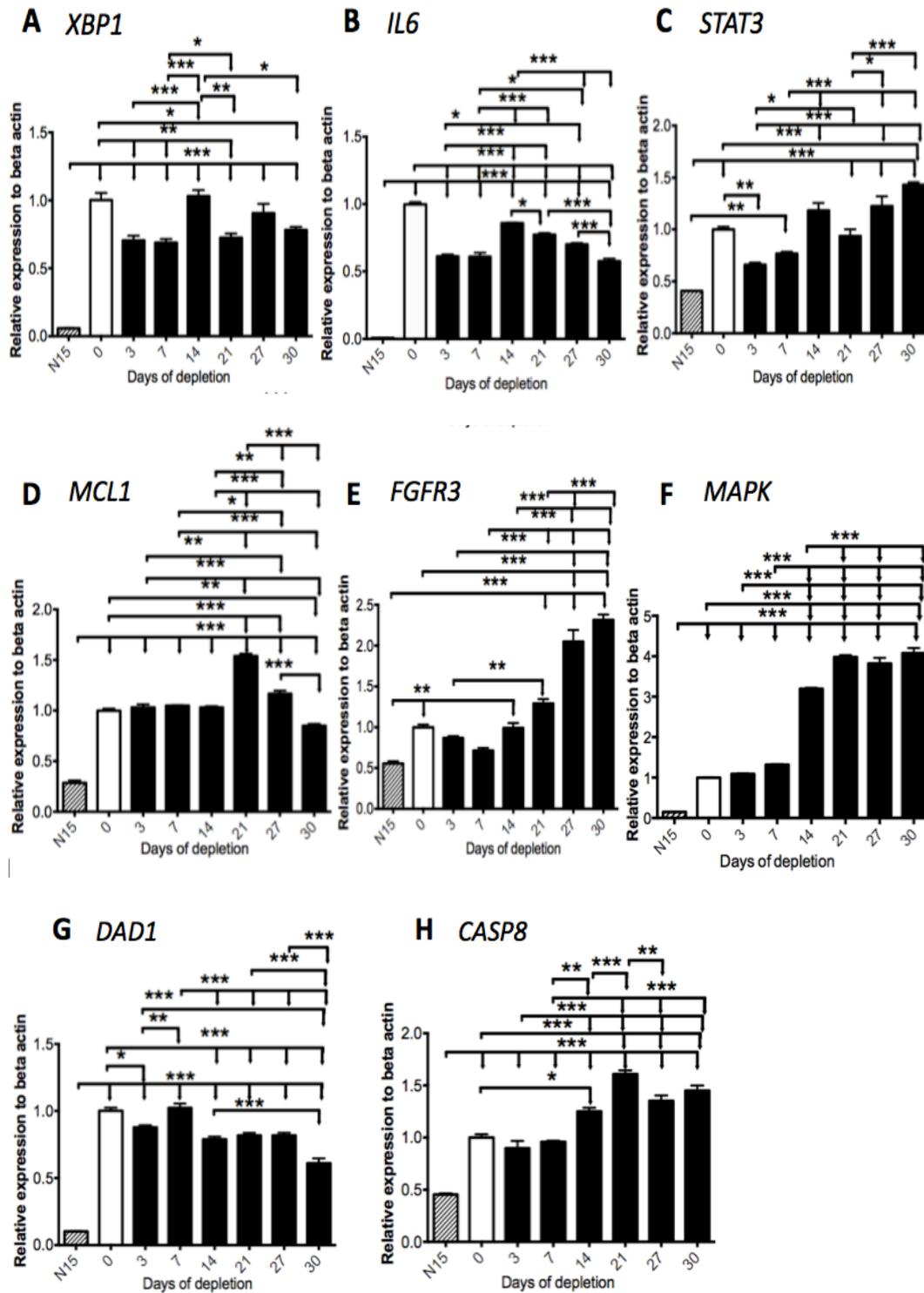
Furthermore, the gene expression profile for *MTSSB* showed increases to its relative expression levels throughout mtDNA depletion, which occurred in a stepwise manner (**Figure 4.9F**). For all the mitochondrial transcription and replication factors (**Figure 4.9**), expression levels in day 0 cells were detected at higher levels than for the non-tumorigenic blood plasma sample, especially for *POLGA* and *TWINKLE*, where expression in the plasma sample was almost negligible.

#### 4.5.6. Expression of the genes regulating multiple myeloma tumor biology

Tumor cells are highly reliant on energy production from glycolysis to maintain their ability to undergo rapid proliferation. Gene expression analysis was performed on selected genes associated with proliferation and apoptosis in multiple myeloma.

For a selection of tumor genes, including *XBPI* (**Figure 4.10A**), *IL6* (**Figure 4.10B**) and *DADI* (**Figure 4.10G**), relative gene expression did not increase above day 0 cells at any time during mtDNA depletion. Instead, gene levels remained above 50% of the gene expression levels for day 0 cells. Changes to gene expression patterns for the survival genes *DADI* and *MCLI* could indicate that it was necessary for the cells to maintain expression levels of these genes close to that of U266 control cells, to ensure cell survival throughout mtDNA depletion.

*XBPI*, *DADI*, *IL6* and also *MCLI* (**Figure 4.10D**) showed that during the later stages of mtDNA depletion, from day 14 or 21 onwards, there was a decrease in the levels of expression for these genes. On the other hand, for the remaining genes, *STAT3* (**Figure 4.10C**), *FGFR3* (**Figure 4.10E**), *MAPK* (**Figure 4.10F**) and *CASP8* (**Figure 4.10H**), expression levels increased during the late stages of mtDNA depletion. For these up-regulated genes, *MAPK* demonstrated the greatest increase with a fold change difference of at least 3 times the levels of day 0 cells by day 21 of mtDNA depletion. From day 21 onwards, *MAPK* expression reached a plateau where expression levels were likely to be at its maximal.

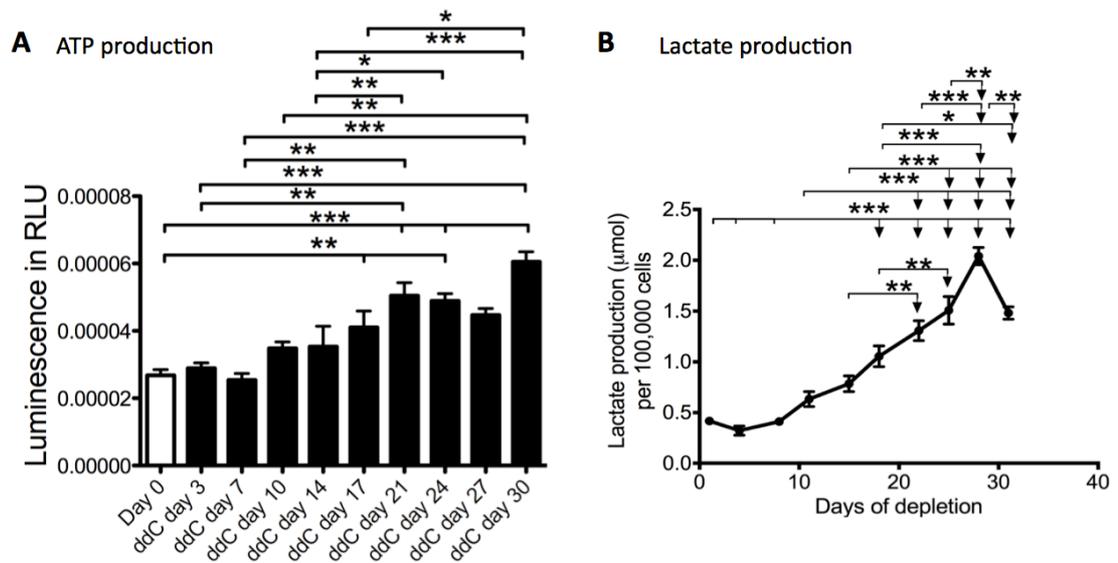


**Figure 4.10.** Gene expression changes to key tumorigenic genes during depletion. (A) XBP1 (B) IL6 (C) STAT3 and (D) MCL1 (E) FGFR3 (F) MAPK (G) DAD1 and (H) CASP8. \* $P = \leq 0.05$ ; \*\* $P = \leq 0.01$ ; \*\*\* $P = \leq 0.001$ .

For all the tumorigenic genes, the blood plasma sample exhibited lower gene expression levels compared to day 0 U266 cells. This observation supports the hypothesis that tumor cells require elevated expression levels of these genes to enable them to persist in the host environment. Moreover, lower expression levels in the blood plasma sample were expected, since non-tumor cells have reduced rates of proliferation compared to their tumorigenic counterparts. Finally, the 2-fold difference between the expression levels of *CASP8* in the blood plasma sample compared to day 0 cells was also consistent with a study performed on genetically identical individuals. In this study, *CASP8* expression was two-fold higher in the individual diagnosed with multiple myeloma, when compared to the healthy control<sup>281</sup>.

#### **4.5.7. ATP and lactate production during mtDNA depletion**

Next, the metabolic characteristics, in the form of ATP and lactate production in U266 cells, were examined in the context of mtDNA depletion. Total ATP production gradually increased in levels during mtDNA depletion (**Figure 4.11A**). However, a temporary decline in ATP production at days 24 and 27 was detected, which preceded peak levels at day 30 of depletion. Total lactate production also increased with mtDNA depletion until day 28 (**Figure 4.11B**), after which levels declined by day 31.

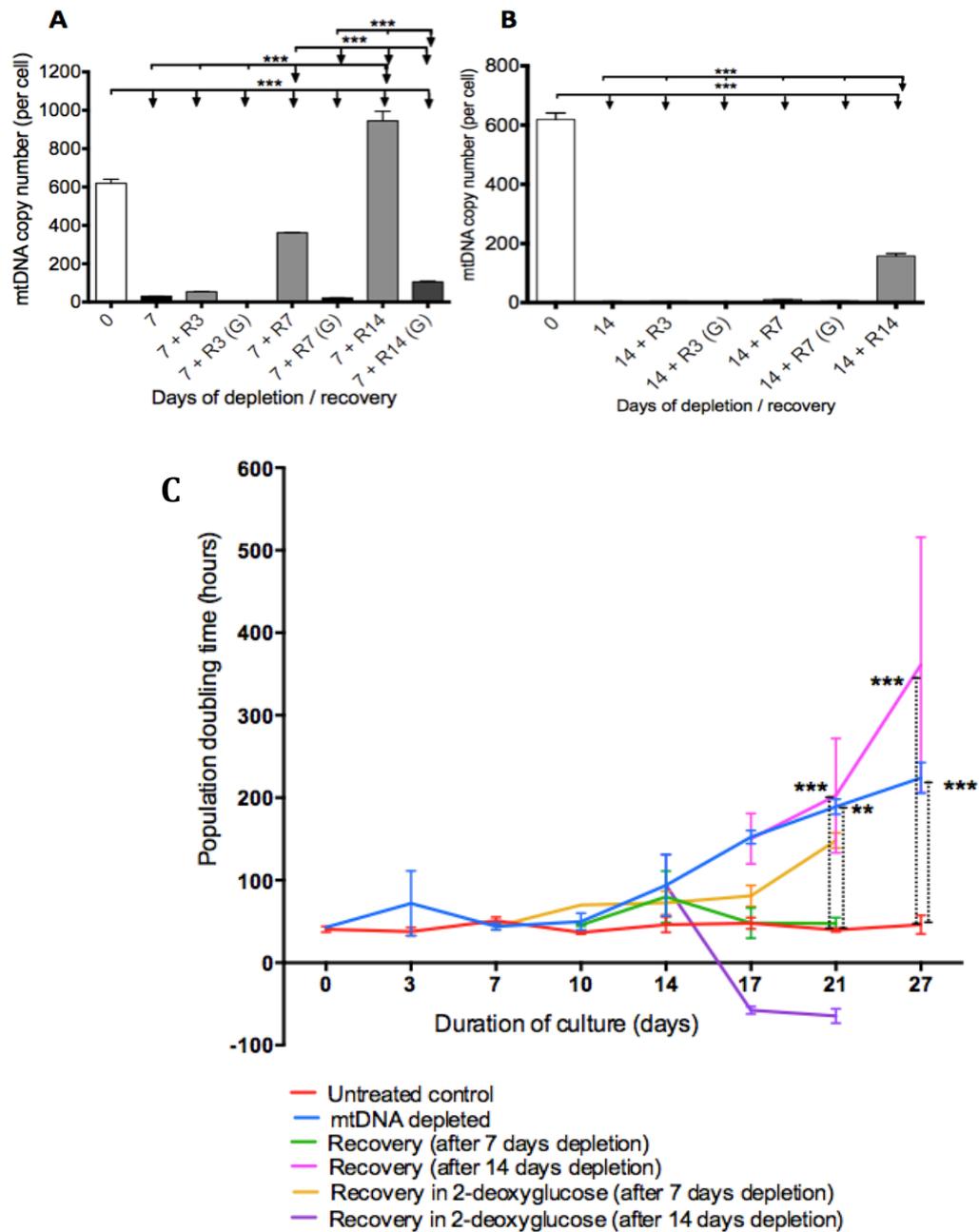


**Figure 4.11.** Changes to ATP (A) and lactate (B) production during mtDNA depletion. \* $P = \leq 0.05$ ; \*\* $P = \leq 0.01$ ; \*\*\* $P = \leq 0.001$ .

#### 4.5.8. Recovery of mtDNA copy number in the presence of glycolysis inhibitor.

Next, the ability to recover mtDNA copy number in depleted U266 cells was examined. This was done in the presence and absence of glycolysis inhibitor, 2-deoxyglucose (2-DG). From the respirometry results, it was previously determined that U266 cells primarily respire via glycolysis, even in the presence of some OXPHOS activity (Table 4.2). Therefore, the use of 2-DG would not only confirm these outcomes, but would also examine the extent to which these cells are reliant on glycolysis.

It was identified that 7 days after mtDNA depletion, U266 cells were able to recover their mtDNA copy numbers above day 0 cells by day 14 of recovery (Figure 4.12A). However, in the presence of 2-deoxyglucose (2-DG), mtDNA



**Figure 4.12.** Recovery of mtDNA copy numbers in U266 cells after mtDNA depletion, cultured in the presence and absence of 2-DG. (A) 7 days of mtDNA depletion, (B) 14 days of mtDNA depletion. ‘R’ represents the number of days of mtDNA replenishment, and (G) represents the addition of 2-DG in culture. (C) Population doubling times during depletion in the presence and absence of recovery, both with and without 2-DG. \*\*  $P = \leq 0.01$ ; \*\*\*  $P = \leq 0.001$ .

copy number was unable to replenish efficiently. For 14-day mtDNA depleted cells, recovery of mtDNA copy number was not noticeable until 14 days of culture in complete media (**Figure 4.12B**). It was not possible to analyse U266 cells cultured in 2-DG during recovery for 14 days (**Figure 4.12B**), as these cells were undergoing significant loss of cell numbers by day 7 of recovery (**Figure 4.12C**).

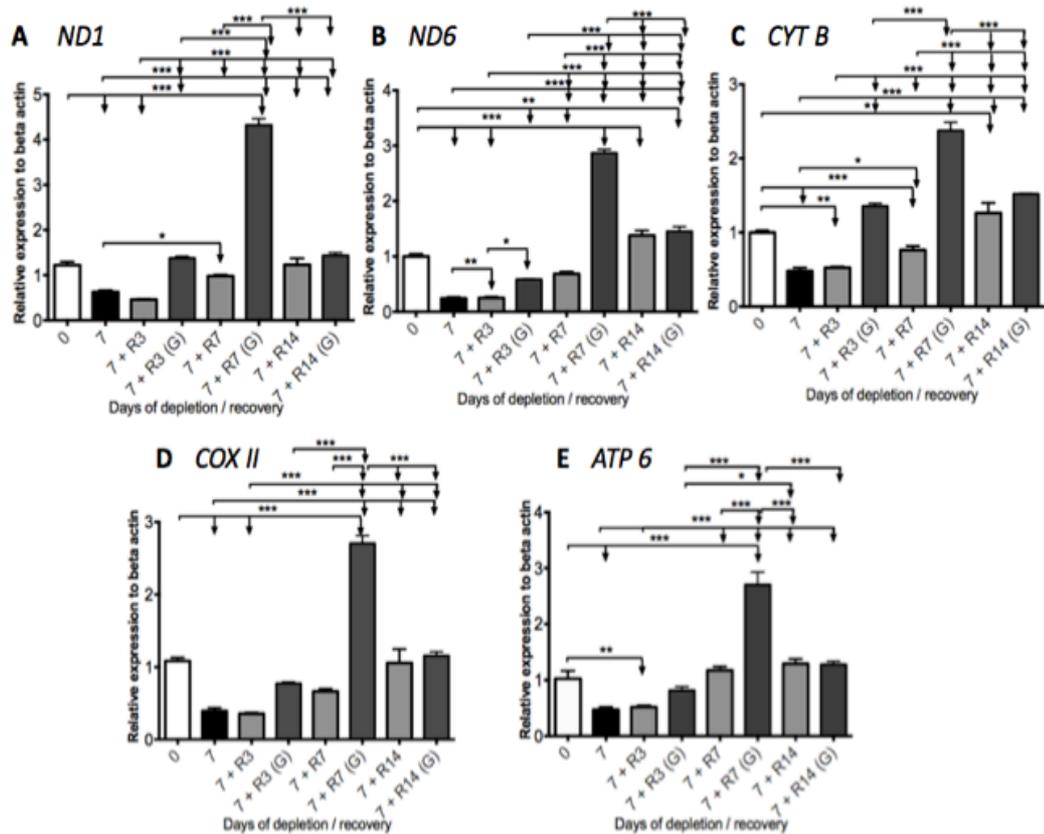
Population doubling times of cells depleted of their mtDNA for 7 days and recovered in the absence of 2-DG, were able to recover proliferation rates similar to the non-depleted cells (**Figure 4.12C**). However, this was not possible for the same cells recovered in the presence of 2-DG, which signified that there is a requirement for functional glycolysis in controlling the proliferation rates of these cells. Furthermore, cells that underwent prolonged mtDNA depletion for 14 days were not able to restore proliferation rates back to normal levels, even in the absence of 2-DG. This observation indicated the presence of irreversible damage at the mtDNA level induced by ddC in these cells, which not only corresponded to results from the copy number analysis (**Figure 4.12B**), but also determined that there was a role for mtDNA in controlling proliferation rates.

In the presence of 2-DG, recovery of 14-day mtDNA depleted cells showed that these cells were not able to proliferate but instead underwent loss of cell population numbers, as negative values represented the time (in hours) required for the cell population to halve rather than to double.

#### **4.5.9. Changes in expression to the mtDNA genes during recovery with 2-DG.**

Since ddC appeared to exert an influence on mtDNA copy numbers, which was an effect dependent upon the duration of cell culture, further experiments were performed to determine whether the effects of ddC were reversible in the presence and absence of 2-DG.

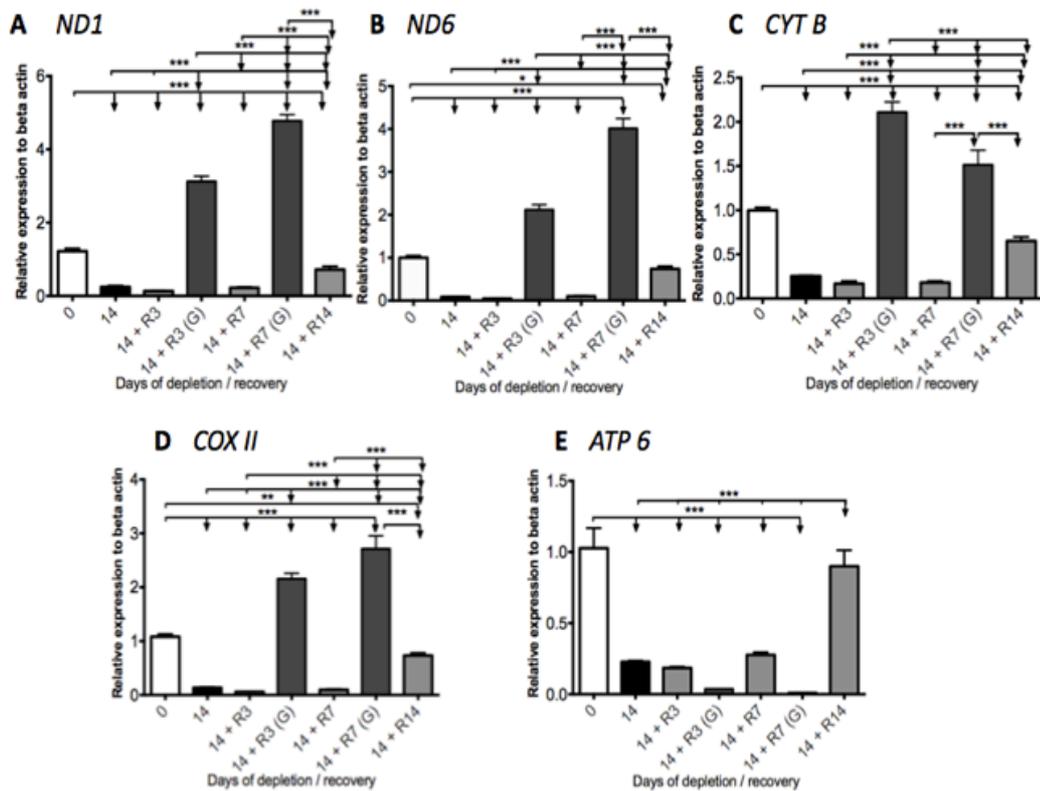
For 7-day mtDNA depleted cells recovered in 2-DG, there was increased expression of the mtDNA-encoded genes, relative to recovery without 2-DG (**Figure 4.13**). Most noticeably, expression levels increased significantly by day 7 of recovery for all genes examined in the presence of 2-DG. This could be due to increased activity of the fatty acid oxidation pathway<sup>282</sup>, resulting in the production of acetyl-coenzyme A (acetyl-CoA) and the generation of reducing factors via the citric acid cycle, for use in the ETC<sup>283</sup>. The sudden requirement for OXPHOS activity under the influence of 2-DG, could positively regulate the expression of these key mtDNA genes. Together the results demonstrate the presence of a compensatory mechanism that acts to restore or up-regulate activity of the ETC following mtDNA depletion, which is enhanced following the inhibition of glycolysis.



**Figure 4.13.** Changes in expression to the mtDNA genes during recovery with 2-DG, after 7 days of mtDNA depletion. Complex I genes (A) *ND1* and (B) *ND6*, Complex III gene (C) *CYT B*, Complex IV gene (D) *COX II* and Complex V gene (E) *ATP 6*. ‘R’ represents the number of days of mtDNA replenishment, and (G) represents the addition of 2-DG in culture. \* $P = \leq 0.05$ ; \*\*  $P = \leq 0.01$ ; \*\*\*  $P = \leq 0.001$ .

In the absence of 2-DG, recovery of the mtDNA gene expression levels after 14 days of depletion (Figure 4.14) occurred less efficiently than cells depleted for only 7 days (Figure 4.13). In the presence of 2-DG, changes to the relative expression levels of the mtDNA genes at day 7 of recovery after 14 days of depletion (Figure 4.14), were similar to the 7-day recovered cells that have been depleted for 7 days (Figure 4.13). Whereas, significant increases were observed to *ND1* (Figure 4.14A), *ND6* (Figure 4.14B) and *COX II* (Figure 4.14D)

expression levels compared to day 0 cells when undergoing recovery with 2-DG after prolonged (14 days) mtDNA depletion, this was not seen for *ATP6* (Figure 4.14E). While it is unlikely that the absence of *ATP6* expression during this time could be due to defects to the transcriptional machinery, it remains possible that there may be defects to the *ATP6* gene structure itself, which prevents the recovery, and therefore detection, of its gene levels. For example, *ATP6* may undergo rapid degradation due to insufficient post-transcriptional modifications that would otherwise stabilize the gene for detection during real-time PCR analysis.



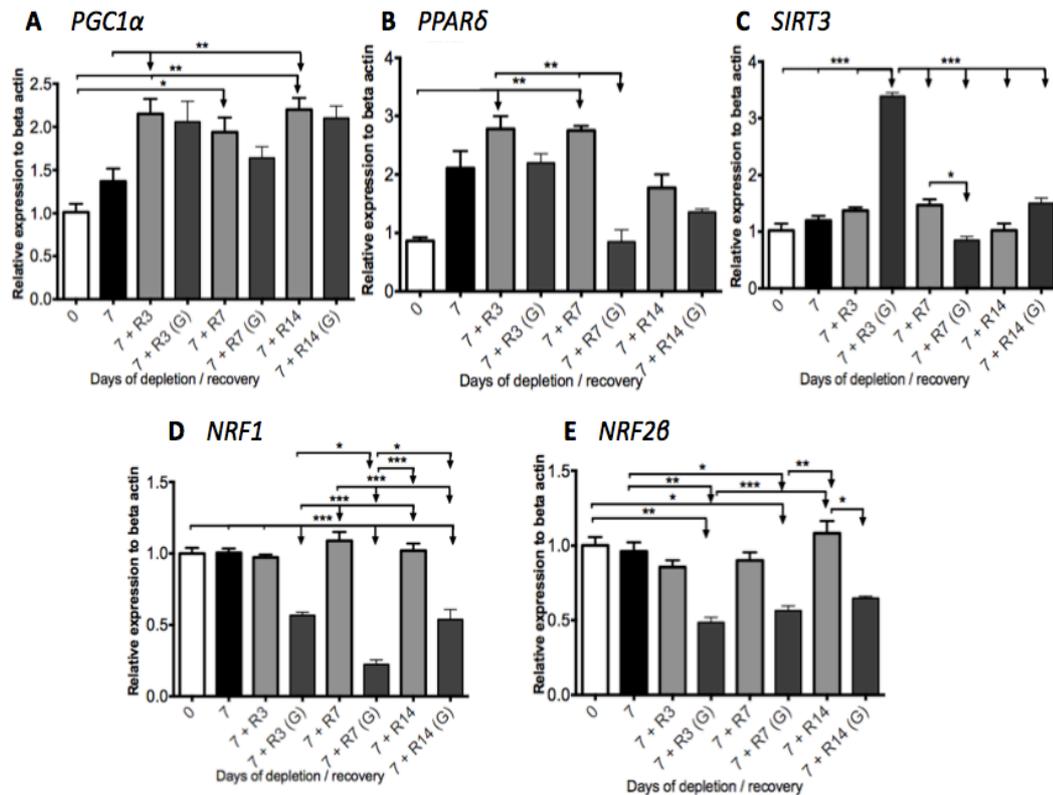
**Figure 4.14.** Expression changes to the mtDNA genes during recovery with 2-DG, after 14 days of mtDNA depletion. Complex I genes (A) *ND1* and (B) *ND6*, Complex III gene (C) *CYT B*, Complex IV gene (D) *COX II* and Complex V gene (E) *ATP 6*. 'R' represents recovery in days, and (G) represents the presence of 2-DG. \* $P = \leq 0.05$ ; \*\* $P = \leq 0.01$ ; \*\*\* $P = \leq 0.001$ .

From these results, it can be determined that overall, in the absence of glycolysis there is an increased need for the cell to up-regulate its transcriptional machinery to enhance the expression of mtDNA encoded genes. Aside from the anomalous result for *ATP6* (**Figure 4.14E**), the general trends in the mtDNA gene expression patterns during recovery in the presence of 2-DG, represent a cellular attempt to switch its primary mode of metabolism from glycolysis to OXPHOS, most likely to maintain cell viability. Both these pathways are essential for energy production, and the lack of both would be detrimental to many aspects of cellular function and to many of the processes involved.

#### **4.5.10. Expression of the upstream mitochondrial regulators during recovery with 2-DG.**

Upon examining the changes in expression levels for upstream regulators involved with mitochondrial biogenesis, several key observations were made. Firstly, in the presence of 2-DG, genes *PGC1 $\alpha$* , *PPAR $\delta$* , *NRF1* and *NRF2 $\beta$*  all decreased their expression levels during recovery after 7 days of mtDNA depletion, relative to cells recovered without 2-DG (**Figure 4.15A, 4.15B, 4.15D and 4.15E**). This suggested that the recovery of expression levels for these genes was more efficient in the absence of 2-DG.

*PGC1 $\alpha$*  was the only mitochondrial biogenesis gene whose levels remained above day 0 cells at all times throughout mtDNA replenishment. The changes in expression of *SIRT3* during recovery with 2-DG was also more variable than recovery without the inhibitor (**Figure 4.15C**). The two-fold increase in *SIRT3* expression at day 3 of replenishment with 2-DG, could indicate an attempt to

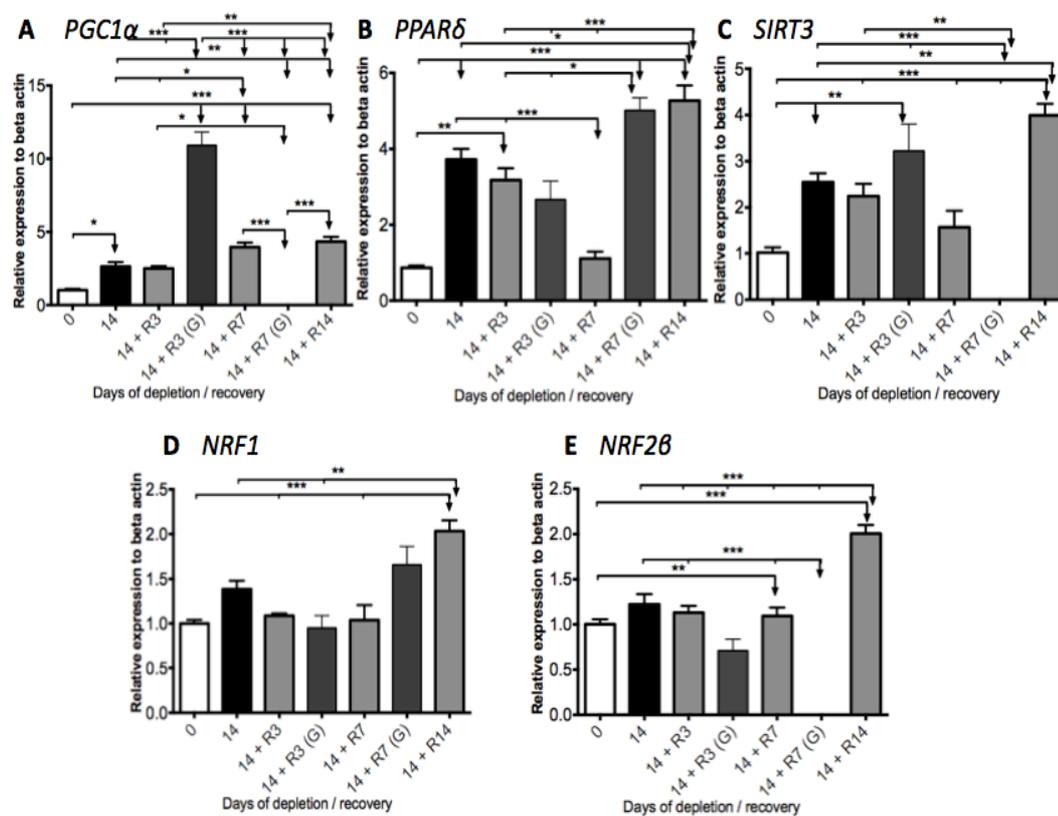


**Figure 4.15.** Expression of the upstream mitochondrial regulators during recovery with 2-DG, after 7 days of depletion. Mitochondrial biogenesis genes (A) *PGC1α*, (B) *PPARδ* and (C) *SIRT3*. Nuclear respiratory factors (D) *NRF1* and (E) *NRF2β*. ‘R’ represents recovery in days, and (G) represents the presence of 2-DG. \* $P = \leq 0.05$ ; \*\*  $P = \leq 0.01$ ; \*\*\*  $P = \leq 0.001$ .

maintain mitochondrial biogenesis since *SIRT3* mediates the effects of *PGC1α*. In the absence of inhibitor, *SIRT3* expression during mtDNA replenishment was close to levels of 7-day mtDNA depleted cells, which can also be observed for both *NRF1* and *NRF2β* at all stages of mtDNA replenishment (Figures 4.15D and 4.15E).

A comparison of gene expression changes after 7 and 14 days of mtDNA depletion revealed that the mitochondrial biogenesis genes had more variable

changes during recovery after 14 days of depletion, which suggests that by this stage there could be irreversible damage to mtDNA. Cells recovered in the absence of 2-DG showed a decline in *PPAR $\delta$*  (Figure 4.16B), *SIRT3* (Figure 4.16C), *NRF1* (Figure 4.16D) and *NRF2 $\beta$*  (Figure 4.16E) levels below that of 14-day mtDNA depleted cells, until day 7 of recovery. These recovered cells also had similar *NRF1* and *NRF2 $\beta$*  expression profiles during the course of recovery.



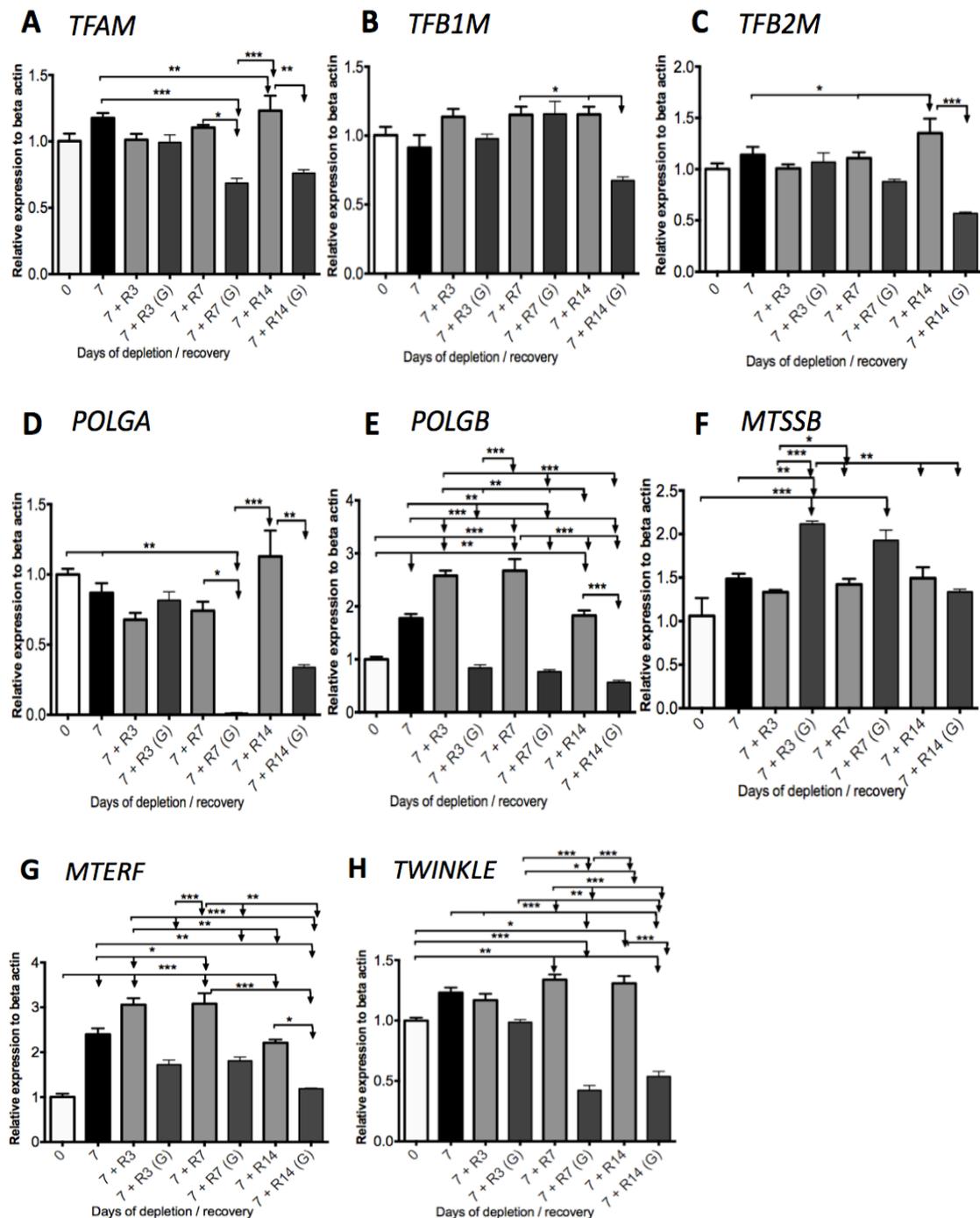
**Figure 4.16.** Expression of the upstream mitochondrial regulators during recovery with 2-DG, after 14 days of depletion. Mitochondrial biogenesis genes (A) *PGC1 $\alpha$* , (B) *PPAR $\delta$*  and (C) *SIRT3*. Nuclear respiratory factors (D) *NRF1* and (E) *NRF2 $\beta$* . ‘R’ represents recovery in days, and (G) represents the presence of 2-DG. \* $P = \leq 0.05$ ; \*\* $P = \leq 0.01$ ; \*\*\* $P = \leq 0.001$ .

However, cells cultured in the presence of 2-DG failed to up-regulate *NRF2 $\beta$*  expression by day 7 of mtDNA replenishment, whereas for *NRF1*, expression levels increased by this stage of recovery.

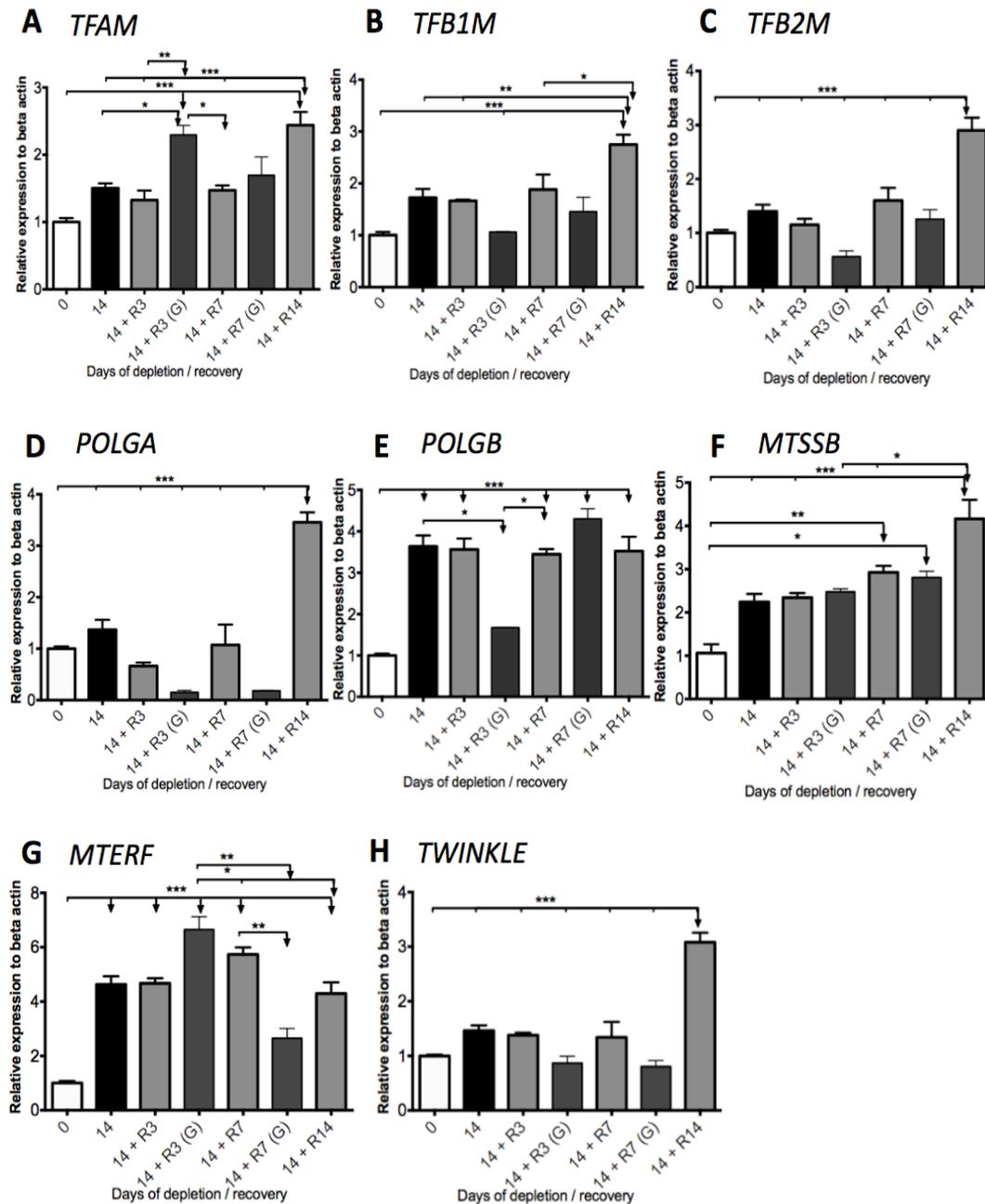
#### **4.5.11. Expression of the mitochondrial transcription and replication genes during recovery with 2-DG**

For cells recovered in the presence of 2-DG, gene expression levels of *POLGB* (**Figure 4.17E**), *MTERF* (**Figure 4.17G**) and *TWINKLE* (**Figure 4.17H**) were lower than for cells cultured in the absence of the inhibitor at all stages of mtDNA replenishment. For *POLGB*, expression in the presence of 2-DG remained low throughout recovery, at levels below 50% of 7-day mtDNA depleted cells. Expression levels of *TFAM*, *TFB1M* and *TFB2M* (**Figures 4.17A-C**) during mtDNA recovery in the absence of 2-DG were close to 7-day mtDNA depleted and day 0 cells throughout recovery. This expression pattern was also observed for *NRF1* and *NRF2 $\beta$*  (**Figures 4.15D and 4.15E**), both products of which control the transcription of *TFAM*, *TFB1M* and *TFB2M*.

*MTSSB* was the only mitochondrial replication gene whose expression in the presence of 2-DG increased above cells recovered in the absence of 2-DG at days 3 and 7 of recovery (**Figure 4.17F**). This change in *MTSSB* expression levels could indicate an attempt to increase efficiency of mitochondrial replication during recovery in 2-DG. However, without similar expression increases for the other mitochondrial replication genes, mtDNA copy number did not undergo efficient mtDNA replenishment as for cells cultured in the absence of glycolysis inhibitor (**Figure 4.12A**).



**Figure 4.17.** Expression of the mitochondrial transcription and replication genes during recovery with 2-DG, after 7 days of mtDNA depletion. Regulators of transcription (A) *TFAM*, (B) *TFB1M*, (C) *TFB2M*, and regulators of mtDNA replication (D) *POLGA*, (E) *POLGB*, (F) *MTSSB*, (G) *MTERF* and (H) *TWINKLE*. ‘R’ represents recovery in days, and (G) represents the presence of 2-DG. \* $P = \leq 0.05$ ; \*\*  $P = \leq 0.01$ ; \*\*\*  $P = \leq 0.001$ .



**Figure 4.18.** Expression of the mitochondrial transcription and replication genes during recovery with 2-DG, after 14 days of mtDNA depletion. Regulators of transcription (A) *TFAM*, (B) *TFB1M*, (C) *TFB2M*, and regulators of mtDNA replication (D) *POLGA*, (E) *POLGB*, (F) *MTSSB*, (G) *MTERF* and (H) *TWINKLE*. ‘R’ represents recovery in days, and (G) represents the presence of 2-DG. \* $P = \leq 0.05$ ; \*\* $P = \leq 0.01$ ; \*\*\* $P = \leq 0.001$ .

Gene expression levels for 14-day mtDNA depleted cells recovered for 14 days in the absence of 2-DG were often double the levels detected at day 3 of recovery under the same culture conditions. This can be seen for *NRF1*, *NRF2 $\beta$*  (**Figures 4.16D-E**), *TFAM*, *TFB1M* (**Figures 4.18A-B**) and *MTSSB* (**Figure 4.18F**). Occasionally, there was a greater than two- fold increase between cells recovered for 3 days and those recovered for 14 days in the absence of 2-DG after 14 days of mtDNA depletion. This was observed for *TFB2M* (**Figure 4.18C**), *POLGA* (**Figure 4.18D**) and *TWINKLE* (**Figure 4.18H**).

During mtDNA replenishment there was a significant up-regulation in the expression levels of *POLGA*, which was suppressed in the presence of 2-DG. Expression levels for the supporting subunit, *POLGB* (**Figure 4.18E**), showed little change during the course of mtDNA replenishment for cells cultured in the absence of 2-DG. Furthermore, gene expression patterns for *MTSSB* during recovery in the absence of 2-DG after 14 days of mtDNA depletion, demonstrated an increasing trend with the duration of mtDNA replenishment (**Figure 4.18F**), which was more apparent than for cells depleted of their mtDNA for 7 days in culture (**Figure 4.17F**).

Cells depleted of their mtDNA for 14 days and recovered for the same length of time, without 2-DG, had *TFAM*, *TFB1M* and *TFB2M* (**Figures 4.18A-C**) gene expression profiles that were similar to *NRF1* and *NRF2 $\beta$*  (**Figures 4.16D-E**). This was similar to the situation described previously for 7-day mtDNA depleted cells during recovery without the inhibitor, which supports the hypothesis that

these changes in gene expression patterns could be due to NRF1 and NRF2 $\beta$  regulating the transcription of these mitochondrial transcription factors.

For most of the mitochondrial transcription and replication genes (*TFB1M*, *TFB2M*, *POLGB* and *MTSSB*), cells cultured in the presence of 2-DG demonstrated an increase in gene expression levels with the duration of mtDNA replenishment. A few genes, *TFAM* and *MTERF*, showed decreases in their expression levels during mtDNA replenishment, and others revealed no change during this recovery period, *POLGA* and *TWINKLE*.

From these observations, it can be proposed that glycolysis plays a major role in the energy metabolism of U266 cells. OXPHOS is also likely to have a role in maintaining survival of these tumor cells. This is because inhibition of glycolysis during mtDNA replenishment resulted in increased levels of expression for the mtDNA-encoded genes, which suggests that for the survival of these tumor cells it is necessary to ensure sufficient OXPHOS occurs, despite the presence of dysfunctional glycolysis. However, overall U266 cells have a greater reliance on glycolysis, since cells depleted of their mtDNA for 14 days and recovered in the presence of 2-DG underwent a reduction in cell population numbers (**Figure 4.12C**).

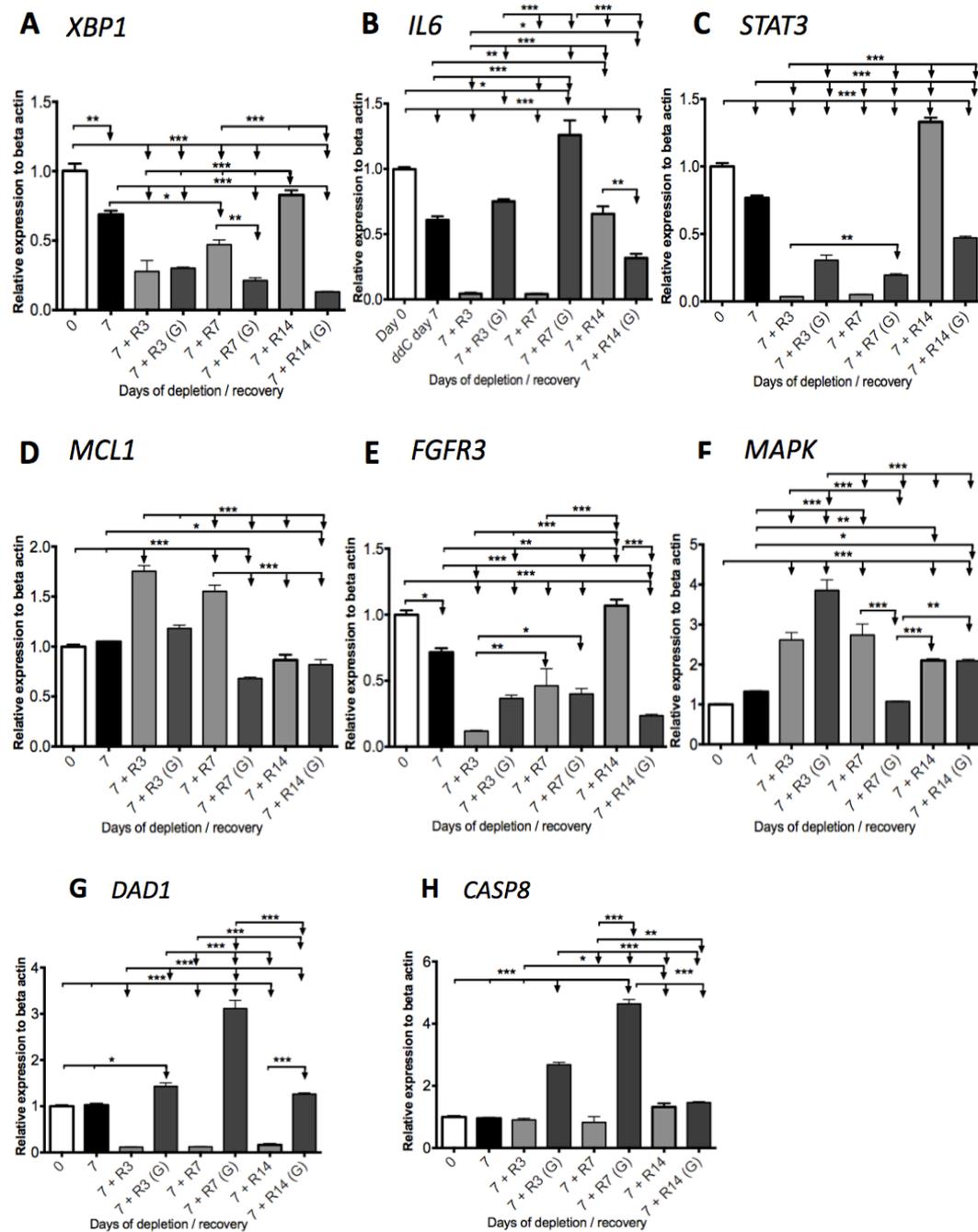
#### **4.5.12. Expression of the multiple myeloma tumor genes during recovery with 2-DG**

Recovery of mtDNA copy numbers after 7 days of mtDNA depletion in cells recovered in the absence of 2-DG led to early up-regulation of some genes

including *MCL1* (**Figure 4.19D**) and *MAPK* (**Figure 4.19F**). Whilst these observations suggested that there is an immediate effect upon mtDNA replenishment to increase tumorigenicity of U266 cells, other genes *XBPI* (**Figure 4.19A**), *IL6* (**Figure 4.19B**), *STAT3* (**Figure 4.19C**) and *FGFR3* (**Figure 4.19E**) revealed up-regulation in their expression levels by the late stages of mtDNA replenishment.

The expression profiles for *IL6* and *STAT3* in cells recovered in the absence of 2-DG, shared very low expression levels until day 7 of mtDNA replenishment, after which there was a significant rise in their levels of gene expression. This effect was more pronounced for *STAT3* than for *IL6*. For *CASP8* little change in expression levels was observed relative to 7-day mtDNA depleted cells (**Figure 4.19H**), whereas for *DADI* there was the failure to up-regulate expression levels throughout mtDNA replenishment (**Figure 4.19G**).

Cells recovered in the presence of 2-DG had variable patterns of expression. *XBPI*, *STAT3* and *FGFR3* had low levels of expression throughout mtDNA replenishment, below 50% of 7-day mtDNA depleted cells. This suggests that suppression of glucose metabolism impaired transcription of these tumor genes. However, in the short-term, the tumorigenic potential of U266 cells recovered in the presence of 2-DG was enhanced by increased transcription of *MAPK*, whose product participates in cellular proliferation. *IL6*, *DADI* and *CASP8* had similar trends in expression levels during recovery after 7 days of mtDNA depletion, with peak levels observed at day 7 of recovery in the presence of 2-DG.

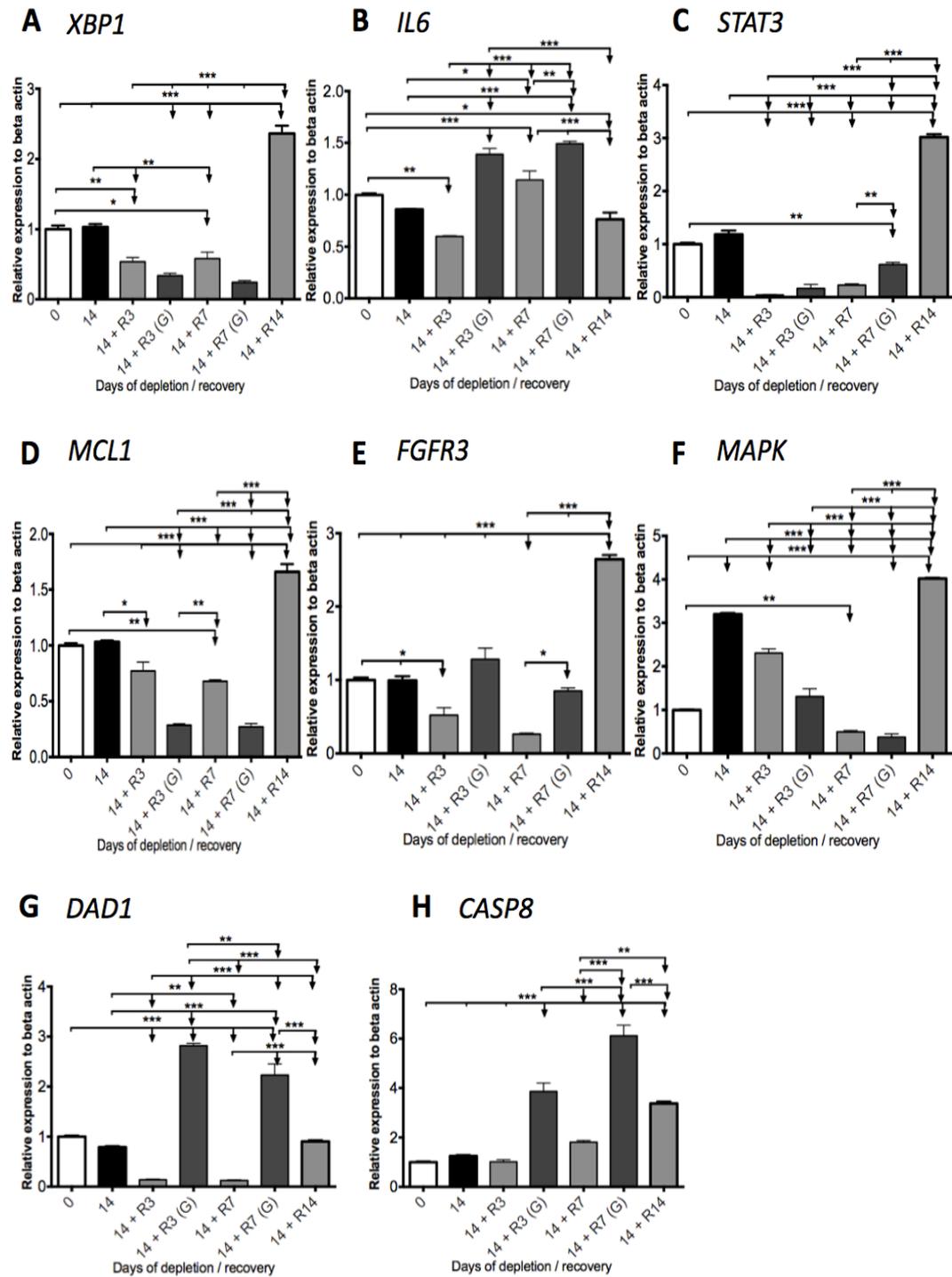


**Figure 4.19.** Expression of the multiple myeloma tumorigenic genes during recovery with 2-DG, after 7 days of mtDNA depletion. (A) XBP1, (B) IL6, (C) STAT3, (D) MCL1, (E) FGFR3, (F) MAPK, (G) DAD1 and (H) CASP8. 'R' represents recovery in days, and (G) represents the presence of 2-DG. \* $P \leq 0.05$ ; \*\*  $P \leq 0.01$ ; \*\*\*  $P \leq 0.001$ .

Altogether, these increases in gene expression were consistent with the hypothesis that inhibition of glycolysis increases tumorigenic properties of U266 cells. Differences between *XBPI* and *IL6* expression profiles for cells cultured with 2-DG, may reflect the need to up-regulate *IL6* expression in situations where *XBPI* expression is low to compensate for the reduced stimulation from *XBPI* on the *IL6* promoter region. This could also be true for the gene expression changes between *STAT3* and its downstream target *MCLI*, in cells recovered in the presence of the glycolysis inhibitor.

Expression levels of *XBPI* (**Figure 4.20A**), *DAD1* (**Figure 4.20G**) and *CASP8* (**Figure 4.20H**) during recovery in the absence of 2-DG, after 14 days of mtDNA depletion, were similar to the gene expression changes observed for *XBPI* during recovery after 7 days of mtDNA depletion (**Figure 4.19A**). Moreover, gene expression levels for *IL6* during mtDNA replenishment after 14-days of mtDNA depletion shared the observation with 7-day mtDNA depleted cells that, in the presence of 2-DG, there were higher levels of *IL6* compared to cells recovered in the absence of 2-DG.

During recovery in the presence of 2-DG, expression levels of *STAT3* (**Figure 4.20C**), *MCLI* (**Figure 4.20D**) and *MAPK* (**Figure 4.20F**) after 14 days of mtDNA depletion, was compromised more than at day 7 days of depletion (**Figure 4.20C**). Moreover, both the expression levels of *MCLI* and *MAPK* were lower in cells that have been recovered in 2-DG, compared to those that were recovered without 2-DG at all stages of mtDNA replenishment.



**Figure 4.20.** Expression of the multiple myeloma tumorigenic genes during recovery with 2-DG, after 14 days of mtDNA depletion. (A) XBP1, (B) IL6, (C) STAT3, (D) MCL1, (E) FGFR3, (F) MAPK, (G) DAD1 and (H) CASP8. 'R' represents recovery in days, and (G) represents the presence of 2-DG. \* $P \leq 0.05$ ; \*\*  $P \leq 0.01$ ; \*\*\*  $P \leq 0.001$ .

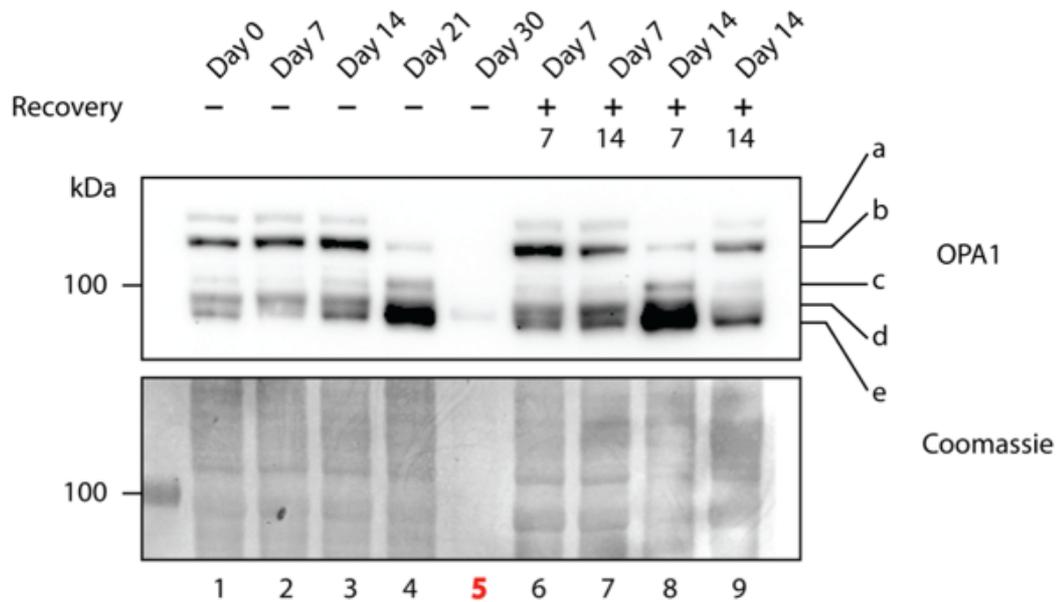
Altogether, 14 days of mtDNA replenishment after 14 days of depletion resulted in the increased ability to recover expression levels for the genes involved with tumor cell survival (*XBPI*, *STAT3*, *MCL1*, *DAD1* and *CASP8*), and cellular proliferation (*FGFR3* and *MAPK*). These observations emphasize that the tumorigenic characteristics of U266 cells are likely to be enhanced following prolonged depletion of their mtDNA (for 14 days) and 14 days of mtDNA recovery, when cultured in the absence of 2-DG.

#### **4.5.13. Determining the levels of apoptosis in mtDNA depleted and/or recovered U266 cells**

OPA1 is a nuclear-encoded mitochondrial protein that is localized to the inner membrane of mitochondria<sup>284</sup>. Processing of OPA1 is stimulated during apoptosis, reflecting the loss of the mitochondrial inner membrane potential, which is independent of any dysfunction to glycolysis or OXPHOS. During mtDNA depletion, levels of apoptosis for U266 cells increased significantly by day 21 of depletion (**Figure 4.21**), which is exemplified by greater expression of the shorter isoforms at positions (d) and (e).

For cells recovered for 14 days after 7 days of mtDNA depletion, there was also an increase in the levels of apoptosis. However, this effect was more pronounced in cells recovered for 7 days after 14 days of mtDNA depletion, where very high levels of apoptosis were detected. By day 14 of recovery after 14 days of depletion, OPA1 processing to its shorter isoform was reduced, indicating reduced levels of apoptosis by this stage. Furthermore, the observation that there was restoration of the long OPA1 isoforms by day 14 of recovery after 14 days

of depletion suggests that any increases to the rates of cell proliferation (**Figure 4.12C**) could be attributed to damage imposed upon the mitochondrion at this stage.



**Figure 4.21.** Western blot analysis showing processivity of OPA1 in mtDNA depleted and/or recovered U266 cells. Processing of the long isoform of OPA1 (b) into its short isoforms (d and e) is shown, with greater expression of the shorter isoform representing high levels of apoptosis in these cells. Lane 5 contained insufficient protein sample and was therefore excluded from analysis. ‘ - ‘ represents mtDNA depletion without recovery, ‘ + ‘ represents mtDNA recovery for the stated number of days. Data were collected by Dr Michael Baker at the Bio21 Institute, Melbourne.

## **4.6. Discussion**

### **4.6.1. MtDNA copy number in normal and cancer cells**

Normal undifferentiated cells usually have low levels of mtDNA copy numbers, which expands with differentiation<sup>285</sup>. These early progenitor cells rely on glycolysis as their primary method for glucose metabolism, and characteristically have faster rates of proliferation<sup>286</sup>. During differentiation, the increase in mtDNA copy numbers is dependent on cell type. Highly metabolically active cells normally have higher requirements for metabolism via OXPHOS, thereby possessing higher mtDNA copy numbers than metabolically inactive cells<sup>287,288</sup>.

The regulation of mtDNA copy number in tumor cells would be expected to follow similar patterns to normal undifferentiated cells, since cancer cells also undergo rapid proliferation. This is supported by others suggesting that mtDNA depletion is responsible for the progression of gastric carcinoma<sup>289</sup>, as well as increased proliferation and reduced sensitivity to apoptosis in osteosarcoma<sup>290</sup> and lung carcinoma cells<sup>291</sup>.

Analysis of the patient tumor samples demonstrated variability in their mtDNA copy numbers (from 280 to 890 copies per cell). Whilst the average mtDNA copy number of all these tumor samples was not significantly different to those in U266 cells. These variations could be due to differences in the mtDNA variants present in these samples, since U266 cells were determined to have a different set of variants compared to patient tumor samples (Chapter 3).

#### **4.6.2. Changes in the gene expression patterns during mtDNA depletion**

The expression of genes participating in the transcription and replication of mtDNA, as well as those associated with the tumor biology of multiple myeloma were much higher in U266 cells compared to the plasma sample. A primary cause of many mtDNA depletion syndromes has been thought to involve the impairment of nuclear-encoded factors responsible for the maintenance of mtDNA, rather than from mutations within the mtDNA itself<sup>292</sup>. However, in this situation, mtDNA depletion acts as the initial cause of metabolic stress, the expression of nuclear-encoded mitochondrial genes were observed to increase during the late stages of mtDNA depletion. This could reflect a compensatory response to protect against the complete loss of mtDNA.

The concept of a compensatory mechanism can also be applied to the gene expression analyses conducted on the key tumorigenic genes for multiple myeloma. It was discovered that expression levels of these genes did not decrease below 50% of levels of U266 control cells at any point during mtDNA depletion, suggesting the presence of a mechanism to preserve and maintain the tumor phenotype of these cells. These results corresponded with others that have suggested increases in nuclear gene expression levels, augmentation of the tumor phenotype and chromosomal instability are common features in mtDNA depleted cells<sup>293</sup>. Moreover, increased nuclear gene expression has been reported in yeast petite mutants that lack mtDNA<sup>294</sup>.

One of the key observations from these results included the up-regulation of tumor proliferation genes, *FGFR3* and *MAPK*, in U266 cells depleted for 14 days

and recovered in the absence of 2-DG (**Figure 4.20E-F**). Despite these changes, these cells proliferated at a significantly slower rate, as observed from their population doubling times (**Figure 4.12C**). The evidence that the reduction in the cell proliferation rates for these U266 cells, depleted and recovered for 14 days each, are attributed specifically to the insufficient replenishment of mtDNA (**Figure 4.12B**) can be explained by the western blot analysis performed on OPA1, which show the presence of intact, functional mitochondria by this stage of recovery.

Results from this Chapter was able to show that decreased expression of the mtDNA-encoded genes during mtDNA depletion led to decreases in the rates of cell proliferation, despite significant up-regulation made to the expression of the nuclear-encoded genes participating in the biogenesis, transcription and replication of mtDNA. These observations emphasise the importance of mtDNA copy numbers in the control of gene expression at both the mitochondrial and nuclear-level.

#### **4.6.3. Changes during mtDNA replenishment in the presence and absence of 2-DG**

Although replenishment of mtDNA copy number could be assisted by the addition of deoxynucleoside monophosphates in culture<sup>295</sup>, for this study the recovery of mtDNA copies was performed by removing the influence of ddC and re-culturing of these cells in fresh media in the absence of ddC. This method allowed for the assessments of the natural rate of mtDNA replenishment after a period of induced stress.

Cells recovered with 2-DG were not able to increase mtDNA copy numbers similar to those cells recovered without 2-DG. This could be due to the reduced availability of metabolic intermediates from the glycolysis pathway that can be used in the pentose phosphate pathway. Such intermediates include ribose-5-phosphate that is required for the biosynthesis of nucleotides, which can be used during mtDNA replication<sup>296</sup>.

The recovery of gene expression after 7 days of mtDNA depletion in the presence of 2-DG, led to increases in levels of the mtDNA-encoded genes, compared to the same cells recovered in its absence. These changes could reflect an attempt to rescue mtDNA-encoded transcript levels when glycolysis is inhibited, thereby attempting to maintain the ability of the ETC to function. However, since the recovery of expression levels for the majority of mitochondrial biogenesis, transcription and replication genes decreased during recovery in the presence of 2-DG after mtDNA depletion, it is possible that the opposite patterns of gene expression observed between the mitochondrial and nuclear genes could be related.

To explain the increased expression of the mtDNA-encoded genes, and the decreased expression of the mitochondrial biogenesis, transcription and replication genes during recovery with 2-DG, the effects induced by 2-DG itself are explored. It has been shown that in the presence of high glucose concentrations, there is the preferential accumulation of the transcription factor complex, MondoA-Mlx, in the nucleus<sup>297</sup>. This heterodimeric complex is

commonly known as a ‘transcriptional biosensor’ that responds to glucose flux, and is responsible for as much as 75% of glucose-induced transcription<sup>298</sup>. MondoA-Mlx is capable of shuttling between the nucleus and mitochondria<sup>299</sup>, and in the presence of glucose it facilitates the binding of MondoA-Mlx to the promoters of target genes, followed by the recruitment of transcriptional coactivators and finally in the activation of gene expression<sup>297</sup>. Therefore, in the presence of 2-DG, which would significantly lower glucose concentrations, the export of MondoA-Mlx from the nucleus could reflect the decrease in the transcription of the nuclear-encoded mitochondrial biogenesis, transcription and replication genes. If the nuclear export of MondoA-Mlx increases its shuttling to the mitochondria, this could increase the transcription of the mtDNA encoded genes, thereby providing a relationship between the transcription of both nuclear and mitochondrial genes, which is dependent on glucose concentrations.

The variable patterns in the expression levels for multiple myeloma tumor biology genes during recovery in the presence of 2-DG, suggest that there may be alternative mechanisms of gene regulation for these cancer genes that is not related to the activities of MondoA-Mlx, but could instead be tumor-specific. Nonetheless, there was the up-regulation of several key cancer survival genes in the absence of functional glycolysis, which indicated the importance of glycolysis in maintaining the survival, and therefore, the tumorigenic phenotype of these cells.

#### **4.7. Conclusion**

MtDNA plays a significant role in the tumor biology of multiple myeloma. The presence of sufficient copies of mtDNA is required to maintain proliferation of these tumor cells, a key characteristic of malignant tumors. Glycolysis was also determined to have an important role in the tumor biology of these cells. The inhibition of glycolysis during mtDNA recovery led to significant increases to the key survival genes *IL6*, *DAD1* and *CASP8*, relative to the cells recovered in the absence of inhibitor, after 7 and 14 days of mtDNA depletion.

These significant increases were not observed for the expression of genes associated with maintaining tumor cell proliferation (*FGFR3* and *MAPK*). This suggests that, in the absence of glycolysis, multiple myeloma cells have reduced capacity to proliferate but attempt to up-regulate the expression of key survival genes as a defense mechanism to avoid a reduction to cell numbers. However, the up-regulation of these survival genes can only protect tumor cells from the loss of cell numbers in the short-term (7 days of mtDNA depletion), and proves ineffective as a defense response after prolonged mtDNA depletion, defined as a period of 14 days. Glycolysis therefore predominately controls the tumorigenic phenotype of these cells, providing enough ATP via this cytoplasmic pathway to support the survival of the cancer cells, whilst being supported on some OXPHOS activity. It is possible that the mtDNA variants identified from Chapter 3, may have a role in contributing towards the tumorigenic phenotype of these multiple myeloma cells.

**Chapter 5: Identification of MtDNA  
Variants in Glioblastoma Multiforme**

## **5.1. Introduction**

To date, a few studies have examined for the presence of variants in the nuclear genome that may contribute towards the progression of GBM<sup>300,301</sup>. One study in particular sequenced the nuclear genome of the U87MG cells, which is a cell line representative of grade IV glioma<sup>182</sup>. Results from this study revealed that sequences encoding for proteins were predominately disrupted in this cell line, due to large-scale chromosomal abnormalities to its genome. Combinations of SNPs, insertions and deletions as well as translocations were identified in this highly mutated cell line. However, there have been no studies that have investigated the potential role of mtDNA variants in the tumor biology of GBM.

To address the role of mtDNA in GBM cells, analyses were performed on the traditional cell lines U87MG<sup>302</sup>, SF-767<sup>302</sup> and HSR-GBM1<sup>303</sup> cells. Often in primary GBM tumors, there is amplification of epidermal growth factor receptor (EGFR) genes<sup>304</sup>, a characteristic believed to be frequently under-represented in cultured GBM cells due to selection pressures that take place *in vitro* to eliminate this gene alteration<sup>305</sup>. To overcome this situation, the GBM 6 and HK301 neurosphere cell lines were used for this investigation, since both express EGFR variant III (EGFRvIII). Other cell lines used in this study included the GBM L1, GBM L2, GBM 4, CSC-014, CSC-020, NO7-152 and BAH-1 cells, which were generated by Associate Professor Terrance Johns (Monash University) from expanded cultures of highly malignant human GBM tumors. Examining these cell lines derived from different GBM patients that may retain patient-specific characteristics in culture will enable analyses to closely represent natural differences in the tumor biology within each of these GBM patients.

The HSR-GBM1 cells were established by Vescovi et al. in 2004<sup>303</sup>. These cells were characterized to display traits closely resembling those of human GBM tumors, which include sharing their capability to migrate *in vivo*<sup>303</sup>. For these reasons, the HSR-GBM1 cells were used as the main cell line for determining whether mtDNA has a role in the establishment of GBM.

## **5.2. Aims and hypothesis**

The aim of this Chapter was to sequence 12 different GBM cell lines, alongside a collection of patient tumor samples to determine whether GBM cells and tumor samples possessed the same variants, or another subset, for the establishment of GBM. The importance of mtDNA towards the development of GBM tumors was also examined.

GBM cell line variants were compared with the non-tumorigenic human neural stem cells to ensure these GBM cell variants have not arisen as an artefact of cell culture. Furthermore, the GBM tumor variants were compared with normal brain biopsy samples to confirm that the identified variants were of tumor origin.

It was hypothesized that GBM cells and tumor samples possess tumor-specific mtDNA variants, which would differ from those identified from the sequencing of multiple myeloma tumors and cells (Chapter 3). Additionally, it was anticipated that GBM cells and tumor samples would share a selection of mtDNA variants characteristic of the tumor phenotype. To determine the importance of mtDNA in the tumor biology of GBM, HSR-GBM1 cells were

depleted of their mtDNA and transplanted into BALB/c nude mice to examine whether the same, or different, variants arise following the replenishment of mtDNA copy numbers as tumor development proceeds *in vivo*. MtDNA variants that become re-established during mtDNA replenishment would likely represent those that are specific to GBM.

### **5.3. Methods**

The methods used are described in the following sections: 2.1.6 to 2.1.8, 2.1.9, 2.4.1, 2.4.2, 2.6, 2.8.1, 2.8.2, 2.9.1, 2.10, 2.16 to 2.24, in addition to those below:

#### **5.3.1. HRM primers used for the screening of GBM mtDNA variants**

Primers were designed for each of the mtDNA variants identified from the Ion Torrent sequencing of various GBM cell lines (**Table 5.1**). All HRM primers produced a product size of between 100bp to 250bp, and all had an optimal annealing temperature of 55°C.

***Table 5.1. Primers used for GBM variant screening on the HRM:***

Variant Position (bp)	Forward primer (5'→3')	Reverse primer (5'→3')	Product size (bp)
194	TGATTCTGCCTCATCTATT	AAGTGGCTGTGCAGACATTC	131
302, 310	CAGGCGAACATACTTACTAAA G	TTAGGCTGGTGTAGGGTTC	184
16186, 16218	CCACCATGAATATTGTACGGT A	TGGCTTTGGAGTTGCAGTTG	147
1386	CAAGGTGTAGCCCATGAGGT	CTGTTCAACTAAGCACTCTAC	132
1682	CACAAAGCACCCAATTACAC	GGTAAATGGTTTGGCTAAGG	116
2130	CACAGAACCCTCTAAATCCC	ACGCTTTCTTAATTGGTGGCTG	145
5253	CTATCTCGCACCTGAAACAAG	CGTAGGTAGAAGTAGAGGTTAAG	196
5752	CAGCTAAGCACCTAATCAAC	TCTAAAGACAGGGTTAGGC	161
6422	CTTCTCCTTACACCTAGCAG	GTAGGACTGCTGTGATTAGG	123
6999	CTCTGAGCCCTAGGATTCAT	CAGTGAATGAAGCCTCCTATG	152
8251, 8252	CTCTGAAATCTGTGGAGCAA C	CTAAGTTAGCTTTACAGTGGG	131
10473	GGATTAGACTGAACGAATTG	AGTAGGGAGGATATGAGGTG	175
10814	CCTAAACCTACTCCAATGCT	AGGAAAAGGTTGGGAACAG	186
11361	GGCTCACTAAACATTCTACTA C	CTGCGCAAGTACTATTGAC	165
11512	CGCTGGGTCAATAGTACTTG	GTTATAATTATGCCTCATAGGGATA G	140
11674	CTGCCTACGACAAACAGAC	GGCAGAATAGTAATGAGGATG	160
12101	CTCACTACCCACCACATTA	CAGTTCTTGTGAGCTTTCTC	216
13061	CAAATCAGCCCAATTAGGTC	GGATTAGTGGGCTATTTCTGCT	154
14160	TCTTCCCACTCATCTAACC	TTAGTAGTAGTTACTGGTTGAAC	109
14426	TCATACTCTTTCACCCACAG	GGTTGTCTTTGGATATACTACAG	137
15264, 15267	GAGGCCAAATATCATTCTGAG	GTTGTTTGATCCCGTTTCGTG	218

### **5.3.2. Culture conditions for the human neural stem cells**

Human neural stem cells were derived from the human embryonic stem cell line (NIH-approved H9 (WA09); Invitrogen)). They were cultured on plates pre-treated with 20mg/ml fibronectin, in media comprised of DMEM/F12 media supplemented with 2% StemPro Neural Supplement, 2mM GlutaMax (all from Invitrogen) and 20ng/ml bFGF and EGF (both from Millipore, Billerica, MA).

### **5.3.3. Normal brain biopsy samples**

Healthy human brain samples were obtained from deceased patients who had donated their brains to the Victorian Brain Bank Network (Human research ethics number 09023B, Southern Health, HREC). Patient ages ranged from 48 to 72, with the average being 59 years of age. All patient samples were absent from any neuropathological disorders and had no known history of neurological disease.

### **5.3.4. GBM patient tumor samples**

GBM tumor samples were obtained from the Victorian Cancer BioBank (Human research ethics number 09023B, Southern Health, HREC). All tumor samples were extracted from patients with a confirmed diagnosed of grade IV GBM. Patient ages ranged from 34 to 80, with the average being 62 years of age.

### **5.3.5. Xenograft tumor models**

Animal work in mice models was approved by the Animal Ethics Committee at Monash University (Approval number: MMCA/2011/76). HSR-GBM1 cells underwent mtDNA depletion with ddC *in vitro* for up to 50 days. Cells depleted

of their mtDNA to levels of 50%, 20%, 3% and 0.2% of the original mtDNA copy numbers were collected and re-suspended in 100µl of complete media (see section 2.1.6) before subcutaneous injection into both flank regions of female BALB/c immunodeficient nude mice (Animal Research Centre, Perth, Australia). All *in vivo* work was carried out by Dr Jacqueline Donoghue. Mice were aged between 4 to 6 weeks old. Extracted tumors were stored in the -80°C until further processing.

## **5.4. Results**

### **5.4.1. Identification of mtDNA variants in GBM cell lines**

12 GBM cell lines were sequenced by Ion Torrent to investigate whether these cells shared the presence of variants, which would identify a common set of variants exists that are representative of GBM cell lines. A total of 13 variants were identified in the non-coding mtDNA region, 4 of which were previously unreported in the comprehensive MitoMap mutations database (2002bp, 2130bp, 2817bp and 3168bp). The majority of variants were present in the D-loop hypervariable segment 1 or hypervariable segment 2 regions (**Table 5.2**). The T→C variant identified at 302bp of the D-loop region possessed the highest level of mutation at 77.9% within the U87MG cells. The variant at 16519bp also had high levels of mutation, ranging from 51.2% to 58.2% in 5 of the 12 GBM lines examined.

Other shared variants within the D-loop region include the 310bp T→C substitution, present in GBM L1, GBM L2 and NO7-152 cells at similar frequencies but was detected at three-fold higher levels within CSC-014 cells. The 302bp variant was identified in 3 different GBM cell lines (GBM L1, CSC-014 and NO7-152), but displayed distinct differences in their levels of mutation depending on the cell line. Several GBM cell lines (GBM L1, GBM L2, GBM 4, CSC-014, CSC-020 and NO7-152) shared the presence of a variant within the 16s rRNA at 2130bp and the origin of light strand replication at 5752bp. Both variants were detected in all 6 GBM lines at similar levels. However, only GBM 6 and U87MG cells shared the 3168bp variant at identical levels reported at 3.7%. Variants observed in HSR-GBM1 cells appeared exclusive to the cell line.

**Table 5.2.** MtDNA variants identified in the non-coding region of 12 GBM cell lines.

	Reference Position	Variant	Percentage change in variant (%)											Gene region	
			HSR-GBM1	GBM L1	GBM L2	GBM 4	GBM 6	CSC 014	CSC 020	NO7 152	SF-767	U87MG	HK301		BAH1
Non-coding region	16186	C→T	4.5												D loop - Hypervariable segment 1, 7S DNA, membrane attachment site
	16218	C→T	16.9												
	16224	T→C										77.9			
	16519	T→C					53.7				55.8	58.2	51.2	52.9	D loop - 7S DNA, membrane attachment site
	194	C→T				4.1									D loop - Hypervariable segment 2, membrane attachment site
	302	A→C		16.4				41.5		39.9					
	310	T→C		5.6	5.0			17.5		6.0					
	1386	T→C	9.9												12s rRNA
	2002 *	G→A					11.1								16s rRNA
	2130 *	A→G		3.9	5.5	4.0		3.6	6.3	7.4					
	2817 *	G→A												11.7	
	3168 *	C→T					3.7					3.7			
5752	A→G		11.3	13.1	11.4		10.5	10.9	9.9					Origin of L-strand replication	

\* Novel mtDNA variants that are not listed in the MitoMap database.

A further 19 variants were detected in the mtDNA coding region (**Table 5.3**), 12 of which were not previously reported in MitoMap (6999bp, 8252bp, 10473bp, 10814bp, 12101bp, 12877bp, 13061bp, 14159bp, 14160bp, 14426bp, 15264bp and 15267bp). The levels of mutation for these 19 coding region variants ranged from 3.1% at position 12101bp within CSC-014 cells to 55.5% in GBM L1 cells at position 8252bp of the *COX II* gene. There were 3 variants within the coding region (6422bp, 8251bp and 11674bp) that caused synonymous amino acid changes. Amongst the 16 variants within the coding region that led to the development of non-synonymous amino acid changes, 5 were predicted by SNPs&GO to impair protein function. However, only 1 variant at 12877bp within *ND5* was predicted to cause dysfunction with high confidence (RI value of 7). This substitution involved a glycine to arginine change (G181R), which was biochemically incompatible since glycine is a small molecule that confers flexibility to the protein structure, whereas arginine is a positively charged residue that can form interactions with other molecules, such as phosphate anions.

Overall, *ND4* had the highest number of mtDNA variants (n = 6), followed by both *ND5* and *ND6* (n = 3). The *ND6* variant at 14160bp was shared between most of the cell lines (in 8 out of 12 cell lines) at relatively low levels (<10%).

**Table 5.3.** *MtDNA variants identified in the coding region of 12 GBM cell lines. MutPred predictions with a P value ≤ 0.05 are shown.*

\* *Novel mtDNA variants that are not listed in the MitoMap database.*

Reference Position	Variant	Percentage change in variant (%)												Gene region	Amino Acid Change	SNPs&GO	Mutpred	
		HSR-GBM 1	GBM L1	GBM L2	GBM 4	GBM 6	CSC 014	CSC 020	NO7 152	SF-767	U87MG	HK301	BAH1				Probability of deleterious mutation	Top 5 predicted features caused by the amino acid mutation
6422	C→T	3.8												COX I	Syn (P)	-	-	
6999 *	G→A	6.5													V366M	Neutral, RI 4	0.454	
8251	G→A		43.5		50.8									COX II	Syn (G)	-	-	
8252 *	C→A		55.5		46.8										P223T	Neutral, RI 9	0.327	
10473 *	C→G						4.3							ND4L	P2A	Neutral, RI 9	0.344	Loss of disorder (P = 0.0496)
10814 *	A→C		6.0		5.3			5.0	5.2					ND4	K19Q	Neutral, RI 8	0.551	Loss of methylation at K19 (P = 0.0012) Loss of ubiquitination at K19 (P = 0.0283)
11361	T→C						6.0								M201T	Neutral, RI 5	0.706	
11512	C→A		6.0		9.6		11.0	6.4	15.2						N251K	Neutral, RI 4	0.495	Gain of methylation at N251 (P = 0.0194)
11674	C→T				3.5										Syn (T)	-	-	
12101 *	T→C				3.8		3.1		4.2						S448P	Neutral, RI 7	0.484	Loss of helix (P = 0.0093) Gain of loop (P = 0.0321)
12102	C→T				3.7				3.2						S448F	Neutral, RI 3	0.472	
12877 *	G→C									27.7				ND5	G181R	Disease, RI 7	0.795	
13043	C→T											3.3			A236V	Neutral, RI 3,	0.786	
13061 *	C→A							4.0							P242Q	Disease, RI 0	0.776	Loss of glycosylation at P242 (P = 0.035)
14159 *	C→G			4.1										ND6	R172P	Disease, RI 3	0.423	Loss of methylation at R172 (P = 0.0305)
14160 *	G→C			4.0	3.5	3.3	7.5		4.3	3.2		8.1	5.0		R172G	Neutral, RI 2	0.442	Loss of methylation at R172 (P = 0.0305)
14426 *	C→T						8.9								G85E	Neutral, RI 7	0.364	Loss of glycosylation at S84 (P = 0.0357) Gain of solvent accessibility (P = 0.0456)
15264 *	C→T	14.1												CYT B	P173L	Disease, RI 4	0.361	
15267 *	C→G	20.5													T174S	Disease, RI 3	0.866	

#### **5.4.2. Validation of the GBM cell line variants using HRM analyses**

All GBM cell line variants identified by Ion Torrent sequencing were confirmed using HRM analysis (**Table 5.4**). The extra variants identified from HRM analyses, are likely to represent variants between levels of 1% and 3%, which were determined as the threshold levels of variant detection for the HRM and Ion Torrent, respectively.

The similarities of each GBM cell line were compared against the variant profile for HSR-GBM1 cells, since these cells represented a well-established, traditional line that closely mimic the biological characteristics of human GBM tumors, and as a result, is commonly used for *in vitro* studies for GBM<sup>306,307</sup>. 3 cell lines (GBM 6, CSC-020 and NO7-152) shared over 80% of variants with HSR-GBM1 cells, whereas 5 (GBM L2, GBM 4, CSC-014, HK301 and BAH-1) shared above 70% of their variant profile with the reference cell line. The variants within U87MG cells had 68% similarity to the HSR-GBM1 cells. However, both GBM L1 and SF-767 cells had variant profiles most distinct from HSR-GBM1 cells, sharing only 56% of variants with this cell line.

**Table 5.4.** Validation of the GBM cell line variants using HRM analyses.

		GBM cell line samples												Gene region	
Reference Position	Variant	HSR-GBM1	GBM L1	GBM L2	GBM 4	GBM 6	CSC 014	CSC 020	NO7 152	HK301	BAH 1	SF-767	U87MG		
Non-coding region	16186	C→T	####										####	D loop	
	16218	C→T	####										####		
	194	C→T	####			####		####	####						
	302	A→C		####		####		####		####		####			
	310	T→C		####	####	####		####		####		####			
	1386	T→C	####												12s rRNA
	2130	A→G		####	####	####		####	####	####			####	####	16s rRNA
	5752	A→G	####	####	####	####		####	####	####					Origin of L-strand replication
Coding region	6422	C→T	####											Cytochrome C Oxidase I	
	6999	G→A	####											Cytochrome C Oxidase II	
	8251	A→G		####		####								NADH dehydrogenase 4L	
	8252	C→A		####		####								NADH dehydrogenase 4	
	10473	C→G		####	####	####		####	####		####				
	10814	A→C		####		####		####	####	####					
	11361	T→C						####							
	11512	C→A		####	####	####		####	####	####		####			
	11674	C→T		####		####		####							
	12101	T→C		####	####	####		####		####		####	####	####	
	12102	C→T		####	####	####				####		####	####	####	
	13061	C→A	####						####				####	NADH dehydrogenase 5	
	14159	C→G		####	####				####					NADH dehydrogenase 6	
	14160	G→C		####	####	####		####	####	####	####	####	####		
	14426	C→T	####	####				####	####				####		
15264	C→T	####	####				####			####		####	Cytochrome B		
15267	C→G	####	####				####			####		####			
Similarity relative to HSR-GBM1		100%	56%	72%	72%	80%	72%	80%	84%	76%	72%	56%	68%		

 Variant present    
  Variant absent    
  Variant present (sites jointly analysed)    
  Variant absent (sites jointly analysed)    
 ##### Variant presence/absence confirmed in both Ion Torrent and HRM analyses

### 5.4.3. Susceptibility of mtDNA regions to the development of mutations

Regions of mtDNA all differ in terms of their nucleotide lengths (bp). To determine the susceptibility of each mtDNA region to the development of mutation, the number of mtDNA variants were normalised against the length of the particular mtDNA region within which it resides.

Within the non-coding region, the 12s rRNA was determined least susceptible to the development of variants (score:  $1.05 \times 10^{-3}$ ), with only 1 variant identified within a 953bp region (**Table 5.5**). On the contrary, the origin of light strand replication, as a 77bp region, had the highest probability of developing mutation (score:  $1.30 \times 10^{-2}$ ), which was 12.38 fold greater than for the 12s rRNA region (**Table 5.6**). The D-loop was the second highest mtDNA non-coding region to be highly susceptible to mutation, as it had a probability score of  $6.24 \times 10^{-3}$ , a fold change 5.95 greater than the 12s rRNA (**Table 5.6**).

Although *ND4* contained the highest number of variants (6) among the coding genes, *ND6* was determined to be the region most susceptible to developing mutation (score:  $5.73 \times 10^{-3}$ ), whilst *COX I* was least susceptible (score:  $1.30 \times 10^{-3}$ ) (**Table 5.5**), which was a difference of 4.41 fold (**Table 5.6**). *CYT B* and *ND5* shared similar probability scores ( $1.75 \times 10^{-3}$  and  $1.66 \times 10^{-3}$ , respectively), which was a 1.06 fold difference. Altogether, the non-coding region had a greater chance ( $2.08 \times 10^{-3}$ ) of developing a mutation than the coding region ( $1.84 \times 10^{-3}$ ) (**Table 5.5**), which was a difference of 1.13 fold.

**Table 5.5.** Susceptibility of mtDNA regions to the development of mutations

Reference Position	Gene region	Nucleotide positions for gene regions (bp)	Size of gene region (bp)	Number of variants	Normalised data	
					Mutation rate in each gene region	Mutation rate grouped for coding/non-coding regions
16186, 16218, 16224, 16519...194, 302, 310	Non-coding region	D-loop	1..576, 16024..16569	1121	7	0.00208
1386		12s rRNA	648..1601	953	1	
2002, 2130, 2817, 3168		16s rRNA	1671..3229	1558	4	
5752		Origin of light strand replication	5721..5798	77	1	
6422, 6999	Coding region	COX I	5904..7445	1541	2	0.00184
8251, 8252		COX II	7586..8269	683	2	
10473		ND4L	10470..10766	296	1	
10814, 11361, 11512, 11674, 12101, 12102		ND4	10760..12137	1377	6	
12877, 13043, 13061		ND5	12337..14148	1811	3	
14159, 14160, 14426		ND6	14149..14673	524	3	
15264, 15267		CYT B	14747..15887	1140	2	

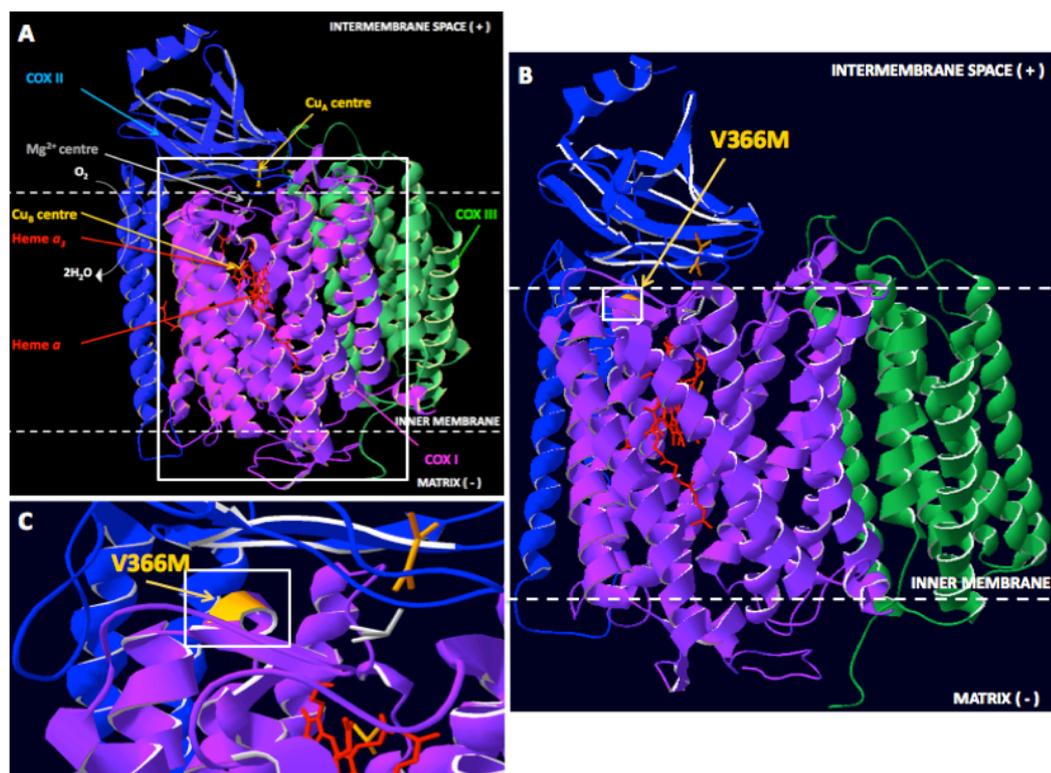
**Table 5.6.** Fold change differences on the susceptibility of mtDNA regions to mutation. Values greater than 1 (black) indicate that the probability of the gene listed on the left is greater than for the gene listed across the top. Values in red inform the opposite.

Gene region	D-loop	12s rRNA	16s rRNA	Origin of L-strand	COX I	COX II	ND4L	ND4	ND5	ND6	CYT B
D loop		5.95	2.43	0.48	4.81	2.13	1.85	1.43	3.77	1.09	3.56
12s rRNA	0.17		0.41	0.08	0.81	0.36	0.31	0.24	0.63	0.18	0.60
16s rRNA	0.41	2.45		0.20	1.98	0.88	0.76	0.59	1.55	0.45	1.46
Origin of L-strand	2.08	12.38	5.06		10.01	4.44	3.84	2.98	7.84	2.27	7.40
COX I	0.21	1.24	0.51	0.10		0.44	0.38	0.30	0.78	0.23	0.74
COX II	0.47	2.79	1.14	0.23	2.26		0.87	0.67	1.77	0.51	1.67
ND4L	0.54	3.22	1.32	0.26	2.60	1.15		0.78	2.04	0.59	1.93
ND4	0.70	4.15	1.70	0.34	3.36	1.49	1.29		2.63	0.76	2.48
ND5	0.27	1.58	0.65	0.13	1.28	0.57	0.49	0.38		0.29	0.94
ND6	0.92	5.46	2.23	0.44	4.41	1.96	1.69	1.31	3.46		3.26
CYT B	0.28	1.67	0.68	0.14	1.35	0.60	0.52	0.40	1.06	0.31	

#### 5.4.4. Protein modeling of the GBM cell line variants

Protein modeling was performed based on the bovine protein structures for COX I, COX II and CYT B, which have protein sequence identities to their human counterparts of 91%, 72% and 78.4%, respectively. The *CYT B* variants at 15264bp and 15267bp were discussed previously in section 3.4.8.1 (**Figure 3.13**), which were found in the multiple myeloma tumor samples.

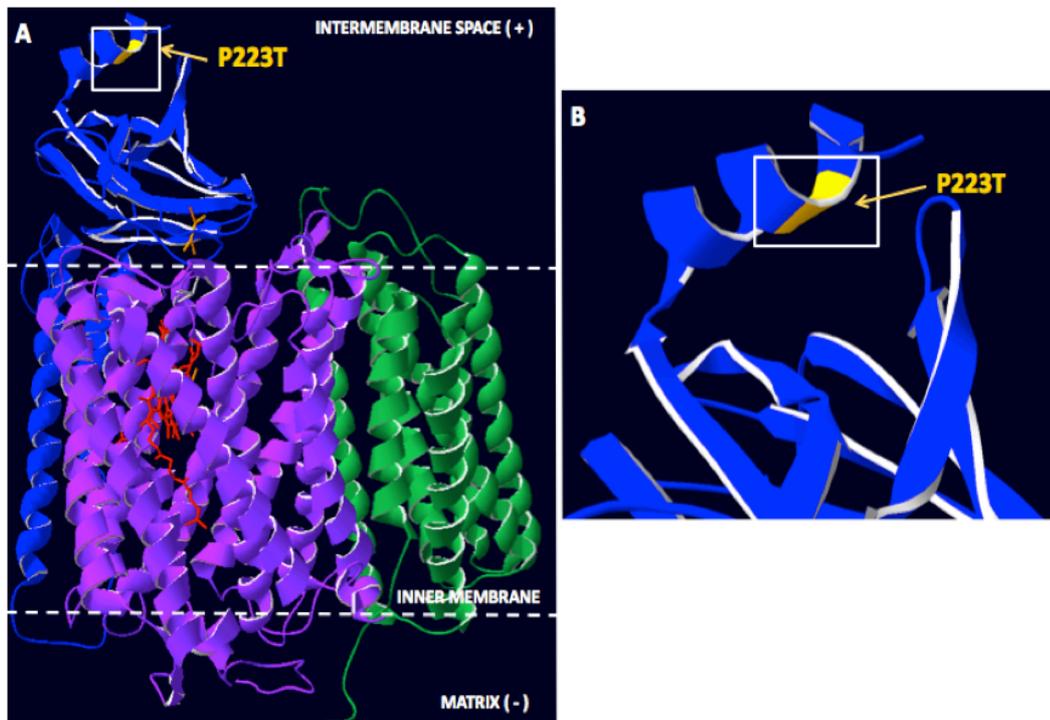
The COX I mutations at position 366, involving a substitution from valine to methionine, was predicted by SNPs&GO to be a neutral mutation. Modeling of



**Figure 5.1.** Modeling of the V366M COX I mutation using the bovine model. (A) The catalytic centre of Complex IV is shown. (B) The V366M mutation resides within a transmembrane helix facing towards the mitochondrial intermembrane space. (C) Closer observation of the V366M mutation.

the GBM-specific mutation identified its location on the 9<sup>th</sup> transmembrane helical domain, where it faces towards the intermembrane space (**Figure 5.1**).

The amino acid change from proline to threonine at position 223 (P223T) within COX II represents another neutral substitution. This mutation was modeled on a helical segment located within the intermembrane space far from the catalytic core area, which supports the prediction of it being a neutral mutation (RI value 9) (**Figure 5.2**).



**Figure 5.2.** Modeling of the P223T COX II mutation using the bovine model. (A) Location of the P223T mutation (yellow). (B) Magnified image of the P223T mutation.

#### 5.4.5. Identification of mtDNA variants in human neural stem cells

Human neural stem cells (hNSCs) were used as a culture control to investigate whether the variants identified across the GBM cell lines were truly representative of those within GBM cells, and have not arisen as an artefact of culture conditions<sup>308</sup>. Sequencing of hNSCs revealed that all variants within the non-coding region of mtDNA resided in the D-loop region, whereas variants within the coding region were mostly identified in the *ND5* region, with one residing in *CYT B*. The variant at position 16519bp in the D-loop had the highest levels of mutation at 51.3%. However, for the majority of variants the levels of mutation were approximately 3%, with the exception of the *CYT B* 15153bp variant at 9.4% (**Table 5.7**). Compared to the GBM cell lines, the hNSCs exhibited much lower numbers of mtDNA variants. Both cell types possessed a different set of mtDNA variants that were not common to one another, except for at position 16519bp within the D-loop region. Only two variants were previously unreported in MitoMap within the hNSCs, which included those at positions 13762bp and 13985bp.

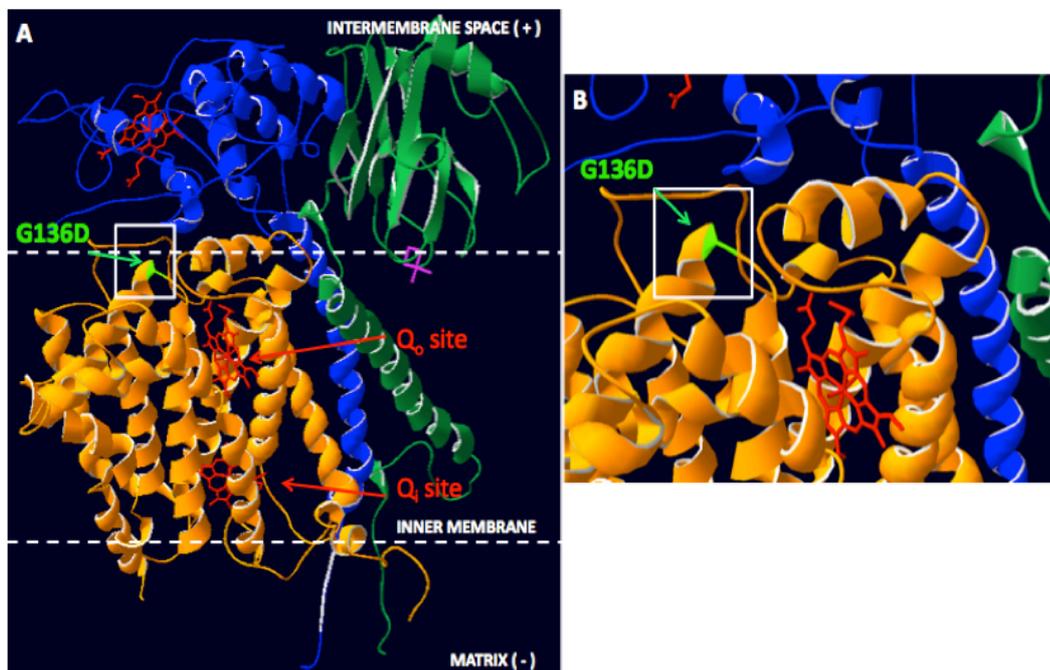
Although predictions from SNPs&GO have indicated that the 13762bp T→C variant may impair protein function, the prediction was given a reliability index of 1. The second *ND5* variant resulted in a synonymous amino acid change that would not influence protein function. The final *ND5* variant identified at 13985bp was predicted to exert a neutral effect on the overall protein structure.

**Table 5.7.** Identification of mtDNA variants in human neural stem cells

	Reference Position	Reference	Variant	Percentage change in variant (%)	Gene region	Amino acid change	SNPs&GO	Mutpred	
								Probability of deleterious mutation	Top 5 predicted features caused by amino acid mutation
Non-coding	16519	T	C	51.3	D loop - Membrane attachment site, 7S DNA				
	496	C	T	3.3	D loop - Hypervariable segment 3				
Coding	13762 *	T	C	3.2	NADH Dehydrogenase 5: ND5	S476P	Disease, RI 1, Uniprot P03915	0.392	<b>Gain of catalytic residue at S476 (P = 0.0028)</b> Gain of glycosylation at P477 (P = 0.068) Loss of helix (P = 0.1299) Loss of disorder (P = 0.2536) Gain of loop (P = 0.2754)
	13984	C	T	3.6		Syn (L)	-		
	13985 *	T	C	3.6		L550P	Neutral, RI 1, Uniprot P03915	0.457	<b>Gain of loop (P = 0.0013)</b> <b>Loss of helix (P = 0.0041)</b> <b>Gain of catalytic residue at P549 (P = 0.0244)</b> <b>Gain of glycosylation at L550 (P = 0.0364)</b> <b>Gain of relative solvent accessibility (P = 0.0479)</b>
	15153	G	A	9.4	Cytochrome B: CYT B	G136D	Disease, RI 8, Uniprot P00156	0.889	Loss of catalytic residue at G136 (P = 0.2619) Loss of sheet (P = 0.3635) Loss of loop (P = 0.3664) Loss of helix (P = 0.3949) Loss of MoRF binding (P = 0.4724)

\* Novel mtDNA variants that are not already listed in the MitoMap database are indicated by asterisks.

The *CYT B* variant at 15153bp was the only mtDNA variant predicted to be disease causing with high confidence (RI value 8). This glycine to aspartate substitution was located in the loop region just before the 8<sup>th</sup> transmembrane-spanning helix structure in the bovine *CYT B* model (**Figure 5.3**). It is possible that this substitution was predicted to be high risk for the onset of disease, as it is in relatively close proximity to Heme b<sub>L</sub> (Q<sub>o</sub> site). However, due to the mutation being located within a flexible loop domain, it could be argued that biologically this variant imposes no risk to protein impairment since these regions are not structurally important features of the overall protein. Furthermore, this variant was detected at relatively low levels (less than 10%) amongst an almost completely wild-type mtDNA sequence, which could mask any negative effects caused by the presence of this variant.



**Figure 5.3.** Modeling the hNSC G136D *CYT B* mutation using the bovine model.

#### 5.4.6. Identification of mtDNA variants in normal brain biopsy samples.

Normal brain samples were sequenced to determine the number of mtDNA variants present, and whether these were similar or different to those variants already identified in the GBM cell lines. Within the non-coding region, 2 novel variants were found that were not previously reported in MitoMap (2232bp and 2647bp). The majority (81%) of non-coding region variants were identified within hypervariable regions 1 and 2 of the D-loop region (**Table 5.8**). Amongst the 13 samples, sample number 05-470 had the highest number of mtDNA variants (12) within the non-coding region, whereas the other 12 samples had less than 5 variants within this same region. Only the variant at 16519bp was shared across the normal brain samples. Overall, the levels of mutation for all non-coding mtDNA variants were below 10%, with the exception of a few samples. Sample 04-273 had an A→G change at position 16293bp, reported at 16.1%, and sample 05-470 had a variant G→C at 2647bp, reported at 12.1% within the 16s rRNA region. Furthermore, 4 samples (07-104 and 07-445, 16-1 and 22-1) had no variant in their non-coding regions.

39 variants were identified across all normal brain samples that resided within the mtDNA-coding region (**Table 5.9**), 14 of which had not been previously reported in MitoMap (4701bp, 4878bp, 4879bp, 8497bp, 8630bp, 9526bp, 9528bp, 9558bp, 11516bp, 11725bp, 11791bp, 12719bp, 13985bp and 14159bp). *CYT B* contained the most variants, accounting for 23% of those identified in the coding region. Sample 09-005 harbored the greatest number of mtDNA variants (11), which were present at ≤6.2%. Sample 08-012 possessed the second-highest number of variants (7) of which two variants were detected at 12.5% at positions

**Table 5.8.** Identification of mtDNA non-coding variants in normal brain biopsy samples

Reference Position	Variant change	Percentage change in variant (%)													Gene region	
		03-965	04-273	05-470	07-104	07-445	07-635	08-012	08-026	09-005	09-152	16-1	22-1	20-1		
16069	C→T			7.2												D loop - Membrane attachment site, Hypervariable segment 1
16126	T→C			5.2												D loop - Membrane attachment site, Hypervariable segment 1, 7S DNA
16145	G→A			4.8												D loop - Membrane attachment site, Hypervariable segment 1, 7S DNA, termination-associated sequence
16179	C→T									4.3						D loop - Membrane attachment site, Hypervariable segment 1, 7S DNA
16222	C→T			5.1												
16261	C→T			9.4												
16293	A→G		16.1													
16356	T→C									3.0						
16519	T→C	57.5		54.5			56.0		57.6	58.3	59.4					
9	G→T			5.8												D loop - Membrane attachment site, 7S DNA
10	T→C			5.8												
66	G→T														3.2	D-loop - Hypervariable segment 2, membrane attachment site, 7S DNA
73	A→G			6.2												D loop - Hypervariable segment 2, H-strand origin, Membrane attachment site, 7S DNA
150	C→T									4.4						
189	A→G							3.0								
195	T→C									4.5						D loop - Hypervariable segment 2, H-strand origin, Membrane attachment site
242	C→T			7.7												D loop - Hypervariable segment 2, H-strand origin, mtTF1 binding site, Membrane attachment site
751	A→T							3.9								12s ribosomal RNA (rRNA)
2232 *	A→T						3.0									16s ribosomal RNA (rRNA)
2647 *	G→C			12.1												
3010	G→A			3.5												

\* Novel mtDNA variants that are not already listed in the MitoMap database are indicated by asterisks.

**Table 5.9.** Identification of mtDNA coding variants in normal brain biopsy samples.

Reference (bp)	Variant change	Percentage change in variant (%)											Gene region	AA change		
		03-965	04-273	05-470	07-104	07-445	07-635	08-012	08-026	09-005	09-152	16-1			22-1	20-1
4646	T→C									3.4					NADH Dehydrogenase 2: ND2	Syn (Y)
4701 *	A→T				4.8											N78Y
4703	T→C				4.8											Syn (N)
4878 *	G→A									4.6						A137T
4879 *	C→G									5.0						A137G
6146	A→G									3.1					Cytochrome C Oxidase I: COX I	Syn (W)
8497 *	A→G		12.5							4.8					ATPase 8: ATP8	Syn (M)
8630 *	A→G											3.7			ATPase 6: ATP6	K35 Stop
8756	T→C								6.5							I77T
8790	G→A					18.4										Syn (L)
8994	G→A								4.7							Syn (L)
9070	T→G									6.2						S182A
9258	C→T					3.2										Syn (L)
9526 *	C→G		5.0				3.8								Cytochrome C Oxidase III: COX III	A107G
9528 *	C→A		9.9	6.4				4.8								P108T
9558 *	C→A					5.8			3.6							P118T
10398	A→G			5.4											NADH Dehydrogenase 3	T114A
10993	G→A			4.6											NADH Dehydrogenase 4: ND4	Syn (M)
11332	C→T									3.4						Syn (A)
11467	A→G									3.1						Syn (L)
11516 *	C→A							4.9								L253M
11725 *	A→T							3.0								Syn (T)
11791 *	C→T											11.8				Syn (L)
12719 *	T→C	3.2														M128T
12774	C→T											9.3				Syn (G)
13984	C→T										3.7				NADH Dehydrogenase 5: ND5	Syn (L)
13985 *	T→C			3.1						3.1	3.7					L550P
14155	C→T										3.8				NADH Dehydrogenase 6: ND6	Syn (G)
14159 *	C→G							12.5							R172P	
14160 *	G→C					8.1		12.5							R172G	
14770	C→A												8.0		Cytochrome B: CYT B	N8K
14823 *	A→C					4.0										N26T
14857	T→C							3.8								Syn (L)
14861 *	G→T							3.6								A39S
14866	C→T									5.4						Syn (C)
15287	T→C					26.9										F181L
15452	C→A			3.2												L236I
15579 *	A→T						3.7									Y278F
15693	T→C									4.9						M316T

\* Novel mtDNA variants that are not already listed in the MitoMap database are indicated by asterisks.

14159bp and 14160bp. Sample 22-1 had only 1 variant, whereas samples, 20-1 and 03-965 had no variant.

Although many variants were identified within the mtDNA-coding region of normal brain samples, they appear to have arisen randomly since very few variants are commonly shared amongst the samples analysed (**Table 5.9**). At most, two variants at positions 9528bp (within *COX III*) and 13985bp (within *ND5*) were shared between 3 samples.

#### **5.4.7. Susceptibility of the mtDNA regions to mutation within normal brain biopsy samples.**

The D-loop region was the region most susceptible to developing a mutation across both the non-coding and coding regions of mtDNA in normal brain samples. The susceptibility score was  $1.52 \times 10^{-2}$  (**Table 5.10**), which was 1.92 fold greater than the second most susceptible region (*CYT B*, score:  $7.89 \times 10^{-3}$ ) (**Table 5.11**). *COX I* was the least susceptible region overall to develop mtDNA variants ( $6.49 \times 10^{-4}$ ). Comparisons between the coding and non-coding regions determined that the coding region had a marginally greater chance ( $3.78 \times 10^{-3}$ ) than the non-coding region ( $3.37 \times 10^{-3}$ ) of acquiring variants, namely 1.12 fold higher for the coding region.

**Table 5.10.** Susceptibility of the mtDNA regions to mutations within normal brain biopsy samples.

Reference Position	Gene region	Nucleotide positions for gene regions (bp)	Size of gene region (bp)	Number of variants detected in gene region	Normalised data			
					Mutation rate in each gene region	Mutation rate grouped for coding/non-coding regions		
16069, 16126, 16145, 16179, 16222, 16261, 16293, 16356, 16519...9, 10, 66, 73, 150, 189, 195, 242	Non-coding	D loop	1..576, 16024..16569	1121	17	0.0152	0.00336	
751		12s rRNA	648..1601	953	1			0.00105
2232, 2647, 3010		16s rRNA	1671..3229	1558	3			0.00193
4646, 4701*, 4703, 4878*, 4879*	Coding region	NADH dehydrogenase 2	4470..5511	1041	5	0.00480	0.00378	
6146		Cytochrome C Oxidase I	5904..7445	1541	1	0.000649		
8497*		ATPase 8	8366..8572	206	1	0.00485		
8630*, 8756, 8790, 8994, 9070		ATPase 6	8527..9207	680	5	0.00735		
9258, 9526*, 9528*, 9558*		Cytochrome C Oxidase III	9207..9990	783	4	0.00511		
10398		NADH dehydrogenase 3	10059..10404	345	1	0.00290		
10993, 11332, 11467, 11516*, 11725*, 11791*		NADH dehydrogenase 4	10760..12137	1377	6	0.00436		
12719*, 12774, 13984, 13985*		NADH dehydrogenase 5	12337..14148	1811	4	0.00221		
14155, 14159*, 14160*		NADH dehydrogenase 6	14149..14673	524	3	0.00573		
14770, 14823*, 14857, 14861*, 14866, 15287, 15452, 15579*, 15693		Cytochrome B	14747..15887	1140	9	0.00789		

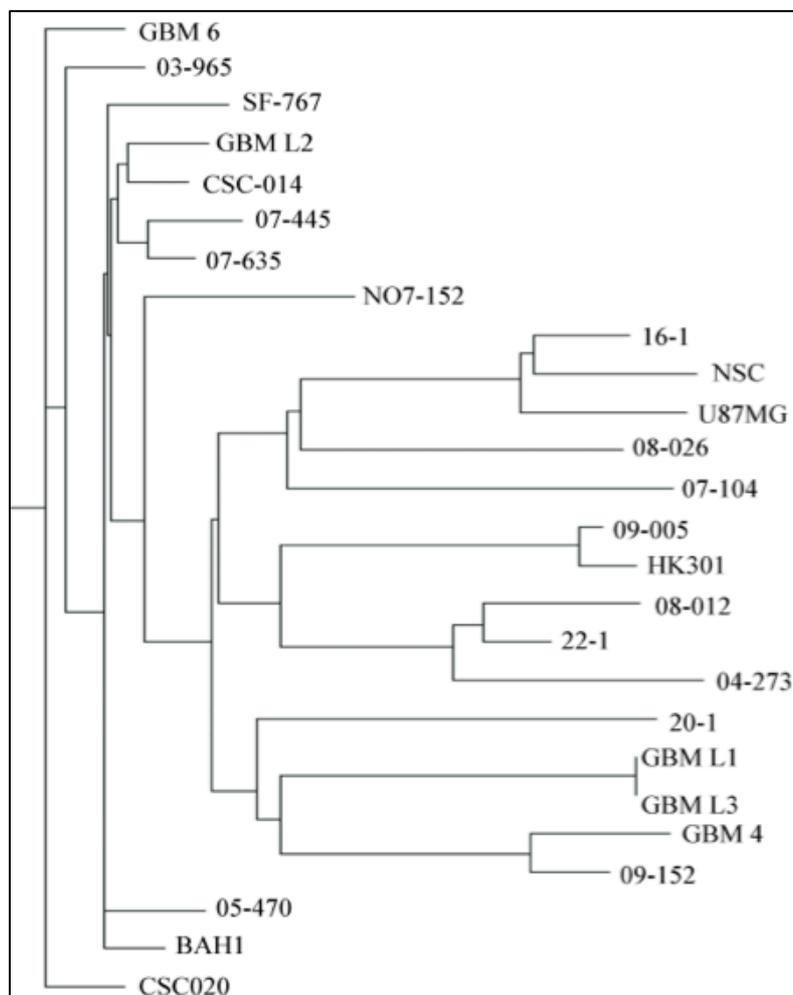
\* Novel mtDNA variants that are not already listed in the MitoMap database are indicated by asterisks.

**Table 5.11.** Fold change differences on the susceptibility of mtDNA regions to mutation, for normal brain biopsy samples. Values greater than 1 (black) indicate that the probability of the gene listed on the left is greater than for the gene listed across the top for the development of variants. Values in red indicate that the gene listed on the left has a lower susceptibility for the development of mutations than the gene listed across the top.

Gene region	D-loop	12s rRNA	16s rRNA	ND2	COX I	ATP8	ATP6	COX III	ND3	ND4	ND5	ND6	CYT B
D-loop		14.45	7.88	3.16	23.37	3.12	2.06	2.97	5.23	3.48	6.87	2.65	1.92
12s rRNA	0.07		0.54	0.22	1.62	0.22	0.14	0.21	0.36	0.24	0.48	0.18	0.13
16s rRNA	0.13	1.84		0.40	2.97	0.40	0.26	0.38	0.66	0.44	0.87	0.34	0.24
ND2	0.32	4.58	2.49		7.40	0.99	0.65	0.94	1.66	1.10	2.17	0.84	0.61
COX I	0.04	0.62	0.34	0.14		0.13	0.09	0.13	0.22	0.15	0.29	0.11	0.08
ATP8	0.32	4.63	2.52	1.01	7.48		0.66	0.95	1.67	1.11	2.20	0.85	0.61
ATP6	0.48	7.01	3.82	1.53	11.33	1.51		1.44	2.54	1.69	3.33	1.28	0.93
COX III	0.34	4.87	2.65	1.06	7.87	1.05	0.69		1.76	1.17	2.31	0.89	0.65
ND3	0.19	2.76	1.51	0.60	4.47	0.60	0.39	0.57		0.67	1.31	0.51	0.37
ND4	0.29	4.15	2.26	0.91	6.71	0.90	0.59	0.85	1.50		1.97	0.76	0.55
ND5	0.15	2.10	1.15	0.46	3.40	0.45	0.30	0.43	0.76	0.51		0.39	0.28
ND6	0.38	5.46	2.97	1.19	8.82	1.18	0.78	1.12	1.98	1.31	2.59		0.73
CYT B	0.52	7.52	4.10	1.64	12.17	1.63	1.07	1.55	2.72	1.81	3.57	1.38	

#### 5.4.8. Phylogenetic analysis on the GBM cell lines and normal brain biopsy samples

Using phylogenetic analysis, it was determined that the GBM cell lines and normal brain samples were from comparable lineages (**Figure 5.4**). Since the normal brain biopsy samples were integrated evenly amongst the tumor cell lines, there was no bias in the genetic distribution of mtDNA. The mtDNA variants identified were therefore not from one specific population or group.



**Figure 5.4.** *Phylogenetic analysis on the GBM cell lines and normal brain biopsy samples. The phylogenetic tree was produced using the phylogeny inference package, PHYLIP.*

#### 5.4.9. Identification of mtDNA variants in GBM patient tumor samples

Ion Torrent sequencing was also performed on 17 GBM patient tumor samples. Within the non-coding region, 5 novel variants were identified (16162bp, 707bp, 2923bp, 7470bp and 12262bp). The D-loop region alone was found to possess 15 variants (**Table 5.12**), the majority of which were discovered in a single GBM tumor (sample 0-8745) at high frequencies within hypervariable regions 1 and 2. Position 16519bp was again a common site for mutation amongst the tumor samples, as variants at all other non-coding sites arose exclusively within each of the tumors. Only the D-loop variants displayed mutation levels as high as 69.1% (sample 0-8745 at positions 16311bp and 195bp).

Sample 0-8745 contained the most variants (13 out of 22) compared to all other tumor samples. For samples other than 0-8745, there were at most 2 variants detected within the non-coding region (samples 0-9022, 0-80243, 10-0214 and 20-2). A few samples (0-9022, 0-7125, 0-8585 and 22-2) exhibited the absence of any mtDNA variant within their non-coding regions. Overall, there was a variable pattern for the presence of variants within the mtDNA non-coding region across the GBM tumor samples, which ranged from 3.1% to 69.1%.

Analysis of the mtDNA coding region revealed that all variants were exclusive to each tumor (**Table 5.13**), as there was the absence of any shared variants amongst the 17 samples. Within the coding region, there were a total of 12 novel variants previously unreported in MitoMap (7428bp, 8088bp, 8447bp, 8491bp, 8558bp, 9577bp, 9578bp, 13346bp, 14159bp, 14588bp, 15354bp and 15660bp). The highest level of mutation detected was 87.9% for the *COXI* variant at

**Table 5.12. Identification of mtDNA non-coding variants in GBM patient tumor samples.**

Reference Position	Variant	Percentage change in frequency (%)															Gene region		
		0-8745 (A)	0-9022 (B)	0-9258 (C)	0-7125 (D)	0-9311 (E)	0-80680 (F)	0-8585 (G)	0-80243 (H)	0-85279 (I)	10-0214 (J)	0-7257 (K)	0-9635 (L)	16-2 (M)	22-2 (N)	20-2 (O)		9-2 (P)	5-2 (Q)
16093	T→C					67.4													D-loop - membrane attachment site, hypervariable segment I
16111	C→T	65.9																	D-loop - membrane attachment site, hypervariable segment I, 7S DNA
16162 *	A→G	44.9																	
16185	C→T							3.1											
16261	C→T									3.1									
16293	A→G	69.0																	
16311	T→C	69.1																	
16319	G→A																4.5		
16519	T→C	17.5	66.5				61.7		64.9	61.7	55.8	63.1	64.3						D-loop - membrane attachment site, 7S DNA
66	G→T															3.1			D-loop - Hypervariable segment 2, membrane attachment site, 7S DNA
146	T→C	51.8																	D-loop - Hypervariable segment 2, H-strand origin, membrane attachment site
151	C→T	68.3																	
152	T→C	68.3																	
153	A→G																	47.8	
195	T→C	69.1																	
707 *	C→T	7.3																	12s ribosomal RNA (rRNA)
1346	A→G																	5.5	tRNA valine
1646	T→C	28.9																	16s ribosomal RNA (rRNA)
2648	T→C	23.4																	
2923 *	G→A		12.3																tRNA serine 1
7470 *	C→T													12.1					tRNA serine 2
12262 *	C→T	12.7																	

\* Novel mtDNA variants that are not already listed in the MitoMap database are indicated by asterisks.

position 6146bp involving an A→G substitution (Sample 0-80243). The second highest level of mutation was reported at 75.7% for the *ND4* variant at position 10873bp (0-9635). The lowest frequency was 3.1% for the C→G substitution at 14159bp within *ND6* (0-8745). Within the coding region, the majority of tumor variants gave rise to either synonymous mutations or was predicted to exert neutral effects on protein function (**Table 5.13**). This suggests that although some variants may have high levels of mutations, not all are capable of causing protein impairment, but instead could represent natural base substitutions.

#### **5.4.10. Susceptibility of the mtDNA regions to mutation for GBM patient tumor samples.**

Although the majority of non-coding region mutations for the GBM tumor samples arose in the D-loop region, normalization of the number of variants to the size of the gene regions revealed that the tRNA serine 2 region possessed the greatest probability of developing a mutation ( $1.72 \times 10^{-2}$ ) (**Table 5.14**). The tRNA serine 2 region had a 1.17 fold greater difference (**Table 5.15**) than both the tRNA valine and tRNA serine 1 regions, which had susceptibility scores of  $1.47 \times 10^{-2}$  (**Table 5.14**). The 16s rRNA region was the least susceptible of all non-coding regions to mutation ( $1.28 \times 10^{-3}$ ). For the coding regions, *ATPase 8* had the highest susceptibility ( $9.71 \times 10^{-3}$ ) (**Table 5.14**), which was 7.48 fold greater (**Table 5.15**) than the least susceptible region, *COXI* ( $1.30 \times 10^{-3}$ ).

**Table 5.13.** Identification of mtDNA coding variants in GBM patient tumor samples. Due to space constraints, GBM tumor samples are labeled according to Table 5.12, and predicted ‘neutral’ (N) and ‘disease’ (D) outcomes from SNPs&GO are shown with its associated reliability index in brackets.

Reference Position	Variant	Percentage change in frequency (%)															Gene region	Amino acid change	SNPs&GO	MutPred			
		(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)	(M)	(N)	(O)				(P)	(Q)	Probability of deleterious mutation	Predicted impact of the amino acid change on protein function #
6146	A→G								87.9										Cytochrome c oxidase I: COX I (Uniprot P00395)	Syn (W)			
7428 *	G→A													17.3						V509M	N (9)	0.259	Gain of disorder (P = 0.0339)
8027	G→A					7.6														A148T	N (8)	0.095	
8088 *	T→C																4.2		Cytochrome c oxidase II: COX II (Uniprot P00403)	L168S	D (5)	0.796	Gain of glycosylation L168 (P = 0.0225) Loss of stability (P = 0.0237) Gain of disorder (P = 0.0265)
8447 *	A→C	4.3																	ATP synthase 8: ATP8 (Uniprot P03928)	M28L	N (1)	0.419	Gain of catalytic residue M28 (P = 0.001) Gain of methylation K27 (P = 0.0411)
8491 *	A→C																48.9			M42I	N (3)	0.261	Gain of methylation K43 (P = 0.0251)
8558 *	C→G															4.1			ATP synthase 6: ATP6 (Uniprot P00846)	A11G	N (2)	0.344	Loss of stability (P = 0.0131)
8994	G→A																38.4			Syn (L)			
9557	C→T												15.2							Syn (P)			
9577 *	T→A																6.5		Cytochrome c oxidase: COX III (Uniprot P00414)	L124Q	N (4)	0.745	Gain of disorder (P = 0.0202)
9578 *	A→C																6.5			Syn (L)			

***Table 5.13. continued***

Reference Position	Variant	Percentage change in frequency (%)															Gene region	Amino acid change	SNPs&GO	MutPred		
		(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)	(M)	(N)	(O)				(P)	(Q)	Probability of deleterious mutation
9986	G→A													24.5				Cytochrome c oxidase: COX III (Uniprot P00414)	Syn (G)			
10873	T→C												75.7					ND4 NADH dehydrogenase 4: ND4	Syn (P)			
11560	A→G	87.3															Syn (W)					
12019	C→T											14.9					Syn (T)					
12091	T→C															3.4	Syn (I)					
12622	G→A											14.7					NADH dehydrogenase 5: ND5		V96I	N (7)	0.533	
12705	C→T														78.5			Syn (I)				
13346 *	C→T											3.6						A337V	D (2)	0.775		
13477	G→A								27.8								(Uniprot P03915)	A381T	N (8)	0.447		
14155	C→T			3.6													NADH dehydrogenase 6: ND6 (Uniprot P03923)	Syn (G)				
14159 *	C→G	3.1																R172P	D (3)	0.423	Loss of methylation R172 (P = 0.0305)	
14160	G→C		3.3					3.2										R172G	N (2)	0.442	Loss of methylation R172 (P = 0.0305)	
14588 *	C→A															5.4		G29V	N (6)	0.797		
15354 *	C→T												3.6				Cytochrome B: CYT B (Uniprot P00156)	T203M	N (1)	0.766		
15660 *	C→T	3.9															P305L	D (2)	0.703	Gain of catalytic residue at P305 (P = 0.0355)		

# Only the changes that were predicted significant (P < 0.05) by MutPred are listed. (\*) previously unreported variants not listed in

MitoMap.

**Table 5.14.** Susceptibility of the mtDNA regions to mutation for GBM patient tumor samples.

Reference Position	Gene region	Nucleotide positions for gene regions (bp)	Size of gene region (bp)	Number of variants detected in gene region	Normalised data		
					Mutation rate in each gene region	Mutation rate grouped for coding/non-coding regions	
66, 146, 151, 152, 153, 195, 16093, 16111, 16162, 16185, 16261, 16293, 16311, 16319, 16519	Non-coding region	D-loop	16024..576	1121	15	0.0134	0.00353
707		12s rRNA	648..1601	954	2	0.00210	
1646		tRNA valine	1602..1670	68	1	0.0147	
2648, 2923		16s rRNA	1671..3229	1558	2	0.00128	
7470		tRNA serine 1	7446..7514	68	1	0.0147	
12262		tRNA serine 2	12207..12265	58	1	0.0172	
6146, 7428*	Coding region	Cytochrome C Oxidase: COXI	5904..7445	1541	2	0.00130	0.00252
8027, 8088*		Cytochrome C Oxidase: COXII	7586..8269	684	2	0.00292	
8447*, 8491*		ATPase 8	8366..8572	206	2	0.00971	
8558*, 8994		ATPase 6	8527..9207	680	2	0.00294	
9557, 9577*, 9578*, 9986		Cytochrome C Oxidase: COXIII	9207..9990	783	4	0.00511	
10873, 11560, 12019, 12091		NADH Dehydrogenase: ND4	10760..12137	1377	4	0.00290	
12622, 12705, 13346*, 13477		NADH dehydrogenase 5: ND5	12337..14148	1811	4	0.00221	
14155, 14159*, 14160, 14588*		NADH dehydrogenase 6: ND6	14149..14673	524	4	0.00763	
15354*, 15660*	Cytochrome B: CYT B	14747..15887	1140	2	0.00175		

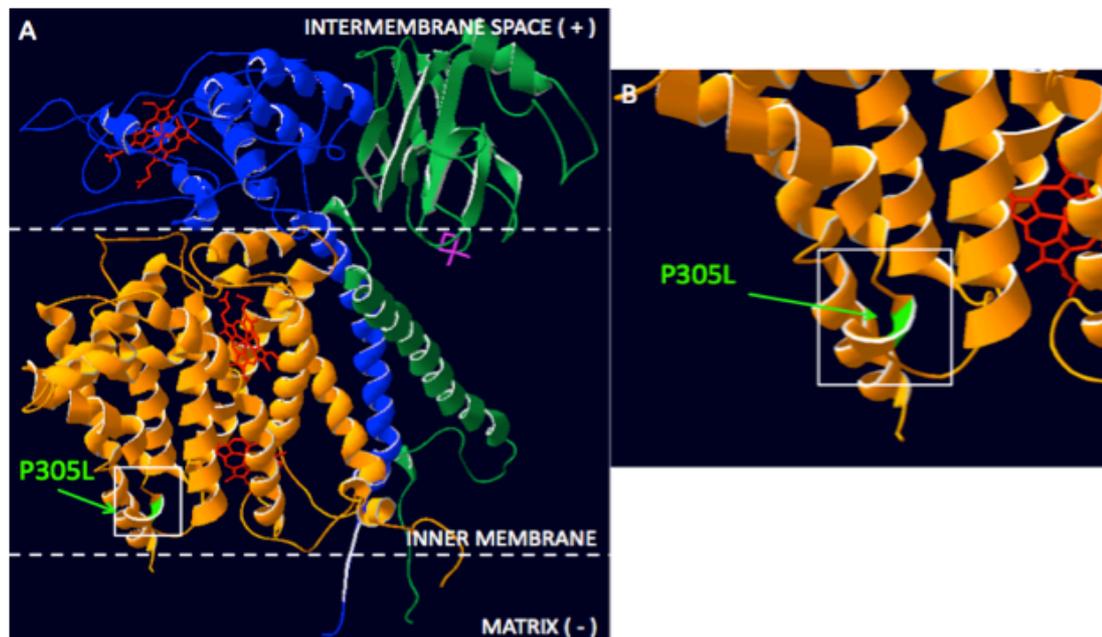
\* Novel mtDNA variants that are not already listed in the MitoMap database are indicated by asterisks.

**Table 5.15.** Fold change differences on the susceptibility of mtDNA regions to mutation, for GBM patient tumor samples. The susceptibility scores calculated for different mtDNA regions were based on Table 5.14. Values greater than 1 (black) indicate that the probability of the gene listed on the left is greater than for the gene listed across the top for the development of variants. Values in red indicate that the gene listed on the left has a lower susceptibility for the development of mutation than the gene listed across the top.

Gene region	D-loop	12s rRNA	tRNA valine	16s rRNA	tRNA serine 1	tRNA serine 2	COX I	COX II	ATP8	COX III	ND4	ND5	ND6	CYT B
D-loop		6.38	0.91	10.42	0.91	0.78	10.31	4.58	1.38	2.62	4.61	6.06	1.75	7.63
12s rRNA	0.16		0.14	1.63	0.14	0.12	1.62	0.72	0.22	0.41	0.72	0.95	0.27	1.19
tRNA valine	1.10	7.01		11.46	1.00	0.85	11.33	5.03	1.51	2.88	5.06	6.66	1.93	8.38
16s rRNA	0.10	0.61	0.09		0.09	0.07	0.99	0.44	0.13	0.25	0.44	0.58	0.17	0.73
tRNA serine 1	1.10	7.01	1.00	11.46		0.85	11.33	5.03	1.51	2.88	5.06	6.66	1.93	8.38
tRNA serine 2	1.29	8.22	1.17	13.43	1.17		13.28	5.90	1.78	3.37	5.94	7.81	2.26	9.83
COX I	0.10	0.62	0.09	1.01	0.09	0.08		0.44	0.13	0.25	0.45	0.59	0.17	0.74
COX II	0.22	1.39	0.20	2.28	0.20	0.17	2.25		0.30	0.57	1.01	1.32	0.38	1.67
ATP8	0.73	4.63	0.66	7.56	0.66	0.56	7.48	3.32		1.90	3.34	4.40	1.27	5.53
COX III	0.38	2.44	0.35	3.98	0.35	0.30	3.94	1.75	0.53		1.76	2.31	0.67	2.91
ND4	0.22	1.39	0.20	2.26	0.20	0.17	2.24	0.99	0.30	0.57		1.32	0.38	1.66
ND5	0.17	1.05	0.15	1.72	0.15	0.13	1.70	0.76	0.23	0.43	0.76		0.29	1.26
ND6	0.57	3.64	0.52	5.95	0.52	0.44	5.88	2.61	0.79	1.49	2.63	3.46		4.35
CYT B	0.13	0.84	0.12	1.37	0.12	0.10	1.35	0.60	0.18	0.34	0.60	0.79	0.23	

#### 5.4.11. Protein modeling of the GBM tumor sample variants.

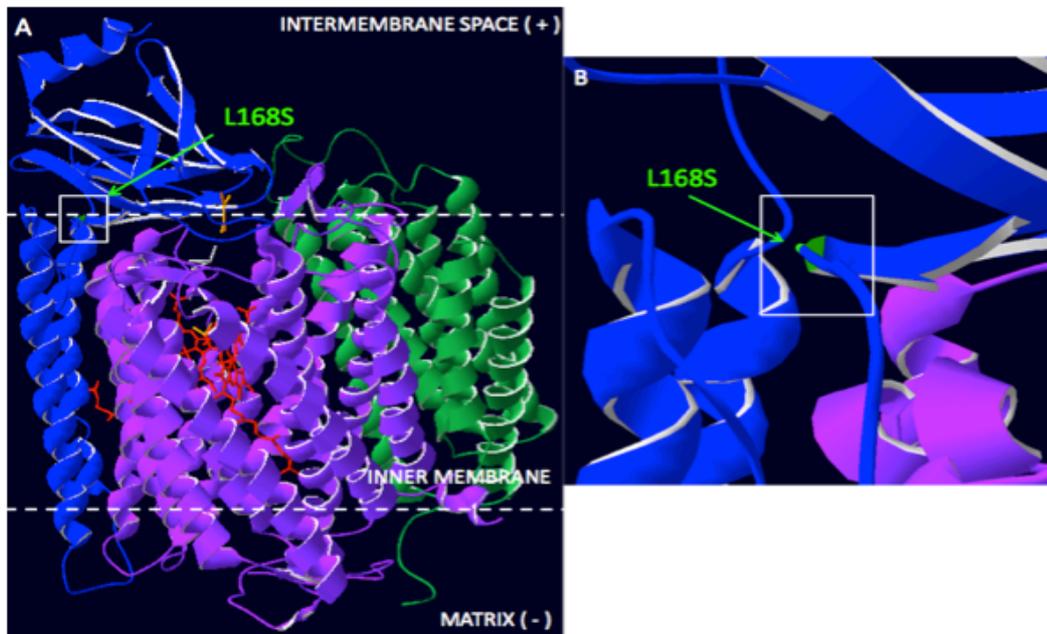
The 4 non-synonymous amino acid changes predicted to impair protein function resided within COX II, ND5, ND6 and CYT B. These represented a gain (R172P) or loss (P305L) of a proline residue in ND6 and CYT B, respectively. Protein modeling for the CYT B P305L mutation determined that it was located on the 16<sup>th</sup> transmembrane-spanning helix (**Figure 5.5**), which is distinctly separated from the other helical domains within CYT B. Its location could be a reason for the low reliability index value of 2 given to the prediction that this mutation may cause protein impairment.



**Figure 5.5.** Modeling of the P305L CYT B mutation using the bovine model.

Although the A337V mutation caused by the 13346bp variant does not involve a biochemically unfavorable amino acid substitution, the central location of the substitution within ND5 may contribute towards protein dysfunction. On the other hand, the L168S protein mutation caused by the 8088bp variant in *COX II* results not only in a biochemically incompatible substitution but also loss of a

conserved residue. This could explain the reliability index of 5 that was associated with this mutation, since leucine is more hydrophobic in nature, whereas serine is more hydrophilic. Based on protein modeling, the L168S mutation resided on a loop region immediately after the 7<sup>th</sup> beta sheet structure on the bovine COX II structure (**Figure 5.6**). If this mutation were to be located within the beta-sheet structure itself, it is likely that the reliability index given to the prediction that this mutation could cause functional impairment might increase.



**Figure 5.6.** Modeling of the L168S COX II mutation using the bovine model.

#### 5.4.12. Comparative analyses between paired GBM tumor samples

Several of the normal brain and GBM tumor samples represented paired samples, which were obtained from an individual patient who had consented to the surgical extraction of the GBM tumor and a small region of healthy tissue from

the opposite, unaffected brain hemisphere. Three pairs of samples were analysed to identify any patterns that may exist between the two samples (**Table 5.16**).

Sequencing of the healthy and tumor tissues from the first individual (sample 16) revealed 3 variants within the normal brain sample (16-1) with the *ND4* variant at position 11791bp exhibiting the highest level of mutation at 11.8% of all the variants present. On the contrary, 8 variants were identified in the tumor sample obtained from the same patient (sample 16-2). The highest level of mutation within the tumor sample was reported at 24.5% within the *COX III* variant located at 9986bp. Most of the variants within the tumor sample had higher levels of mutation than those found in the normal sample, except for the variants at positions 13346bp within *ND5* and 15354bp within *CYT B*, both of which were reported at 3.6% in the tumor sample but were absent in the normal samples.

Sequencing of the healthy and tumor tissues from the second individual (sample 22) revealed only 1 variant within *CYT B* at position 14770bp in the normal sample (22-1) at 8%. In this case, no variants were detected within the tumor sample (22-2). For the third individual (sample 20), only 1 variant at 66bp was detected in the normal brain tissue (20-1) located within the D-loop hypervariable region 2 at 3.2%. In the tumor sample of this same patient, the same D-loop variant was present at a similar frequency of 3.1%. However, there was the presence of an additional 3 variants, 1 within the D-loop at position 153bp (47.8%), another within *ATP6* at 4.1% (8558bp) and finally 1 variant within the *ND5* region at position 12705bp at 78.5%. For this third individual,

**Table 5.16.** Identification of mtDNA variants in paired samples from 3 individual GBM patients. Normal brain samples are represented by '-1' at the end of the sample number, whereas tumor samples are represented by '-2'.

	Reference Position	Variant	Percentage change in variant (%)						Gene region	Amino Acid Change	SNPs&GO	Mutpred	
			16-1	16-2	22-1	22-2	20-1	20-2				Probability of deleterious mutation	Predicted impact of the amino acid change on protein function #
Non-coding	66	G→T					3.2	3.1	D loop Hypervariable Segment 2				
	153	A→G						47.8	D loop H-strand origin				
	7470	C→T		12.1					tRNA serine 1				
Coding region	7428	G→A		17.3					COX I (Uniprot P00395)	V509M	Neutral RI 9	0.259	Gain of disorder (P = 0.0339)
	8558	C→G						4.1	ATP 6 (Uniprot P00846)	A11G	Neutral RI 2	0.344	Loss of stability (P = 0.0131)
	8630	A→G	3.7							K35 Stop			
	9557	C→T		15.2					COX III (Uniprot P00414)	Syn (P)			
	9986	G→A		24.5						Syn (G)			
	11791	C→T	11.8						ND4	Syn (L)			
	12019	C→T		14.9						Syn (T)			
	12622	G→A		14.7					ND5 (Uniprot P03915)	V96I	Neutral RI 7	0.533	
	12705	C→T					78.5			Syn (I)			
	12774	C→T	9.3							Syn (G)			
	13346	C→T		3.6						A337V	Disease RI 2	0.775	
	14770	C→A			8.0				CYT B (Uniprot P00156)	N8K	Neutral RI 1	0.446	Gain of methylation at N8 (P = 0.0039) Gain of MoRF binding (P = 0.0159) Gain of ubiquitination at N8 (P = 0.0452)
	15354	C→T		3.6					CYT B	T203M	Neutral RI 1	0.766	

# Only the changes that were predicted significant (P < 0.05) by MutPred are listed.

the presence of these extra mtDNA variants in the tumor sample arose in the key mtDNA regions. This is because defective ATP6 function prevents the efficient production of ATP and defects to ND5 are thought to be partially responsible for the onset of metastasis in tumors<sup>222</sup>.

#### **5.4.13. Variants in mtDNA depleted and recovered HSR-GBM1 cells**

To identify mtDNA variants that are important for GBM tumor biology, the traditional HSR-GBM1 cell line underwent mtDNA depletion *in vitro*, followed by recovery *in vivo*. This approach was used to determine whether mtDNA variants remained the same or differed during the establishment of a tumor from mtDNA depleted HSR-GBM1 cells. MtDNA depletion was performed using ddC so that cells contained 50% (mtDNA<sup>50</sup>), 20% (mtDNA<sup>20</sup>), 3% (mtDNA<sup>3</sup>) and 0.2% (mtDNA<sup>0.2</sup>) of their original mtDNA content. These were then injected into BALB/c nude mice, together with non-depleted HSR-GBM1 cells that were used as a control, to generate tumors.

HRM analysis was performed on the DNA from each of the tumors (**Table 5.17**), for the panel of variants identified during sequencing of the 12 GBM cell lines from **Tables 5.2 and 5.3**. 10 of the 22 tumors had 100% similarity to control HSR-GBM1 cells that had not undergone mtDNA depletion. These 10 tumors were representative of the different mtDNA depletion categories (mtDNA<sup>100</sup>, mtDNA<sup>50</sup>, mtDNA<sup>20</sup>, mtDNA<sup>3</sup> and mtDNA<sup>0.2</sup>). A further 9 tumors shared 93.8% of variants with the HSR-GBM1 control cells (2 mtDNA<sup>100</sup> tumors, 5 mtDNA<sup>50</sup> tumors, 1 mtDNA<sup>20</sup> tumor and 1 mtDNA<sup>0.2</sup> tumor), 2 tumors (1 mtDNA<sup>100</sup> and 1

**Table 5.17.** Identification of mtDNA variants in depleted HSR-GBM1 that have been recovered in BALB/c nude mice, using HRM analysis. Variants in yellow are shared across all tumors formed, and those in orange have been acquired during tumor formation, which are representative of the variants identified in the other GBM cell lines that were screened.

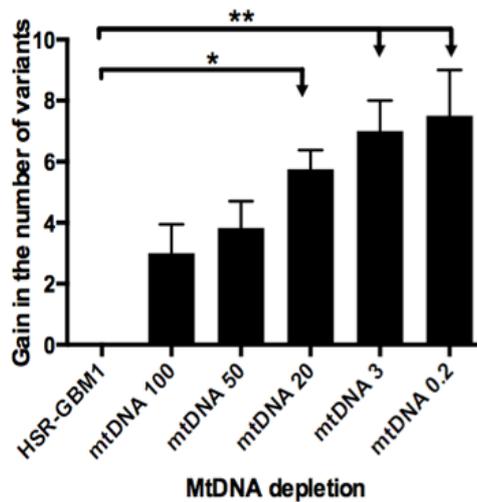
	Reference Position	Variant	Gene region	Cell line	Control tumor samples				MtDNA depleted mice tumor samples																	
				HSR-GBM1	mtDNA <sup>100</sup>				mtDNA <sup>50</sup>				mtDNA <sup>20</sup>		mtDNA <sup>3</sup>				mtDNA <sup>0.2</sup>							
Non-coding region	16186	C→T	D-loop																							
	16218	C→T																								
	194	C→T																								
	302	A→C																								
	310	T→C																								
	1386	T→C	12s rRNA																							
	2130	A→G	16s rRNA																							
	5752	A→G	Origin of L-strand replication																							
Coding region	6422	C→T	Cytochrome C Oxidase I																							
	6999	G→A																								
	8251	G→A	Cytochrome C Oxidase II																							
	8252	C→A																								
	10473	C→G	NADH dehydrogenase 4L																							
	10814	A→C	NADH dehydrogenase 4																							
	11361	T→C																								
	11512	C→A																								
	11674	C→T																								
	12101	T→C																								
	12102	C→T																								
	13061	C→A	NADH dehydrogenase 5																							
	14159	C→G	NADH dehydrogenase 6																							
	14160	G→C																								
	14426	C→T																								
15264	C→T																									
15267	C→G	Cytochrome B																								
Similarity to HSR-GBM1 (%)				-	93.8	87.5	93.8	100	100	93.8	93.8	93.8	93.8	93.8	100	100	100	93.8	81.3	100	87.5	100	100	100	93.8	100
Gain in variants to HSR-GBM1				-	1	3	6	1	4	7	2	1	4	5	4	6	7	4	6	10	8	6	7	4	6	9

 Variant present     
  Variant absent     
  Variant present (sites jointly analysed)     
  Variant absent (sites jointly analysed)

mtDNA<sup>3</sup>) shared 87.5% of variants with HSR-GBM1 cells and 1 mtDNA<sup>20</sup> tumor shared only 81.3%.

A number of variants in the non-coding region were present in all tumors formed, which included 4 variants in the D-loop region at 16186bp, 16218bp, 302bp and 310bp, and 1 variant in the 12s rRNA region at position 1386bp. Within the coding region, 1 *COX I* variant at 6999bp, 2 within *ND4* at 11361bp and 11674bp, 2 within *ND6* at 14159bp and 14160bp, and 2 within *CYT B* at positions 15264bp and 15267bp were shared by all tumor samples. Several variants were acquired during tumor formation (at positions 10473bp, 10814bp, 12101bp, 12102bp, 13061bp, 14426bp). These were representative of the variants identified within the other GBM cell lines (**Tables 5.2 and 5.3**).

Overall, there was an increase in the number of mtDNA variants that were representative of other GBM cell line variants with the duration of mtDNA depletion, followed by recovery *in vivo* (**Figure 5.7**). A significant gain in the number of variants relative to HSR-GBM1 cells was not detected until mtDNA copy number was depleted to 20% of the original numbers, which could indicate the mtDNA<sup>50</sup> stage to possess the lowest mtDNA content that is sufficient for maintaining normal function of the mitochondrion. The positive correlation that was seen in **Figure 5.7** could suggest that, with the duration of mtDNA depletion, there is an increasing predisposition for the GBM tumors to develop GBM-specific variants during re-population of mtDNA. This is likely to have arisen by the selection of mtDNA molecules after large-scale depletion, in order to drive tumor development of GBM.



**Figure 5.7.** The correlation between mtDNA depletion and the gain of mtDNA variants in HSR-GBM1 cells recovered in BALB/c nude mice. The number of variants that were acquired is relative to control HSR-GBM1 cells that had not been depleted or transplanted into mice. \*  $P < 0.05$ , \*\*  $P < 0.01$ .

#### 5.4.14. Critical mtDNA variants required for the tumor biology of GBM.

The GBM patient tumor samples were screened by HRM using the same panel of variants that were identified during Ion Torrent sequencing of the GBM cell lines (Tables 5.2 and 5.3). 8 variants were identified to be common in all human GBM tumor samples (highlighted in yellow in Table 5.18) and amongst these variants, few were shared with the HSR-GBM1 variants that were previously identified common for the establishment of a tumor in the BALB/c nude mice models (highlighted in yellow in Table 5.17).

These shared variants between human GBM tumors and those HSR-GBM1 tumors formed in the mice models included the D-loop variants at positions 302bp and 310bp, the *COX I* variant at 6999bp and the *ND4* variant at 11361bp (Table 5.18). The finding that these variants were common during the establishment of tumor formation in mice and humans, suggests that there are a certain set of variants that are required to initiate GBM tumor formation, with a small number that may initiate tumor formation in a species-dependent manner.

The variants that may be specific to tumor formation in mice includes the D-loop variants 16186bp and 16218bp, the 12s rRNA variant at 1386bp, the *ND4* variant at 11674bp, the *ND6* variants at 14159bp and 14160bp and those within *CYT B* at positions 15264bp and 15267bp. For humans, species-specific variants for tumor formation includes the D-loop variant 194bp, the *COX I* variant at 6422bp and the *COX II* variants at 8251bp and 8252bp. Variations in the tumor microenvironments in humans and mice may be reasons to explain for these differences.

**Table 5.18.** Identification of mtDNA variants in human GBM tumor samples using HRM analysis. Variants highlighted in yellow represent those that are shared across all tumor samples.

	Reference Position	Variant	GBM tumor samples												Gene region
			09-311	09-258	08-585	09-022	08-0243	08-0680	09-635	085-279	07-257	07-125	10-0214	08-745	
Non-coding region	16186	C→T	■	■	■	■	■	■	■	■	■	■	■	■	D-loop
	16218	C→T	■	■	■	■	■	■	■	■	■	■	■	■	
	194	C→T	■	■	■	■	■	■	■	■	■	■	■	■	
	302	A→C	■	■	■	■	■	■	■	■	■	■	■	■	
	310	T→C	■	■	■	■	■	■	■	■	■	■	■	■	
	1386	T→C	■	■	■	■	■	■	■	■	■	■	■	■	12s rRNA
	2130	A→G	■	■	■	■	■	■	■	■	■	■	■	■	16s rRNA
	5752	A→G	■	■	■	■	■	■	■	■	■	■	■	■	Origin of L-strand replication
Coding region	6422	C→T	■	■	■	■	■	■	■	■	■	■	■	■	Cytochrome C Oxidase I
	6999	G→A	■	■	■	■	■	■	■	■	■	■	■	■	Cytochrome C Oxidase II
	8251	G→A	■	■	■	■	■	■	■	■	■	■	■	■	
	8252	C→A	■	■	■	■	■	■	■	■	■	■	■	■	
	10473	C→G	■	■	■	■	■	■	■	■	■	■	■	■	NADH dehydrogenase 4L
	10814	A→C	■	■	■	■	■	■	■	■	■	■	■	■	NADH dehydrogenase 4
	11361	T→C	■	■	■	■	■	■	■	■	■	■	■	■	
	11512	C→A	■	■	■	■	■	■	■	■	■	■	■	■	
	11674	C→T	■	■	■	■	■	■	■	■	■	■	■	■	
	12101	T→C	■	■	■	■	■	■	■	■	■	■	■	■	
	12102	C→T	■	■	■	■	■	■	■	■	■	■	■	■	
	13061	C→A	■	■	■	■	■	■	■	■	■	■	■	■	NADH dehydrogenase 5
	14159	C→G	####	####	■	####	####	####	■	■	■	■	■	■	NADH dehydrogenase 6
	14160	G→C	####	####	■	####	####	####	■	■	■	■	■	■	
	14426	C→T	■	■	■	■	■	■	■	■	■	■	■	■	Cytochrome B
15264	C→T	■	■	■	■	■	■	■	■	■	■	■	■		
15267	C→G	■	■	■	■	■	■	■	■	■	■	■	■		

Variant present    
 Variant absent    
 Variant present (sites jointly analysed)    
 Variant absent (sites jointly analysed)    
#### Variant presence/absence confirmed in both Ion Torrent and HRM analyses

#### 5.4.15. Overview of the mtDNA variants identified across the cells and tissue samples

Although, the number of GBM tumors analysed were not sufficient to draw reliable conclusions from, so far the D-loop region was determined to contain the highest number of mtDNA variants in the non-coding region for the GBM cell lines, tumor and normal brain samples (**Table 5.19**). However, apart from the normal brain biopsy samples, the D-loop region was not the region that was most susceptible to mutation. For the GBM cell lines, this was identified as the origin of light strand replication and for the GBM tumor samples, this was tRNA serine 2. Across all cells and tissue types, the mitochondrial rRNA regions were least susceptible to the development of mutation in the non-coding region of mtDNA.

**Table 5.19.** Table showing an overview of the sequencing results from the GBM cells, tumors and normal brain tissue.

		GBM cell lines	GBM tumors	Normal brain tissue
Non-coding region	Most variants	D-loop		
	Fewest variants	12s rRNA, Origin of L-strand replication	tRNA valine, tRNA serine 1, tRNA serine 2	12s rRNA
	Most susceptible to mutation	Origin of L-strand replication	tRNA serine 2	D-loop
	Least susceptible to mutation	12s rRNA	16s rRNA	12s rRNA
Coding region	Most variants	NADH Dehydrogenase 4	Cytochrome C Oxidase III, NADH Dehydrogenase 4, NADH Dehydrogenase 5, NADH Dehydrogenase 6	Cytochrome B
	Fewest variants	NADH Dehydrogenase 4L	Cytochrome C Oxidase I, Cytochrome C Oxidase II, ATPase 8, ATPase 6, Cytochrome B	Cytochrome C Oxidase I, ATPase 8, NADH Dehydrogenase 3
	Most susceptible to mutation	NADH Dehydrogenase 6	ATPase 8	Cytochrome B
	Least susceptible to mutation	Cytochrome C Oxidase I		

Within the coding region, *ND4* harbored the most number of variants within the cell lines, whilst *ND6* was determined most susceptible to mutation within this category. Equal numbers of variants were identified at high levels of mutation in the *COX III*, *ND4*, *ND5* and *ND6* regions for the GBM tumors. However, the region most susceptible to mutation was *ATPase 8* for the tumor samples. *ND4* was shared as a region containing the most variants in both the GBM cell lines and tumor samples.

The presence of high levels of mutation in the patient tumor samples could relate to greater compromise of the ETC to function compared to, for example, normal brain tissue. Within the coding region, *CYT B* contained the highest number of mtDNA variants and was the region most susceptible to mutation in normal brain tissues. Multiple regions (*COX I*, *ATP8*, *ND3*) were also identified to contain the fewest mtDNA variants within these normal brain biopsy samples. Despite differences between the GBM cell lines, tumor and normal brain biopsy samples, all were identified to share *COX I* as the mtDNA-coding region that was least susceptible to the development of mutation (**Table 5.19**).

## **5.5. Discussion**

### **5.5.1. MtDNA variants in the GBM cell lines, tumors and normal brain tissue**

Many of the variants identified in the GBM cell lines and normal brain samples represent novel variants that have not yet been published in the MitoMap database for mtDNA mutations. The few that have been reported include those at positions 1386bp, 8251bp, 11361bp, 11674bp and 12102bp within the GBM cell lines. However, despite being reported within the MitoMap database, none of these variants have so far been associated with GBM. All mtDNA variants identified within the GBM cell lines were absent from the normal brain tissue samples, which suggest that GBM cells possess a different set of mtDNA variants that could be responsible for its tumor phenotype.

Across the GBM cell lines, GBM tumors and normal brain samples, 19 of the 22 tRNA regions together with *NDI* were revealed as mtDNA regions that were completely free of mutations. A possible reason for the absence of mutation within *NDI* could be that the protein represents one of the most conserved subunits of Complex I, as it is involved with ubiquinone binding, one of the earliest stages of OXPHOS<sup>309</sup>. Consequently, if a mutation were to arise within this region, it is likely that protein impairment would occur.

The observation that all tRNA regions within the GBM cell lines and normal brain samples have intact wild-type sequences suggests that these samples do not carry impairment to their protein translational machinery. However, variants within the 3 tRNA regions detected in the GBM tumors (tRNA valine, tRNA serine 1 and 2)

could suggest dysfunction at the level of protein synthesis for these samples. Noticeably these tRNA variants do not include the A→G substitution at position 8344bp<sup>310</sup>, nor the A→G substitution at 3243bp<sup>311</sup>, both of which have been proposed as ‘hotspot’ regions for mutations, representing the earliest discovered variants within the tRNA regions<sup>312</sup>.

### **5.5.2. Reliability of the Ion Torrent sequencing results**

None of the variants identified reside within homopolymer stretches of sequence, which have been identified as regions particularly prone to sequencing errors when using the Ion Torrent. The variant at position 16218bp is located in the region directly before a series of T residues that reside between 16218bp and 16227bp<sup>313</sup>, providing evidence against it arising as a false positive variant. Furthermore, the polycytosine-rich region within the D-loop between positions 303bp and 309bp has been reported to be susceptible to an insertion of a C nucleotide, resulting in the length heteroplasmy at this site<sup>314</sup>. Adjacent to this region, position 310bp is thought to be vulnerable to damage caused by electrophiles and free radicals<sup>315</sup>, which could explain the detection of the high levels of mutation at positions 302bp and 310bp within the GBM cell lines. Also, frequent detection of the 16519bp variant within the D-loop across the GBM cell lines, tumor and normal samples, suggests that this variant is non-tumorigenic and may be present within the general population.

### **5.5.3. The impact of GBM mtDNA variants on protein function**

Few *ND5* variants located at positions 13346bp within the GBM tumor samples and 13061bp within the GBM cell lines were predicted to impair protein function. It is

likely that these variants could cause dysfunction at the transcript and/or the protein level. Mammalian cells require a minimum of 60% transcriptional activity of *ND5* to maintain normal function. Below this threshold, the rate of protein synthesis for *ND5* becomes insufficient to sustain the proper assembly of Complex I<sup>316</sup>. In addition, protein synthesis of *ND5* is also tightly coupled to the respiration rate of Complex I<sup>316</sup>, which suggests that mtDNA variants within *ND5* may have multiple implications on protein function.

Recently, it was hypothesized that functional Complex I is important for initiating the Warburg effect and in sustaining tumor growth by promoting the adaptation of cells to conditions of hypoxia<sup>317</sup>. Gene mutations affecting *ND6* has been isolated as the single most important cause of defective protein function, as mutant transcripts of *ND6* have been associated with reduced respiratory capacity in cells<sup>318</sup>. If this hypothesis for *ND6* were true, then any variants found within the Complex I subunits, for example, mtDNA variants in *ND2*, *ND4* and *ND5*, will likely be better tolerated by the protein, and will therefore not impair, or will exert minimal disruption, to the overall function of the Complex I<sup>319</sup>.

#### **5.5.4. MtDNA variants as initiators of GBM**

MtDNA depletion of HSR-GBM1 cells and their subsequent replenishment of copy number in the BALB/c immunodeficient mice models has demonstrated that *de novo* mtDNA variants are acquired during the tumor establishment of GBM. The finding that the novel variants acquired during mtDNA replenishment are representative of the variants identified from the additional GBM cell lines, suggests that clonal

amplification of the already existing mtDNA variants does not occur during replenishment *in vivo*. It is likely that there is a common set of mtDNA mutations that are characteristic of GBM, which can be targeted for the development of therapeutic treatments or used in the clinical setting as a diagnostic tool for GBM. Based on the results from **Tables 5.17 and 5.18**, the common set of mtDNA variants identified important for GBM tumor biology were determined as D-loop variants A→C and T→C at positions 302bp and 310bp, respectively, as well as the *COX I* variant at 6999bp (G→A) and the *ND4* variant at 11361bp (T→C).

## **5.6. Conclusion**

It remains to be determined whether mtDNA variants act as a cause of cancer or if they arise as a consequence of the event<sup>320</sup>. It has been proposed by some that unfavorable mtDNA mutations could undergo negative selection in cancer, which suggests that tumor establishment is dependent on wild-type mtDNA copies<sup>321</sup>. For example, it was suggested that lower levels of heteroplasmy could be due to negative selection against deleterious mutations<sup>52</sup>. However, based on the results of this Chapter, there is evidence to support the concept that mtDNA variants could act as possible initiators of GBM.

Replenishment of mtDNA after cells had been depleted of their mtDNA appears to be essential for the establishment of a GBM tumor. The accumulation of mtDNA variants that occurs during this process is likely to interfere with essential processes, such as, protein complex assembly and the binding of factors that would result in the shuttling of electrons through the ETC. A positive correlation was reported between

the number of variants gained and the duration of mtDNA depletion for GBM cells recovered in BALB/c nude mice models. In a situation where there is defective OXPHOS, energy production would be compromised, forcing cells to utilize glycolysis as their primary pathway for ATP production. This is true for GBM cells, which respire predominately via the Warburg effect in accordance with results from other studies<sup>322</sup>, as well as the respirometry measurements obtained from Chapter 4 (**Table 4.2**).

# **Chapter 6: The Role of MtDNA in Glioblastoma Multiforme**

## **6.1. Introduction**

It was previously determined that U266 cells depleted of their mtDNA undergo changes to their expression levels for the mtDNA encoded genes, as well as those genes involved with mitochondrial biogenesis, transcription and replication (Chapter 4). In this Chapter, the HSR-GBM1 cell line was investigated to examine whether similar changes to gene expression levels can be induced. HSR-GBM1 cells were depleted in the same manner as before, using ddC. However, these cells were recovered in conditioned media in the absence of ddC. The purpose of this was to more closely mimic recovery in the tumor microenvironment, since cells cultured under such conditions would be supported by both cellular substrates and secreted factors. It has been reported that GBM conditioned media is abundant in supplies of secreted factors, such as VEGF, EGF and PDGF-AA<sup>323</sup>. Therefore, it is hoped that the recovery of the GBM cells in conditioned media will provide a more accurate measurement of gene expression changes, since it would closely represent the natural conditions in which these cells thrive.

One study that investigated the correlation between mtDNA copy number and gene expression levels in GBM cells identified that mtDNA depletion increased the expression of key mitochondrial transcription and replication factors *TFAM*, *TFB1M*, *TFB2M* and *POLG*<sup>280</sup>. Additionally, it was determined that the higher the tumor grade of the astrocytoma, the lower the mtDNA copy number, for which GBM had the highest tumor grade (IV) and also the lowest mtDNA copy number<sup>280</sup>. This association has also been reported for other cancer types, including breast<sup>324</sup> and ovarian<sup>325</sup> cancer.

On the other hand, others have suggested that changes to gene expression levels are correlated with the metabolic characteristics of cells. This hypothesis was based on a study investigating levels of transcript production in isolated mitochondria, which demonstrated that the rate of RNA synthesis was linked with changes to the mitochondrial ATP levels<sup>326</sup>. Low levels of ATP were thought to stimulate the transcription of mtDNA genes, whereas high ATP levels were thought to prevent further transcription of these mtDNA genes. These suggestions are consistent with the previous observations in the multiple myeloma U266 cells, which showed a similar compensatory effect whilst respiring predominately via glycolysis (Chapter 4).

One of the main factors that maintain the tumorigenic capacity of cancer cells is their failure to undergo differentiation<sup>29</sup>. Apart from increased tumor cell proliferation, there is also the requirement to enhance angiogenesis of the brain vasculature to support the growth of newly established tumor cells. Occasionally, brain hemorrhage and oedema arise as side effects accompanying the increased angiogenesis in GBM, due to poor maintenance of the intratumoral blood-brain barrier, which is regulated by astrocytes, a type of glial cell<sup>327,328</sup>. There is also diffuse infiltration of the astrocytes, as these have direct access to the blood vessels of the brain, encouraging metastasis<sup>329</sup>. Furthermore, their ability to infiltrate distal regions of the brain via migration along tracks of the white matter and extracellular membrane proteins of the basement membrane<sup>330,331</sup> complicates the possibility of treatment via curative surgical resectioning<sup>332</sup>, since GBM is an astrocytic glioma.

There has been increasing interest in the reprogramming of these GBM cells by differentiation to reduce the tumorigenic phenotype of these cells<sup>333,334</sup>. GBM cells can be differentiated into astrocytes and neurons. Both these matured cell types are in close communication with one another in normal individuals, as together they are essential for the proper development of synapses. However, more importantly, their modes of metabolism are believed to be coupled<sup>335</sup>.

According to the astrocyte-neuron lactate shuttle hypothesis, astrocytes preferentially metabolise glucose via glycolysis to produce lactate that is secreted into the extracellular space and subsequently taken up by the neighboring neurons. Neurons reportedly lack activity of the bi-functional enzyme 6-phosphofructo-2-kinase/ fructose-2,6-bisphosphatase isoform 3, which is required for the generation of fructose 2,6-bisphosphate during glycolysis<sup>336</sup>. However, the secretion of lactate from the astrocytes and its subsequent uptake into neurons, allows for the stimulation of the citric acid cycle<sup>337</sup>. This is because lactate can be used as a substrate in neurons to by-pass glycolysis and, as a result, neurons reportedly have a greater reliance on OXPHOS as their primary metabolic pathway<sup>338,339</sup>.

Often in clinical practice, the degree of cellular differentiation in tumor cells is important in defining the stage of tumor progression, which influences prognosis and treatment options available to the affected patient<sup>340</sup>. Gene expression analysis is one method that can be used to determine the stage of cancer progression. For example, during neuronal differentiation of human prostate

cancer cells induced by staurosporine, an inhibitor of protein kinases, there were changes to the expression levels of genes associated with cell proliferation, replication and malignancy<sup>341</sup>. Furthermore, expression of the *MYB* gene is thought to be responsible for maintaining human breast cancer cells in their undifferentiated state, which, in its absence, induces differentiation and thereby sensitizing cancer cells to apoptosis<sup>342</sup>.

## **6.2. Aims and hypothesis**

The aim of this Chapter was to first identify whether the changes to gene expression levels during mtDNA depletion of HSR-GBM1 cells, both with and without recovery in conditioned media, induces similar outcomes to the U266 cells (Chapter 4). It was hypothesized that mtDNA depletion of the HSR-GBM1 cells would result in a compensatory response, whereby there would be an up-regulation to, at least, the mtDNA-encoded genes. This effect has already been observed in the U266 cells, and has been reported in other studies<sup>326</sup>. Furthermore, this hypothesis is supported by the findings that mtDNA depleted and recovered HSR-GBM1 cells harbor greater numbers of mtDNA variants (Chapter 5), which is likely to impact on the expression levels of both the mitochondrial and nuclear-encoded genes involved with maintaining mitochondrial function. However, due to biological differences, variability in the patterns of gene expression changes during mtDNA depletion and replenishment was also expected between the two tumor types.

Second, this Chapter also aims to determine whether there is a relationship between gene expression levels and the state of differentiation in GBM cells.

Astrocytes and neurons differentiated from HSR-GBM1 cells were used for this investigation. The results would indicate whether or not the tumorigenic characteristics of HSR-GBM1 cells were lost upon differentiation. Both astrocytes and neurons were analysed since evidence suggests that there is interdependence in the metabolism of the two cell types. It was hypothesized that gene expression levels would be lower in the differentiated cells, compared to the non-differentiated HSR-GBM1 cells, since differentiated cells are likely to metabolize energy more efficiently than cancer cells.

### **6.3. Methods**

The methods used are described in the following sections: 2.1.6 to 2.1.9, 2.4.1, 2.5.1, 2.6 to 2.8.2, 2.9.2 and 2.11. The primers used during gene expression analyses for the HSR-GBM1 cells were from Table 4.1 of Chapter 4. Additional methods are described below:

#### **6.3.1. Differentiation of HSR-GBM-1 cells into neurons and astrocytes**

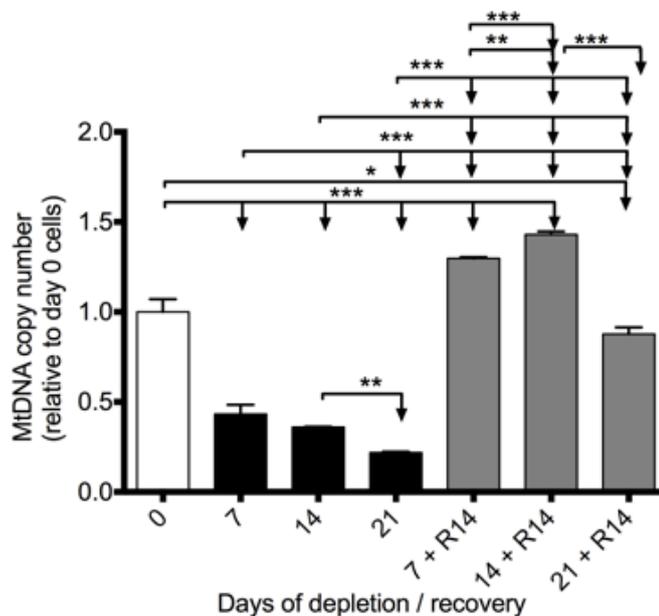
Undifferentiated HSR-GBM-1 cells were harvested and mechanically dissociated to form a single cell suspension.  $1.5 \times 10^4$  cells/cm<sup>2</sup> were seeded onto 6-well plates treated beforehand with 20µg/ml poly-L-ornithine, and 10µg/ml laminin (both: Sigma, St Louis, MO). For differentiation into neuronal cells, plated HSR-GBM-1 cells were grown in neural induction media consisting of 1x Neurobasal media, supplemented with 2mM GlutaMAX, 2% B27 supplement, 0.1mM β-mercaptoethanol (all from Gibco), 10ng/ml Brain Derived Neurotrophic Factor (BDNF; Millipore), and 100µM of Adenosine 3', 5'-cyclic monophosphate (cAMP; Sigma). For differentiation into astrocytes, induction media comprised of 1x High Glucose DMEM, supplemented with 2mM GlutaMAX, 1% N2 supplement, 0.1mM β-mercaptoethanol and 2% FBS (all from Gibco).

Media was changed every 2 days for both astrocytes and neuronal cells, and due to differences in their rates of proliferation, were passaged once the cultures reached approximately 80% confluency.

## 6.4. Results

### 6.4.1. MtDNA content in HSR-GBM1 cells recovered in conditioned media

HSR-GBM1 cells were depleted of their mtDNA and recovered in conditioned media. Results revealed that there was a progressive decline in the mtDNA copy number of HSR-GBM1 cells during mtDNA depletion to day 21 (**Figure 6.1**). After 7 days of mtDNA depletion, HSR-GBM1 cells lost over half of their total mtDNA copy numbers. However, the rate of decline in copy number was less pronounced from day 7 of mtDNA depletion onwards.



**Figure 6.1.** MtDNA content in HSR-GBM1 cells depleted of their mtDNA and recovered in conditioned media. Non-depleted HSR-GBM1 cells are shown in white, those cells that have undergone mtDNA depletion are displayed in black and those that have undergone recovery after depletion are displayed in grey. 'R14' represents 14 days of mtDNA replenishment. \*  $P \leq 0.05$ , \*\* $P \leq 0.01$ , \*\*\* $P \leq 0.001$

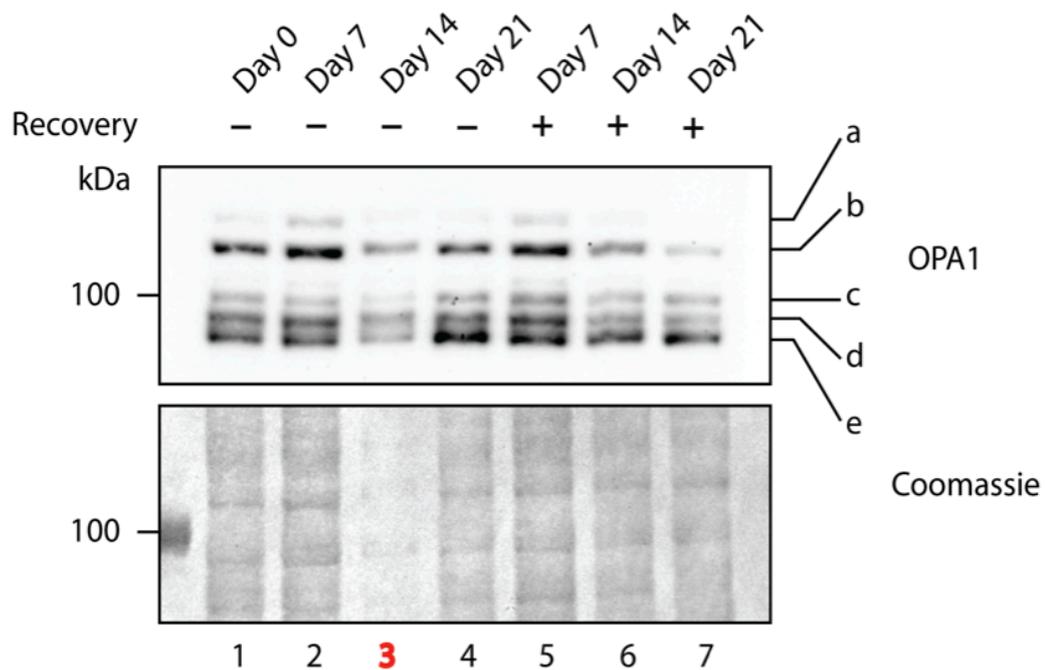
After 14 days of recovery in conditioned media, following 7 days of mtDNA depletion, mtDNA copy number increased significantly above day 0 cells. There were further increases seen for cells recovered for 14 days after 14 days of mtDNA depletion. Noticeably, 21 days of mtDNA depletion saw an increase in mtDNA copy numbers after 14 days of mtDNA replenishment, which did not exceed levels reported for day 0 HSR-GBM1 cells. This pattern likely suggests that prolonged mtDNA depletion of HSR-GBM1 cells perturbs a feedback mechanism that prevents efficient mtDNA replenishment. To test this hypothesis, gene expression analyses were performed on mtDNA depleted HSR-GBM1 cells. However, before doing so, the levels of apoptosis in the HSR-GBM1 cells during mtDNA depletion and recovery were determined.

#### **6.4.2. Determining the levels of apoptosis in mtDNA depleted and/or recovered HSR-GBM1 cells**

Western blot analysis was performed to assess OPA1 protein expression during the course of mtDNA depletion, both with and without recovery in conditioned media, for HSR-GBM1 cells. As described previously in section 4.4.11, processing of the OPA1 protein occurs during apoptosis, which is independent of dysfunction to glycolysis or OXPHOS activity.

Results demonstrated that the long isoform of OPA1 was stable throughout mtDNA depletion, and the cells did not undergo much apoptosis, even by day 21 of mtDNA depletion with recovery for 14 days, as indicated by the processing of the long OPA1 isoform to the shorter isoform (**Figure 6.2**). Comparing these outcomes to the mtDNA copy number results (**Figure 6.1**), show that the loss of

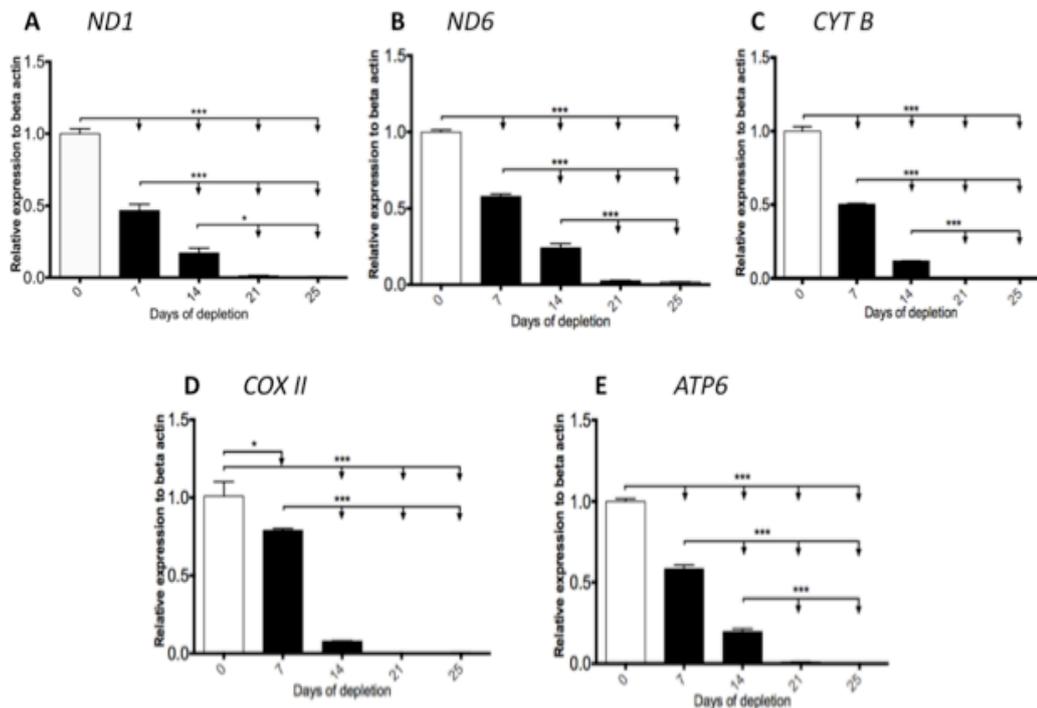
mtDNA copies up to day 21 of depletion does not itself result in apoptosis. Instead, it is likely that the accumulation of mtDNA variants during the replenishment of copy number after a period of prolonged mtDNA depletion, as identified from Chapter 5, could be responsible for causing apoptosis following the selection of mutant mtDNA copies.



**Figure 6.2.** Western blot analysis for OPA1 expression in mtDNA depleted and/or recovered HSR-GBM1 cells. Processing of the long isoform of OPA1 (b) into its short isoforms (d and e) is shown, with greater expression of the shorter isoform representing high levels of apoptosis in these cells. Lane 3 contained an insufficient concentration of protein and was therefore excluded from the analysis. ‘ - ‘ represents mtDNA depletion without recovery, whereas ‘ + ‘ represents mtDNA replenishment for 14 days in HSR-GBM1 conditioned media. Data were collected by Dr Michael Baker at the Bio21 Institute, Melbourne.

### 6.4.3. MtDNA gene expression changes in depleted HSR-GBM1 cells

For all mtDNA-encoded genes, there was the progressive loss of expression throughout mtDNA depletion (**Figure 6.3**).



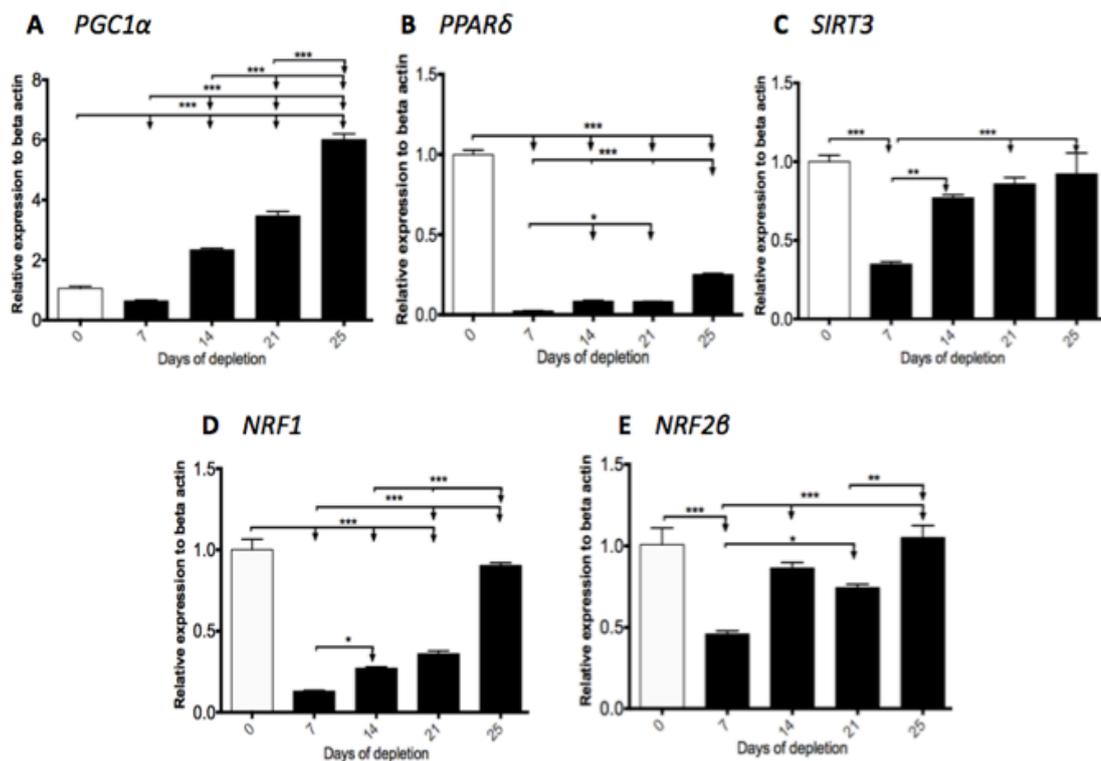
**Figure 6.3.** Expression analysis on key mtDNA genes in depleted HSR-GBM1 cells. Complex I genes (A) *ND1* and (B) *ND6*, Complex III gene (C) *CYT B*, Complex IV gene (D) *COX II* and Complex V gene (E) *ATP 6*.

With the exception of *COX II*, all other mtDNA-encoded genes decreased their expression levels to half of that for the day 0 HSR-GBM1 cells by day 7 of mtDNA depletion. By day 21 of depletion, all mtDNA-encoded genes had negligible levels of expression.

### 6.4.4. Expression of the upstream mitochondrial regulators

All the key genes associated with mitochondrial biogenesis exhibited progressive increases to their levels of expression during mtDNA depletion of HSR-GBM1

cells (**Figure 6.4**). *PGC1 $\alpha$*  demonstrated the greatest fold increase during depletion, which was almost 6 times higher than at day 25 of mtDNA depletion when compared to day 0 HSR-GBM1 cells. For *SIRT3*, *NRF1* and *NRF2 $\beta$* , expression levels significantly decreased during the initial 7 days of mtDNA depletion, which preceded a subsequent rise in expression levels through to day 25 of depletion, reaching levels that matched, or were close to levels of day 0 cells. *PPAR $\delta$*  was the only gene whose expression levels decreased significantly by day 7 of mtDNA depletion, but failed to increase to levels a third of that detected for day 0 cells.



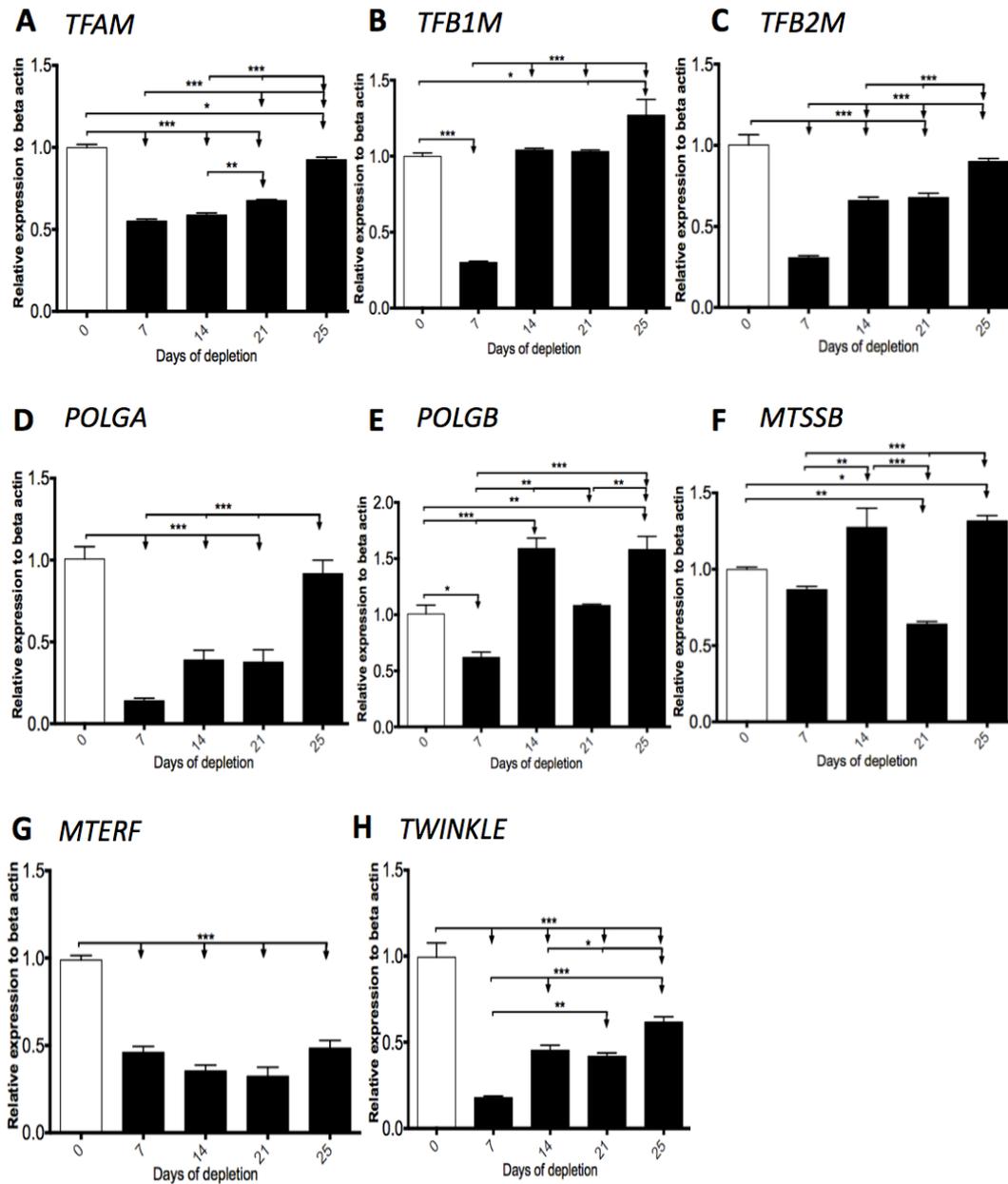
**Figure 6.4.** Gene expression of the upstream mitochondrial regulators in depleted HSR-GBM1 cells. Mitochondrial biogenesis genes (A) *PGC1 $\alpha$* , (B) *PPAR $\delta$*  and (C) *SIRT3*. Nuclear respiratory factors (D) *NRF1* and (E) *NRF2 $\beta$* .

The inability to up-regulate *PPARδ* levels of expression during mtDNA depletion could be compensated for by the significant increases in *PGC1α*. This is because *PGC1α* activity is regulated by *PPARδ*, which means that if *PGC1α* is able to self up-regulate its expression levels in the absence of sufficient *PPARδ* levels, then the changes in the expression levels of *PPARδ* during mtDNA depletion of HSR-GBM1 cells would be of lesser importance than the other mitochondrial biogenesis genes.

#### **6.4.5. Expression of the mitochondrial transcription and replication factors**

Generally, mtDNA depletion stimulated up-regulation of all the mitochondrial transcription and replication factors, with the exception of *MTERF* (**Figure 6.5G**). During the early stages (day 7) of mtDNA depletion, all genes displayed significant decreases in expression levels, which preceded a rise in levels with prolonged mtDNA depletion. For the genes of the mitochondrial transcription factors *TFAM*, *TFB1M* and *TFB2M* (**Figure 6.5A-C**), there were differences in the rates at which the expression levels were up-regulated. For example, *TFAM* exhibited a gradual rise in expression levels from day 7 of mtDNA depletion onwards, whereas *TFB1M* and *TFB2M* showed minimal changes in their expression levels between days 14 and 21 of mtDNA depletion. For *TFB1M* and *TFB2M* significant increases in gene expression levels followed, which were apparent at day 25 of mtDNA depletion. Unlike *TFB2M*, *TFB1M* expression levels increased above day 0 HSR-GBM1 cells by day 25 of mtDNA depletion.

The change in patterns of expression for *POLGA* (**Figure 6.5D**) were similar to that displayed for *TFB2M*. Furthermore, *POLGB* (**Figure 6.5E**) expression



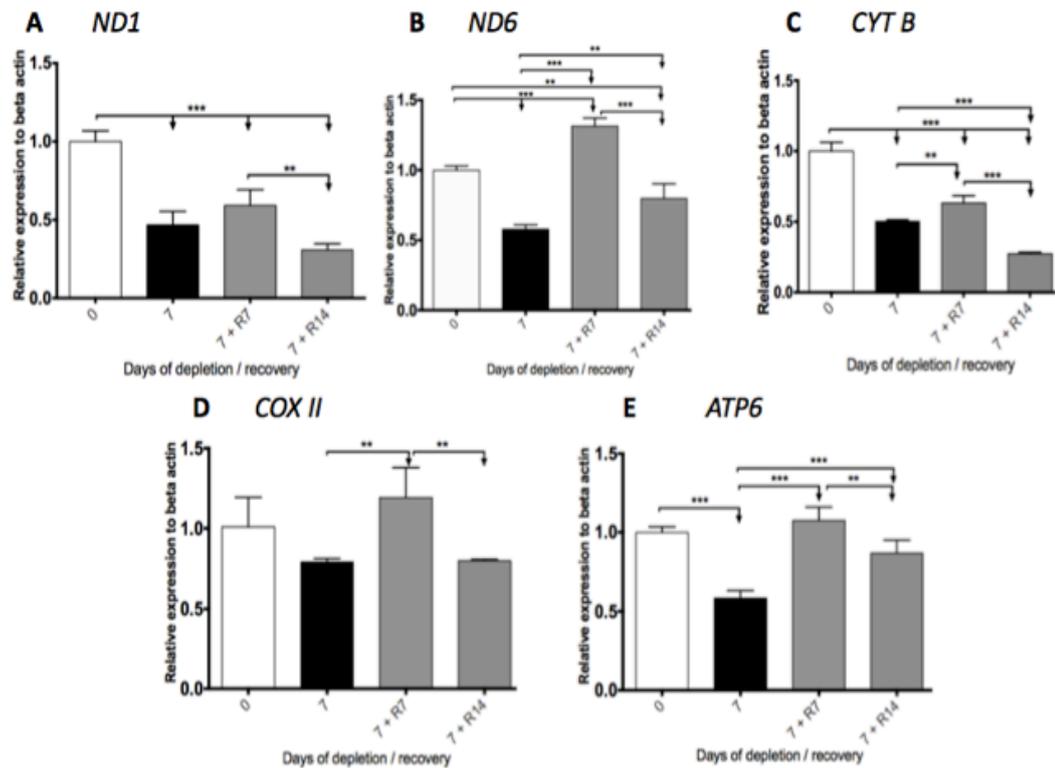
**Figure 6.5.** Expression of the mitochondrial transcription and replication factors. Regulators of transcription (A) *TFAM*, (B) *TFB1M*, (C) *TFB2M*, and regulators of mtDNA replication (D) *POLGA*, (E) *POLGB*, (F) *MTSSB*, (G) *MTERF* and (H) *TWINKLE*.

patterns shared similarities with *MTSSB* (**Figure 6.5F**) during mtDNA depletion, since both show a transient but significant decrease in expression levels at day 21 of mtDNA depletion. Similar to *MTERF*, the ability to up-regulate *TWINKLE* expression levels appeared to be compromised during mtDNA depletion (**Figure 6.5H**), as *TWINKLE* levels reached only 50% of those reported for day 0 cells by day 25 of mtDNA depletion.

Down-regulation of *MTERF* expression during mtDNA depletion ensures the up-regulation of other mitochondrial transcription and replication factors under conditions of mtDNA stress, since the MTERF protein functions to terminate the transcription of mitochondrial genes. Together these outcomes suggest that the mtDNA replication factors are not damaged during the course of mtDNA depletion.

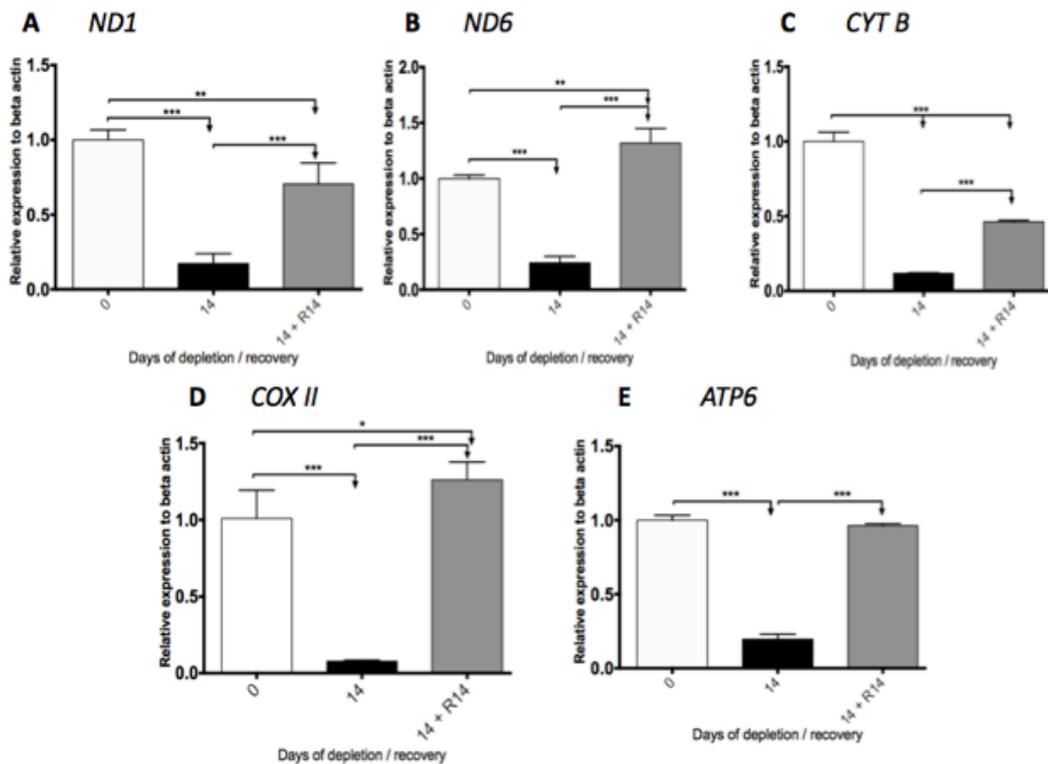
#### **6.4.6. Expression of the mtDNA genes during recovery of HSR-GBM1 cells**

For the mtDNA-encoded genes, only *ND6* (**Figure 6.6B**) had significant increases to its expression levels above those reported for day 0 HSR-GBM1 cells, at day 7 of recovery. Nonetheless, peak expression levels for all the mtDNA-encoded genes during mtDNA replenishment also occurred at this time (at R7). Cells recovered for 14 days exhibited a reduction in the detectable levels of expression for all these genes relative to day 7 mtDNA recovered cells.



**Figure 6.6.** Expression of the mtDNA genes during recovery of 7-day depleted HSR-GBM1 cells. Depleted cells were replenished in conditioned media. Complex I genes (A) ND1 and (B) ND6, Complex III gene (C) CYT B, Complex IV gene (D) COX II and Complex V gene (E) ATP 6. 'R' represents recovery in days.

On the other hand, mtDNA depletion of HSR-GBM1 cells for 14 days followed by 14 days of recovery in conditioned media showed that, for all mtDNA-encoded genes, there were significant increases to their relative gene expression levels when compared against the 14-day mtDNA depleted cells (**Figure 6.7**). Only ND6 (**Figure 6.7B**) and COX II (**Figure 6.7D**) were able to up-regulate their gene expression levels higher than those reported for day 0 HSR-GBM1 cells.

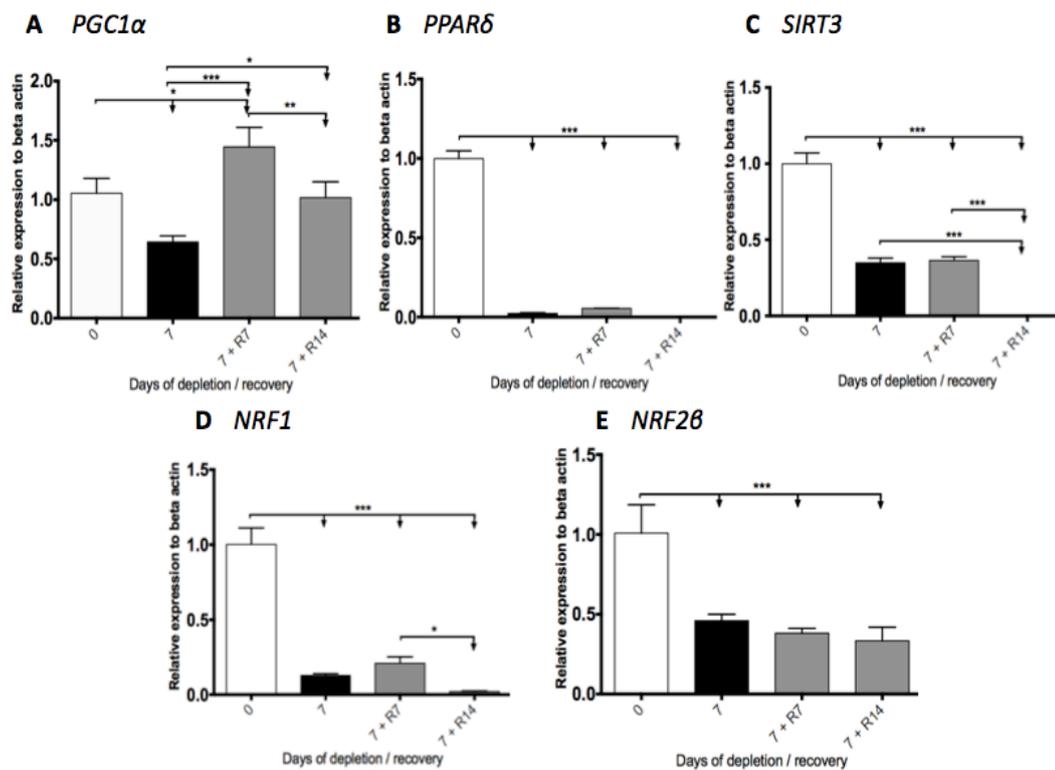


**Figure 6.7.** Expression of the mtDNA genes during recovery of 14-day depleted HSR-GBM1 cells. Depleted cells were replenished in conditioned media. Complex I genes (A) ND1 and (B) ND6, Complex III gene (C) CYT B, Complex IV gene (D) COX II and Complex V gene (E) ATP 6. ‘R’ represents recovery in days.

Overall, comparing 14-day mtDNA depleted cells with the results obtained from 7-day mtDNA depleted cells, it can be seen that increasing the duration of mtDNA depletion to 14 days resulted in the enhanced efficiency of the mitochondrial transcription machinery to up-regulate the expression of mtDNA-encoded genes.

#### 6.4.7. Expression of the upstream mitochondrial regulators during recovery

For all mitochondrial biogenesis and nuclear respiratory factor genes, with the exception of *PGC1 $\alpha$*  (Figure 6.8A), there were insignificant changes to gene expression levels between 7-day mtDNA depleted HSR-GBM1 cells and those cells replenished of their mtDNA copy number for 7-days in conditioned media (Figures 6.8B-E).

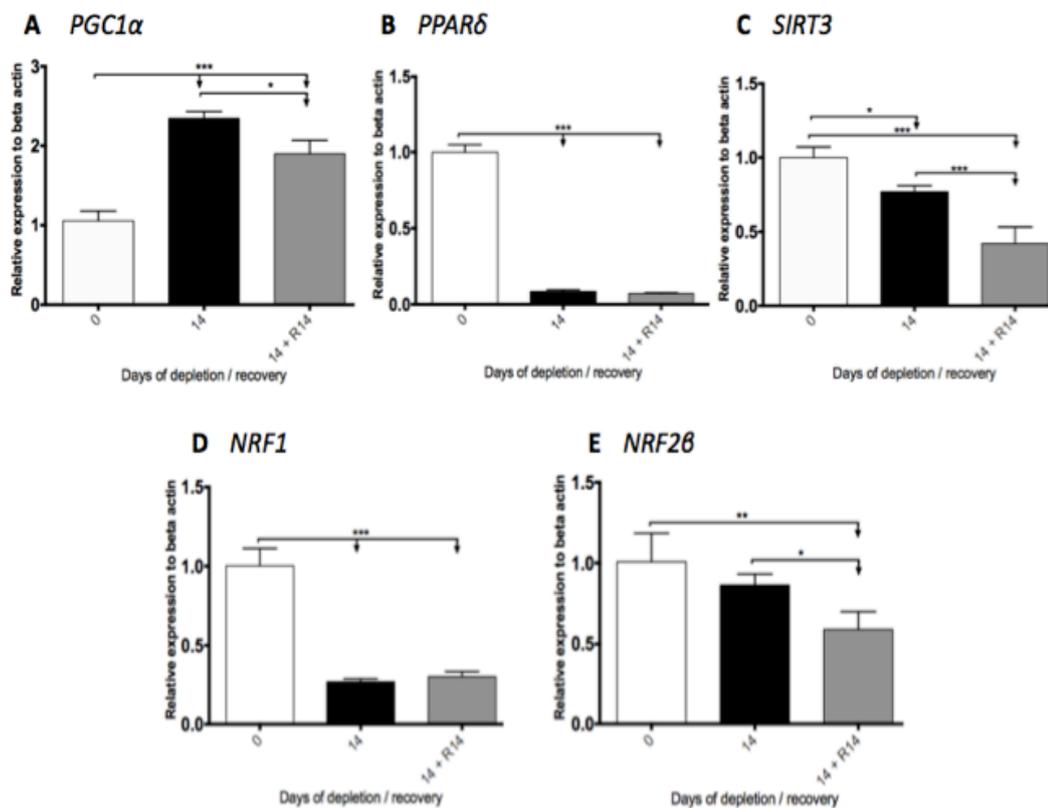


**Figure 6.8.** Expression of the upstream mitochondrial regulators during recovery of 7-day depleted HSR-GBM1 cells. Mitochondrial biogenesis genes (A) *PGC1 $\alpha$* , (B) *PPAR $\delta$*  and (C) *SIRT3*. Nuclear respiratory factors (D) *NRF1* and (E) *NRF2 $\beta$* . 'R' represents recovery in days.

*PPAR $\delta$*  was the only gene that had negligible levels of gene expression at day 7 of mtDNA depletion, which remained the same throughout the course of mtDNA replenishment (Figure 6.8B). *SIRT3* (Figure 6.8C), *NRF1* (Figure 6.8D) and

*NRF2* $\beta$  (Figure 6.8E) all had gene expression levels below 40% of day 0 HSR-GBM1 cells, whereas *PGC1* $\alpha$  increased its transcript levels above day 0 cells at day 7 of mtDNA replenishment. By day 14 of recovery in conditioned media, *PGC1* $\alpha$  expression decreased to levels that matched day 0 HSR-GBM1 cells.

After 14 days of mtDNA depletion with recovery in conditioned media, there were further decreases to the levels of expression for the mitochondrial biogenesis and the nuclear respiratory factor genes (Figure 6.9). Only *NRF1* was able to up-regulate gene expression relative to 14-day mtDNA depleted cells,



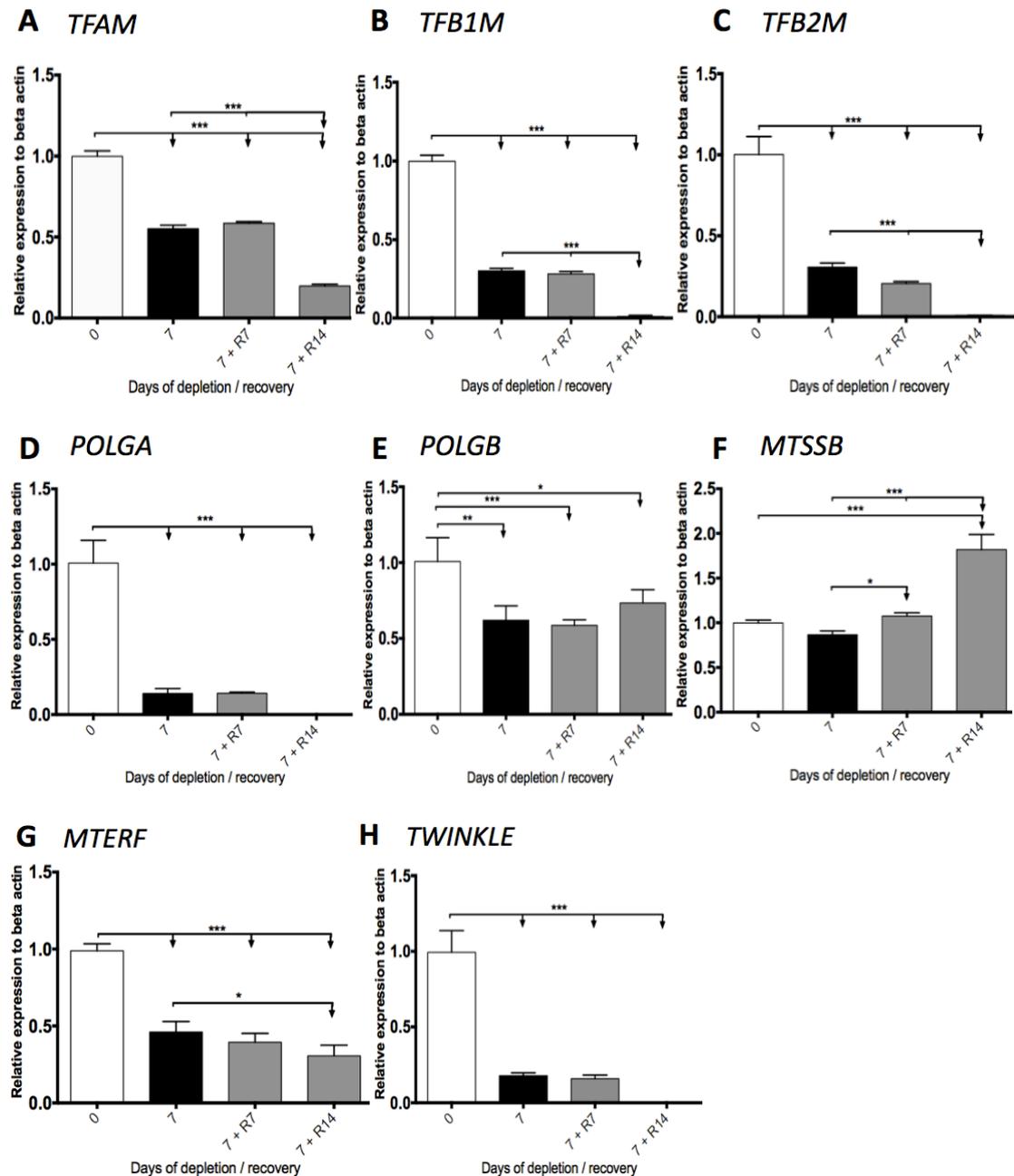
**Figure 6.9.** Expression of the upstream mitochondrial regulators during recovery of 14-day depleted HSR-GBM1 cells. Mitochondrial biogenesis genes (A) *PGC1* $\alpha$ , (B) *PPAR* $\delta$  and (C) *SIRT3*. Nuclear respiratory factors (D) *NRF1* and (E) *NRF2* $\beta$ . 'R' represents recovery in days.

although this increase was statistically insignificant (**Figure 6.9E**). *PGC1 $\alpha$*  (**Figure 6.9A**), *SIRT3* (**Figure 6.9C**) and *NRF2 $\beta$*  (**Figure 6.9E**) all had decreased levels of gene expression at day 14 of recovery in conditioned media relative to the mtDNA depleted cells. Similar to cells recovered of their mtDNA after 7 days of mtDNA depletion, *PPAR $\delta$*  displayed almost negligible levels of expression, which did not increase even after 14 days of recovery (**Figure 6.8B**).

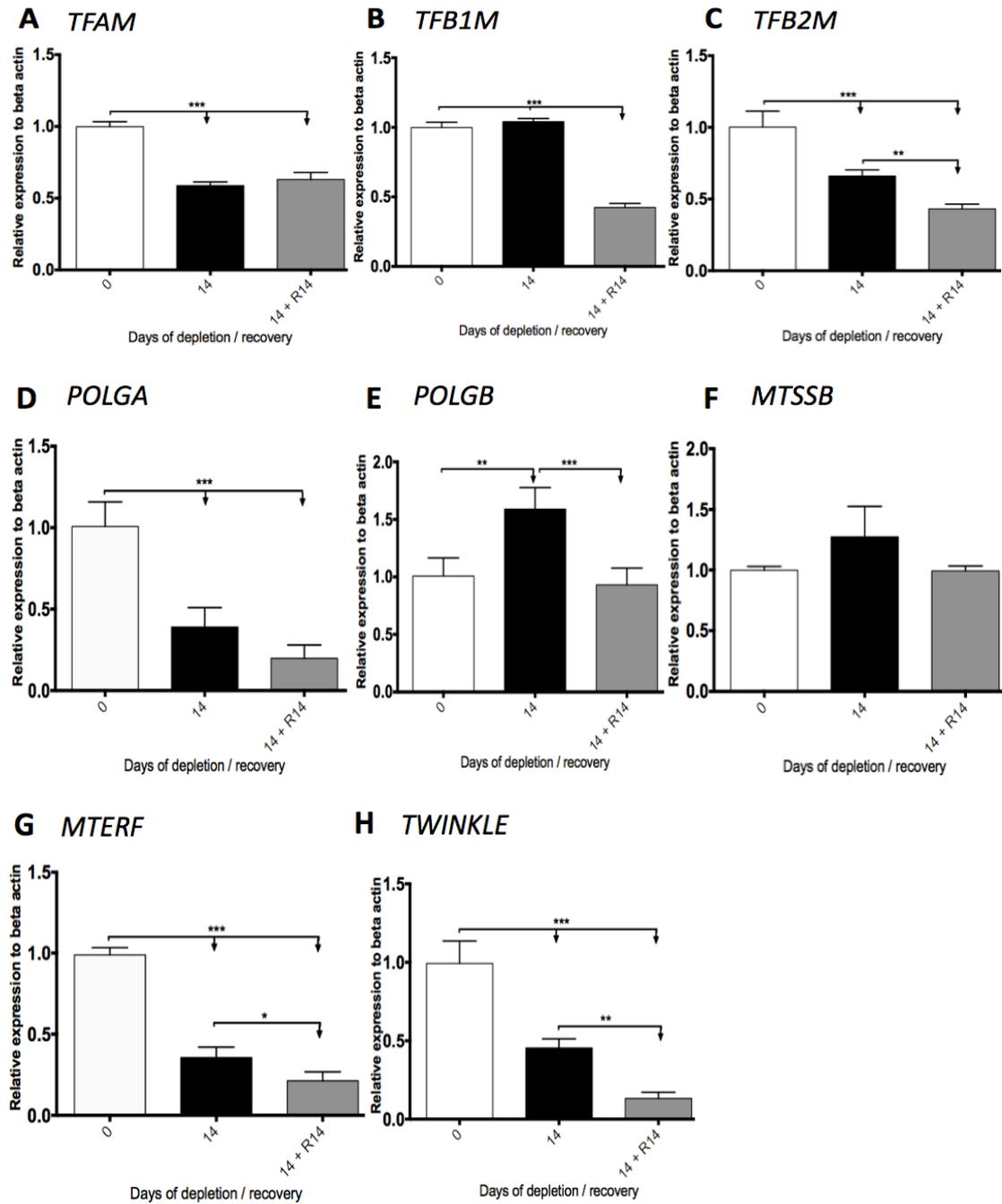
#### **6.4.8. Expression of the mitochondrial transcription and replication factors during recovery**

*TFB1M* expression during recovery after 7 days of depletion was similar to the expression patterns for *TFB2M* (**Figure 6.10C**), *POLGA* (**Figure 6.10D**) and *TWINKLE* (**Figure 6.10H**). Only *POLGB* (**Figure 6.10E**) and *MTSSB* (**Figure 6.10F**) were up-regulated during recovery after 7 days of mtDNA depletion.

After 14 days of mtDNA depletion, all mitochondrial transcription and replication genes, with the exception of *TFAM* (**Figure 6.11A**), had decreased gene expression levels after 14 days of mtDNA replenishment relative to the 14-day mtDNA depleted cells. For the majority of these genes the decreases in the levels of gene expression were significant, except for *POLGA* (**Figure 6.11D**) and *MTSSB* (**Figure 6.11F**). Changes to the expression levels of *TFAM* showed slight but insignificant increases at day 14 of recovery in conditioned media compared to 14-day mtDNA depleted cells. These changes to the expression of the mitochondrial transcription and replication factors confirm that depletion for 14 days prevents the efficient recovery of expression for these genes when cells are recovered for 14 days in conditioned media.



**Figure 6.10.** Expression of the mitochondrial transcription and replication genes during recovery of 7-day mtDNA depleted HSR-GBM1 cells. Regulators of transcription (A) *TFAM*, (B) *TFB1M*, (C) *TFB2M*, and regulators of mtDNA replication (D) *POLGA*, (E) *POLGB*, (F) *MTSSB*, (G) *MTERF* and (H) *TWINKLE*. ‘R’ represents recovery in days.

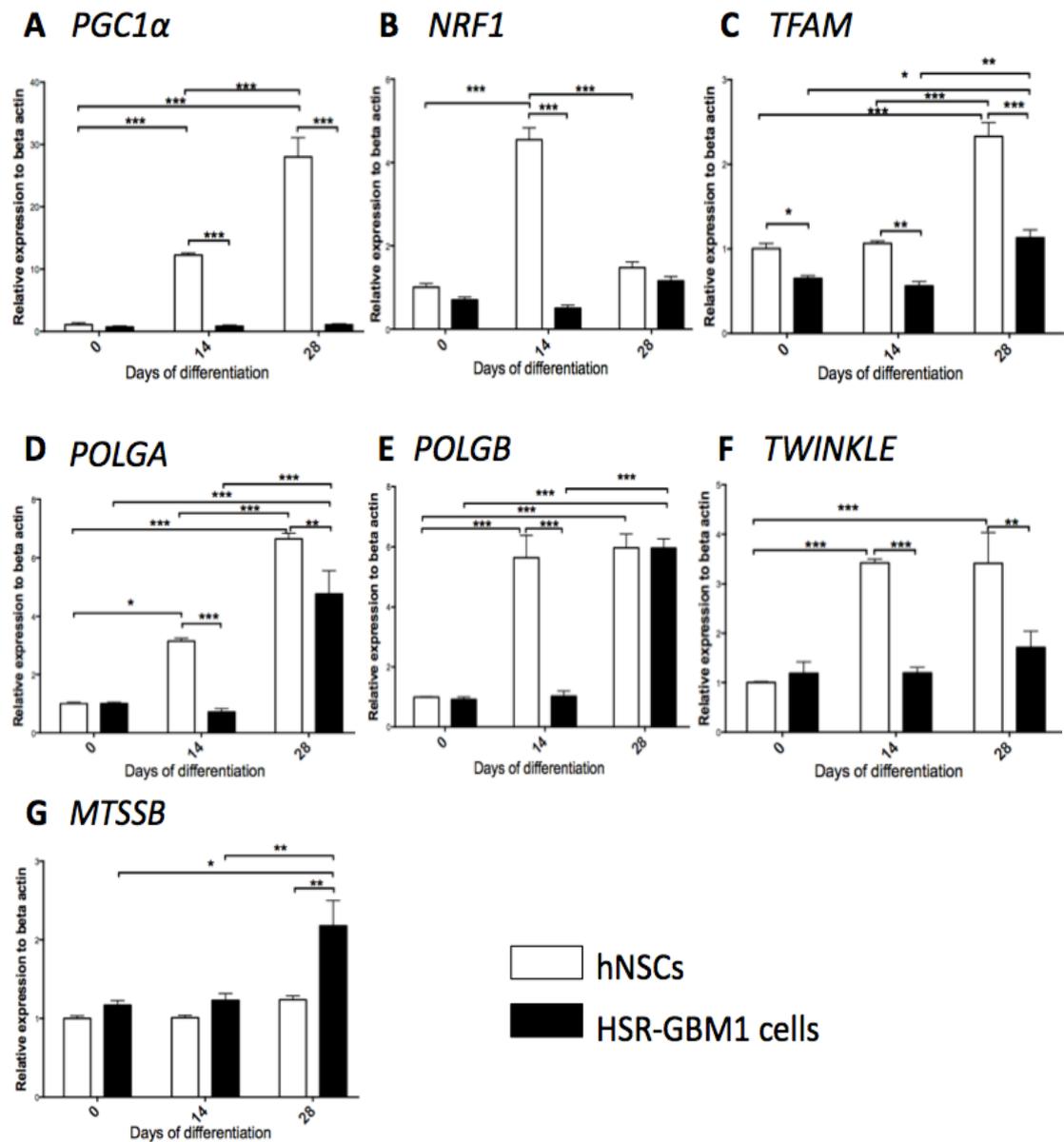


**Figure 6.11.** Expression of the mitochondrial transcription and replication genes during recovery of 14-day mtDNA depleted HSR-GBM1 cells. Regulators of transcription (A) *TFAM*, (B) *TFB1M*, (C) *TFB2M*, and regulators of mtDNA replication (D) *POLGA*, (E) *POLGB*, (F) *MTSSB*, (G) *MTERF* and (H) *TWINKLE*. ‘R’ represents recovery in days.

#### 6.4.9. Changes in gene expression in differentiated astrocytes

During the differentiation of hNSCs to astrocytes there were progressive increases to the expression levels of *PGC1 $\alpha$*  (**Figure 6.12A**) and *POLGA* (**Figure 6.12D**), both of which are key genes for the maintenance of mtDNA. *POLGB* (**Figure 6.12E**) and *TWINKLE* (**Figure 6.12F**) showed early up-regulation of expression to peak levels by day 14 of hNSC-astrocyte induced differentiation, whereas there was late up-regulation of *TFAM* (**Figure 6.12C**) expression, which peaks at day 28 of differentiation. For the expression levels of *NRF1* (**Figure 6.12B**) there was a significant but transient increase during hNSC differentiation, which occurs at day 14 and decreases back to levels similar to day 0 hNSCs by day 28 of differentiation. Insignificant changes to gene expression levels were reported for *MTSSB* throughout the duration of hNSC differentiation (**Figure 6.12G**).

For HSR-GBM1 cells, there were insignificant changes to the expression of genes *PGC1 $\alpha$* , *NRF1* and *TWINKLE* during differentiation into astrocytes. The expression profiles for *TFAM*, *POLGA*, *POLGB* and *MTSSB* demonstrated that significant changes in gene expression did not occur until day 28 of astrocyte differentiation.

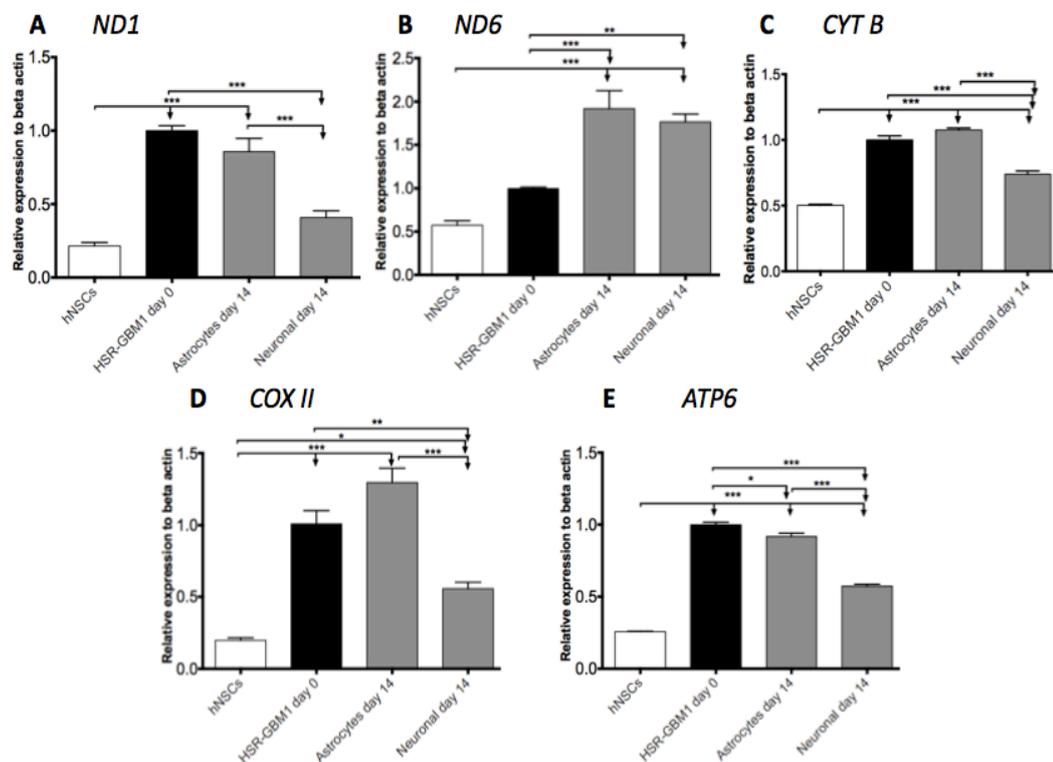


**Figure 6.12.** Expression analysis for the mitochondrial maintenance genes in astrocytes. The key genes involved with mitochondrial biogenesis, transcription and replication factors are shown during differentiation of the hNSCs and HSR-GBM1 cells to astrocytes, after 14 and 28 days in astrocyte-induction media.

#### 6.4.10. Expression of the mtDNA genes after differentiation into astrocytes and neurons

Gene expression changes in astrocytes and neurons differentiated from HSR-GBM1 cells were examined. The purpose of this was to determine whether there were differences in the gene expression levels between the two mature cell types.

*ND1* (Figure 6.13A) and *ATP6* (Figure 6.13E) shared similar patterns of gene expression after 14 days of differentiation into astrocytes and neurons. For *ND6* (Figure 6.13B), *CYT B* (Figure 6.13C) and *COX II* (Figure 6.13D), expression levels



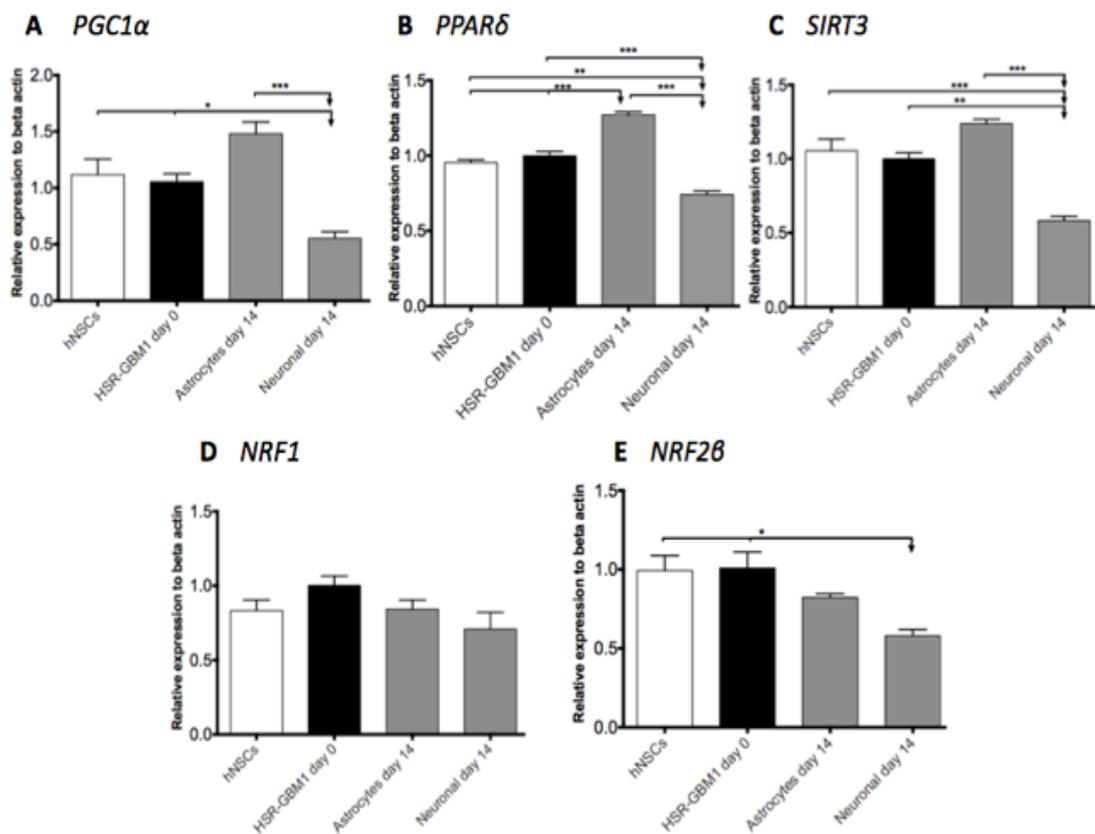
**Figure 6.13.** Expression of the mtDNA genes after differentiation into astrocytes and neurons. HSR-GBM1 cells were differentiated for 14 days into the two cell types. Complex I genes (A) *ND1* and (B) *ND6*, Complex III gene (C) *CYT B*, Complex IV gene (D) *COX II* and Complex V gene (E) *ATP 6*.

in the astrocytes were higher than for the undifferentiated HSR-GBM1 cells. However, for all mtDNA-encoded genes, expression levels in the neuronal cells were lower than observed in the astrocytes.

#### 6.4.11. Expression of the upstream mitochondrial regulators after differentiation into astrocytes and neurons

Astrocytes and neurons shared similar patterns of expression for the mitochondrial biogenesis factors *PGC1 $\alpha$* , *PPAR $\delta$*  and *SIRT3* (Figure 6.14A-C).

There were increases to the levels of *PGC1 $\alpha$* , *PPAR $\delta$*  and *SIRT3* in the astrocytes



**Figure 6.14.** Expression of the upstream mitochondrial regulators after differentiation into astrocytes and neurons. Differentiation was performed for 14 days on HSR-GBM1 cells cultured in induction media. Mitochondrial biogenesis genes (A) *PGC1 $\alpha$* , (B) *PPAR $\delta$*  and (C) *SIRT3*. Nuclear respiratory factors (D) *NRF1* and (E) *NRF2 $\beta$* .

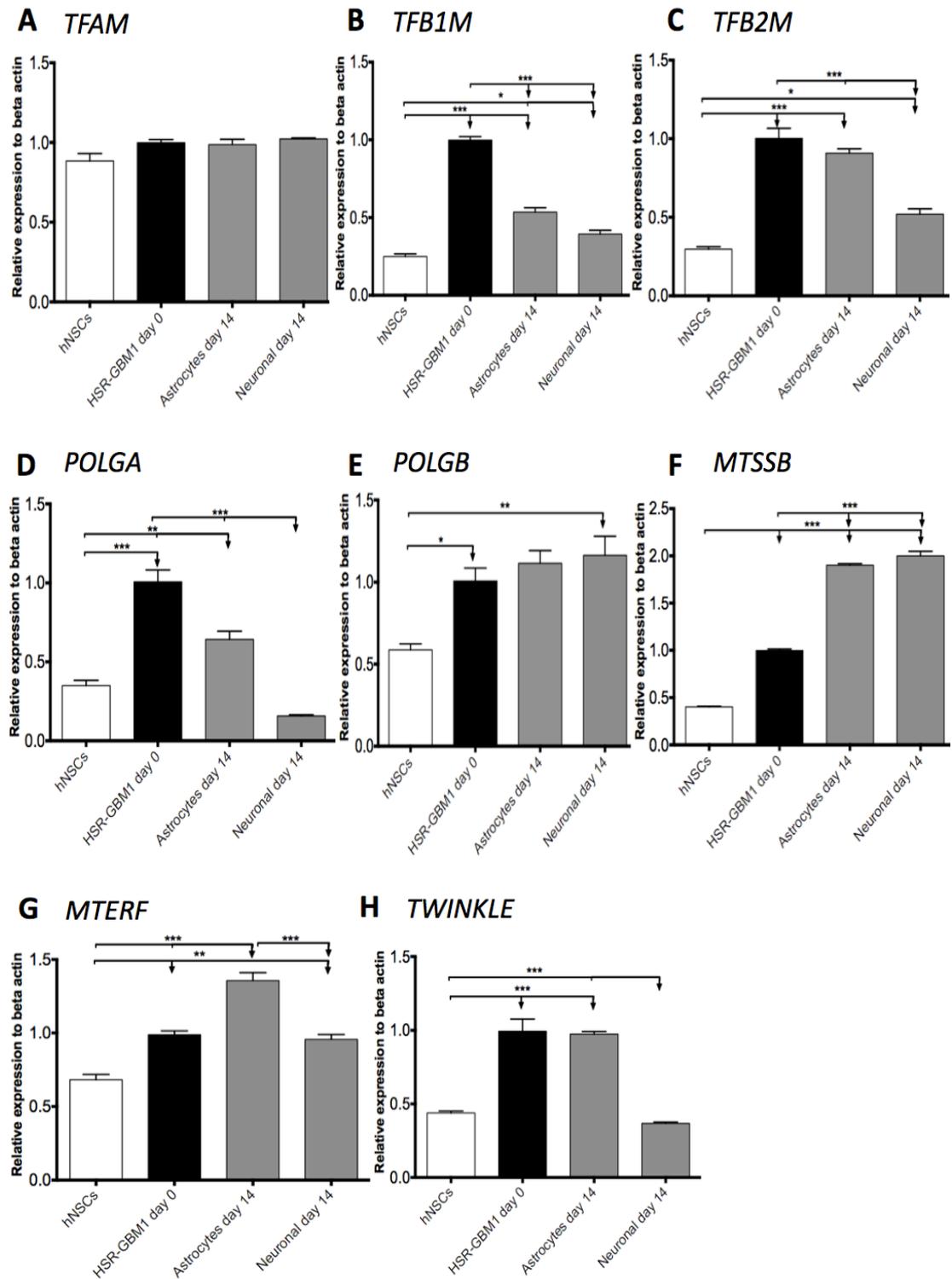
above that of undifferentiated HSR-GBM1 cells, whereas expression of these genes in the neuronal cells decreased as low as 50% of levels below HSR-GBM1 cells. On the other hand, *NRF1* and *NRF2 $\beta$*  (**Figure 6.14D-E**) both showed that astrocytes and neurons have reduced expression levels of these genes relative to undifferentiated HSR-GBM1 cells.

#### **6.4.12. Expression of the mitochondrial transcription and replication genes after differentiation into astrocytes and neurons**

Gene expression levels of *TFB1M* (**Figure 6.15B**), *TFB2M* (**Figure 6.15C**), *POLGA* (**Figure 6.15D**) and *TWINKLE* (**Figure 6.15H**) in 14-day astrocytes and neurons were reduced compared to undifferentiated HSR-GBM1 cells. However, both *POLGB* (**Figure 6.15E**) and *MTSSB* (**Figure 6.15F**) increased their expression levels in these differentiated cells relative to their undifferentiated counterpart. The up-regulation in *MTSSB* expression in astrocytes and neurons was more pronounced compared to the increases in *POLGB* expression relative to undifferentiated cells. *TFAM* was the only gene whose expression did not change following 14 days of differentiation into astrocytes and neurons (**Figure 6.15A**).

Gene expression changes observed in 14-day astrocytes and neurons suggest that there is limited replication of mtDNA in these cells, because expression of the key mitochondrial transcription and replication genes were reduced compared to undifferentiated HSR-GBM1 cells. It is likely that differentiated cells metabolise glucose more efficiently, and as such, OXPHOS is likely to predominate in these mature astrocytes and neuronal cell types for the generation of ATP. This

hypothesis is supported by recent data proposing that the astrocyte-neuron lactate shuttle is not required for OXPHOS to take place<sup>343</sup>, thereby opposing the traditional hypothesis that, astrocyte metabolism is closely coupled to neurons, to supply these cells with lactate as a substrate for stimulating OXPHOS activity.



**Figure 6.15.** Expression of the mitochondrial transcription and replication genes after differentiation into astrocytes and neurons. Regulators of transcription (A) *TFAM*, (B) *TFB1M*, (C) *TFB2M*, and regulators of mtDNA replication (D) *POLGA*, (E) *POLGB*, (F) *MTSSB*, (G) *MTERF* and (H) *TWINKLE*.

## **6.5. Discussion**

### **6.5.1. Gene expression changes in HSR-GBM1 cells recovered in conditioned media**

Results from day 7 and day 14 mtDNA depleted HSR-GBM1 cells recovered of their mtDNA copy numbers in conditioned media, suggested that prolonged mtDNA depletion for 14 days, was generally less effective for increasing the expression levels of key mitochondrial biogenesis and nuclear respiratory factor genes. Moreover, recovery of gene expression levels was only effective during short-term culture (7 days) in conditioned media after 7 days of mtDNA depletion.

It was previously found that mtDNA depletion itself does not damage the mitochondrial transcription and replication factors (section 6.4.4). Therefore the inability to rescue gene expression levels of these factors during recovery, especially after day 14 of mtDNA depletion, is likely to be due to damage to a feedback mechanism between the mitochondria and nucleus in coordinating the regulation of expression for these genes during mtDNA replenishment. Whilst mtDNA copy number analysis (**Figure 6.1**) showed the up-regulation of copy number during recovery after 7 and 14 days of depletion, there could be a delay in the nuclear transcription of the mitochondrial associated genes in response to increases in mtDNA copy number, which further supports the hypothesis for a feedback mechanism.

A key observation during recovery in conditioned media after 7 days of mtDNA depletion, was the inability of HSR-GBM1 cells to up-regulate the expression

levels of *TFAM* (Figure 6.10A), *TFB1M* (Figure 6.10B) and *TFB2M* (Figure 6.10C), which contradicted results from another study that reported up-regulation in the expression of these genes during depletion<sup>280</sup>. One could interpret these results to indicate that recovery in conditioned media may not provide sufficient metabolites and factors to support the activity of the mitochondrial transcriptional machinery. Upon collection of the conditioned media from a culture of control HSR-GBM1 cells, the cells may have exhausted certain supplies of nutrients within the media in order for these cells to produce additional factors that can be secreted into the surrounding media to support its growth and proliferation. Without fresh supplies of certain nutrients that cannot be produced by the cells alone, the culture of cells in conditioned media may not be optimal to maintain the expression levels of genes that are important for the maintenance of mtDNA.

Although the regulation of mitochondrial gene expression levels is thought to be largely dependent on the metabolic activities of cells, evidence has emerged to suggest that post-transcription modifications can occur independently of the cell's metabolic status<sup>344</sup>. This could affect, for example, rates of transcript degradation, and may provide an alternative explanation for the differences between the recovery of gene expression levels in HSR-GBM1 and U266 cells.

#### **6.5.2. Gene expression analysis in astrocytes and neurons differentiated from HSR-GBM1 cells**

Within the subventricular zone of the brain, there is a population of astrocyte-like neural stem cells, which under abnormal regulation may give rise to GBM<sup>303,345</sup>.

Alternatively, de-differentiation of mature glial cells can lead to the same outcome, forming a GBM tumor<sup>346</sup>. These biological findings that connect hNSCs with HSR-GBM1 cells are the primary reasons why the hNSCs were used as a non-tumorigenic control when analyzing for gene expression differences between day 14 and day 28 astrocyte-differentiated cells (**Figure 6.12**).

It has been reported that, unlike astrocytes, neurons are less able to adapt to changes in energy metabolism. This is because inhibiting mitochondrial respiration in neurons using nitric oxide, induced depletion of ATP levels and led to cellular apoptosis<sup>347</sup>, whereas astrocytes responded to the same event by increasing glycolysis in an attempt to prevent decreases to ATP levels, therefore avoiding apoptosis<sup>347</sup>. The inability of neurons to adapt to metabolic change is further emphasized by their undergoing apoptosis when the rates of glycolysis<sup>348</sup> or glycogen synthesis are increased<sup>349</sup>. For these reasons, neuronal cells were not analysed for changes in gene expression levels during the course of differentiation from HSR-GBM1 cells.

Changes in gene expression for the key factors associated with mitochondrial biogenesis, transcription and replication, comparing days 14 and 28 of astrocyte differentiation from the hNSCs and HSR-GBM1 cells (**Figure 6.12**) revealed that for most genes, except *MTSSB*, the hNSCs had higher gene expression levels compared to astrocytes differentiated from HSR-GBM1 cells. A possible explanation for this could be that astrocytes derived from hNSCs may develop certain characteristics leading towards a tumorigenic phenotype, since as

mentioned previously, astrocyte-like hNSCs are capable of developing into GBM tumor *in vivo*<sup>303,345</sup>.

## **6.6. Conclusion**

MtDNA depletion of HSR-GBM1 cells shows similar results to mtDNA depletion of U266 cells. However, overall, HSR-GBM1 cells appear to up-regulate gene expression levels less efficiently compared to U266 cells, which might be variable depending on the tumor type examined. The increase in the number of mtDNA variants that were gained during recovery of mtDNA depleted HSR-GBM1 cells may be a contributing factor towards the reduced ability of these cells to recover their gene expression levels as high as that previously observed for U266 cells. HSR-GBM1 cells also appear to have more stable expression of the long isoform of OPA1 compared to U266 cells, when undergoing mtDNA depletion, both with and without mtDNA recovery.

The observation that day 14 astrocytes and neurons had reduced expression of the key replication genes could indicate high efficiency for glucose metabolism, and therefore OXPHOS is likely to predominate in these mature cell types. It has been suggested by other studies that cells with high levels of oxygen consumption have reduced ability to form tumors<sup>350-352</sup>. Therefore, lower gene expression levels in astrocytes derived from HSR-GBM1 cells could also reflect their loss of tumorigenic properties during the differentiation process. Therefore, mtDNA is important in the initiation of a GBM tumor.

## **Chapter 7: General discussion**

### **7.1. The role of mtDNA in cancer stem cells and in tumor biology**

One of the most widely accepted concepts for explaining cancer initiation relates to the presence of cancer stem cells that may be responsible for sustaining tumor growth within a heterogenous population of tumor cells<sup>353</sup>. These cancer stem cells retain some of the properties of pluripotency and represent a rare population of cells within the tumor. Due to their high proliferative capacity, cancer stem cells are more likely the cause of tumor relapse after chemotherapy treatment, since the clonal expansion of one drug-resistant cell could give rise to the re-establishment of a tumor<sup>354</sup>.

From the results of this thesis, mtDNA has been identified to modulate the tumorigenic phenotypes of cancer stem cells from both multiple myeloma and GBM. Support for this was provided by evidence showing the acquisition of mtDNA variants during mtDNA replenishment (Chapters 3 and 5), as well as changes to the expression levels of key genes involved with its tumor biology, mitochondrial biogenesis, transcription and replication during mtDNA depletion and recovery (Chapters 4 and 6). These findings could complement the current cancer stem cell hypothesis, which states that a tumor arises as a single mutated cell that has gained a proliferative advantage to establish a tumor<sup>355</sup>, since the cause of the mutation could arise at the mtDNA level.

However, the alternative method by which a tumor may arise could be through the de-differentiation of mature cell types to form undifferentiated stem cells, via the accumulation of genetic alterations overtime<sup>356</sup>. This was demonstrated to occur in intestinal epithelial cells, as a result of enhanced signaling via the NF-

κB pathway and the increased activation of Wnt<sup>357</sup>. In terms of changes induced to the mitochondria during the process of cellular de-differentiation, it is likely that energy production within the mitochondria will change to meet the energy demands imposed by the cancer cells. This is because several studies have been able to demonstrate that mitochondria within human induced pluripotent stem cells (iPSCs) can alter certain properties associated with mitochondrial gene expression, mtDNA copy number, morphology, ATP and lactate production levels to match those that are characteristic of early stem cells<sup>358</sup>. This effectively ‘reprograms’ the mitochondrion to match the metabolic demands of the cell type<sup>359</sup>. Either way, the ability of the mitochondria to adapt to different cellular conditions and energy requirements may be key to the formation of a tumor, including those of multiple myeloma and GBM.

## **7.2. MtDNA variants**

MtDNA variants reported in this thesis have only included those of base substitutions. This is because the Ion Torrent platform is not ideal for the detection of small insertions and deletions (indels), consisting of 1 or 2 nucleotides. During the variant selection procedure, many of the indels were detected at very low levels of less than 10, out of a genome coverage of at least 1500 or more sequencing reads. It was likely that these had arisen as sequencing errors, with the majority identified as residing within homopolymer sequence regions, which is a known problem for the Ion Torrent sequencing platform<sup>181</sup>. There was also the absence of large-scale deletions from all sequencing results, which would have been less likely to represent false positive results. In this thesis, only the mtDNA variants that are likely to represent true variants are

included, which means that there may be additional mtDNA variants that are not reported, since the variant selection criteria is also quite stringent.

In Chapter 3, the GC content of a gene sequence was explored as a possible factor that could influence mutation rates across the mtDNA. However, the GC content of a sequence region is not the only factor that could influence susceptibility of a region to mutation. It was suggested by others to contribute approximately 10% towards variability in the rate of substitutions across the human genome<sup>360</sup>. Differences in the rate at which these regions evolve could also be a factor.

A comprehensive study reported that whilst the rate of substitution for synonymous amino acid changes are approximately the same for all mtDNA coding genes, differences are observed for non-synonymous mutations<sup>361</sup>. The *cytochrome c oxidase* region exhibited the slowest rate of substitution, whereas *ATP8* appeared the fastest in terms of undergoing evolutionary change<sup>361</sup>. Moreover, the fold change for the evolutionary rate of small rRNA was reported to be 20 times greater in mitochondria than in the nucleus, whereas for tRNAs, this fold change was 100 times greater<sup>361</sup>. Therefore, the natural rates of mtDNA evolution are likely influence the susceptibility of a region to mtDNA mutations.

### **7.3. Compensatory mechanisms during mtDNA depletion**

Changes to the levels of ATP production are often associated with morphological alterations of the mitochondria<sup>362</sup>. Depletion of mtDNA is therefore likely to induce a transient protective mechanism to prevent the complete loss of ATP

production. One possibility for mediating this effect involves the fusion of mitochondria, which permits the exchange of mitochondrial components to minimize the stresses that may be imposed upon individual mitochondrion, and thereby partially supporting mitochondrial function during periods of stress<sup>363</sup>.

An alternative mechanism to mitochondrial fusion has been reported in plants, where the stress of cultivation conditions can induce the up-regulation of expression for the nuclear-encoded  $F_0$  subunit, which is a key component of the ATP synthase complex<sup>364</sup>. After its translocation into the mitochondria, ATP production levels increase, a situation that emphasizes the bi-directional communication between the nucleus and mitochondria. Other theories have proposed that mitochondria or mtDNA can be mobilized intercellularly via cytoplasmic extensions<sup>365</sup>, a concept that was developed to prove the transfer of plastids across plant cells via cell-wall junctions<sup>366</sup>. However, for both these mechanisms, whether these occur in humans currently remains elusive.

#### **7.4. The involvement of mtDNA in the initiation of GBM**

Sequencing of paired samples of GBM provide an important insight into understanding whether some mtDNA variants are specific to tumors or are non-discriminative between tumor and normal tissues. From Chapter 5, the results demonstrated that tumor samples generally have greater numbers of mtDNA variants, which are also reported to have high levels of mutation. For all but one mtDNA variant (at 66bp, Table 5.16), the nucleotide substitutions were specific to either the normal or tumor tissue.

Earlier studies have also examined paired glioma samples. In 2001, Kirches et al. sequenced 17 GBM samples for comparison against their corresponding blood samples<sup>367</sup>. However, the level of accuracy from their sequencing method was not as high as other studies that utilise next generation sequencing as a primary approach to identifying variants. From this study, Kirches reported variants in only 7 of the 17 tumors analysed. Later in 2002, the same group published a short report to extend their analyses, sequencing 10 pairs of astrocytomas and their corresponding blood samples<sup>368</sup>. However, they reported the absence of any sequence variant present in these 10 paired samples. Moreover, in another study that focused on mtDNA mutations in mouse brain tumors, the absence of pathogenic variants across the entire mitochondrial genome was reported in both induced and spontaneous cases of tumor development<sup>369</sup>. Together these studies emphasise the importance of the sequencing techniques used for the identification of mtDNA variants that will provide highly accurate and meaningful data.

A genome-wide expression screen of GBM tumor tissue and their adjacent normal counterparts has been performed, which demonstrated that both these tissue types possessed similarities as well as differences in their patterns of gene expression<sup>370</sup>. As anticipated, it was found that genes participating in angiogenesis were up-regulated in the tumor tissue, alongside genes associated with cell proliferation and formation of the extracellular matrix. Genes that were down-regulated in the tumor sample, included those that were essential for development of the nervous system. In adjacent normal tissue, it was discovered that, unexpectedly, similar to the GBM cells, there was the enhanced expression

of genes involved with tumor malignancy, including those participating in cell motility. Gene expression of tumor suppressors was also down-regulated in the normal tissue. These observations led to the suggestion that perhaps there is 'leakage' of genes from the tumor cells that can invade the surrounding normal tissue. If this were true, then it remains possible for defective mitochondrial genes, such as those involved with mitochondrial biogenesis, transcription and replication to 'leak' and promote tumor establishment in neighboring normal cells. However, results from this thesis collectively support the role of mtDNA in the tumor biology of both GBM and multiple myeloma.

#### **7.5. The initiation of cancer from the transformation of normal cells**

The concept of gene leakage from tumor cells to normal cells counteracts the hypothesis that perhaps normal cells fuel tumor growth via the secretion of various factors. This was proposed in 2012, where a study demonstrated that the loss of Pten, a tumor-suppressor<sup>371</sup>, in the tumor stroma (normal cells) resulted in the down-regulation of microRNA, miR-320<sup>372</sup>. As a consequence, there was an increase in the secretion of oncogenic factors that function to enhance proliferation and malignancy of neighboring stromal cells. Normal cells are then transformed into cancer cells, thereby promoting tumor growth and malignancy<sup>372</sup>.

Earlier theories related to the initiation of cancer from normal cells include the reverse Warburg effect proposed by Pavlides et al in 2009<sup>373</sup>. This hypothesis suggested that tumor stromal cells were encouraged to undergo transformation by the adjacent tumor cells, into wound-healing stromal cells, such that these

stromal cells will predominately utilise aerobic glycolysis (Warburg effect) for ATP production. The secretion of the metabolites pyruvate and lactate can then be taken up and used by the neighboring cancer cells during the citric acid cycle, and the ATP produced via OXPHOS in the cancer cells can then be used to further promote cellular proliferation and therefore tumor growth<sup>373</sup>. Inhibitors of lactate transport are thought to target these stromal cells and induce their apoptosis via accumulation of intracellular lactate levels<sup>373</sup>. Glycolysis inhibitors such as 2-DG and dichloro-acetate, which cannot be metabolized by cells, were also suggested as a therapeutic approach to target these transformed stromal cells<sup>374</sup>.

#### **7.6. The correlation between mtDNA copy number and gene expression levels**

During the progressive depletion of mtDNA, all factors, regardless of mitochondrial or nuclear origin, experienced an increase in the relative expression levels proportional to the duration of mtDNA depletion, up until a certain point. However, up-regulation of these factors does not necessarily correspond to increased mtDNA copy number. For example, POLG activity is dependent on the function of supporting factors during mtDNA replication, such as TWINKLE and MTSSB, both of which must precede activity of POLG to permit the multi-subunit complex to carry out its function of synthesizing nascent mtDNA strands, complementary to the mitochondrial template strand. Therefore, the absence of a single factor participating in mitochondrial transcription and replication could be rate-limiting for the processes of mtDNA transcription and replication.

Results from a protein expression study performed in HeLa cells complements some of the gene expression results that were reported in Chapters 4 and 6 during mtDNA depletion and recovery. The protein expression study, led by Seidel-Rogol and Shadel, identified that the key mitochondrial transcription proteins TFAM and mtRNA polymerase had similar expression profiles during mtDNA depletion and replenishment<sup>375</sup>. They reported that recovery of expression for both factors to day 0 levels occurred at a rate slower than the recovery of mtDNA copy number<sup>375</sup>. It was suggested that during mtDNA depletion, the cell attempts to prevent the accumulation of transcripts as a resource reserve. Under normal situations in the absence of mtDNA depletion, the expressions of TFAM and mtRNA polymerase are generally in excess of the required transcription levels. More specifically, 15 molecules of TFAM have been estimated to be present per mtDNA copy in human cells<sup>376</sup>, which is thought to be optimal for the initiation of transcription from the light strand promoter region<sup>377</sup>. Significant decline in the mtDNA transcription levels mediated by TFAM is believed to occur only when there is a 10-fold decrease to the TFAM/mtDNA copy number ratio<sup>378</sup>. Therefore, it is possible that, even during mtDNA depletion, low levels of TFAM may still be sufficient to support mtDNA transcription from the origin of light strand promoter region.

### **7.6.1. Multiple myeloma**

For multiple myeloma cells, it was identified in Chapter 4 that recovery of mtDNA copy number after 7 days of mtDNA depletion, was significantly higher than day 0 cells at day 14 of recovery. Cells depleted for 14 days of their

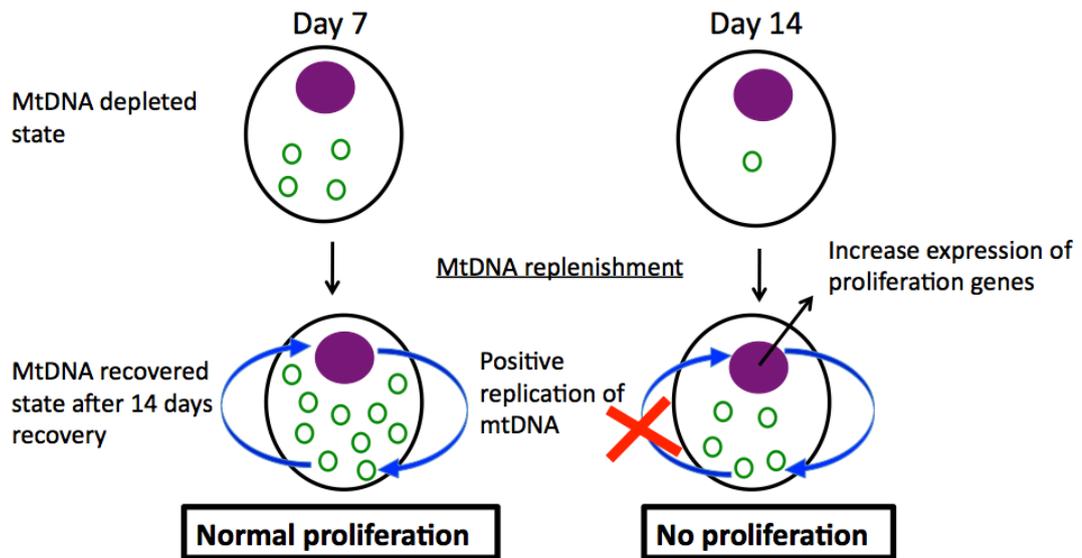
mtDNA, were not able to effectively increase their mtDNA copy numbers even after 14 days of recovery, suggesting recovery was impaired after prolonged exposure to mtDNA depletion. *TFAM* expression levels by day 7 and 14 of mtDNA depletion already exceeded levels of day 0 cells. With the duration of recovery, expression of *TFAM* progressively increased from day 3 of mtDNA replenishment onwards, which did not fall below levels of day 0 cells at any point during the recovery process for both day 7 and 14 mtDNA depleted cells. Together, these observations confirm that the accumulation of *TFAM* expression is important during the recovery of copy number in multiple myeloma cells.

### **7.6.2 GBM**

MtDNA copy number and *TFAM* expression levels in the HSR-GBM1 cells revealed different patterns of change. MtDNA copy numbers were significantly up-regulated relative to day 0 cells by day 14 of recovery after days 7 and 14 of mtDNA depletion. Both these depleted cells were unable to increase *TFAM* expression levels above day 0 cells at all stages of recovery, which contradicts the theory that excessive expression of *TFAM* must be present for mtDNA copy number to undergo efficient recovery. It is possible that the up-regulation of several other key mitochondrial transcription and replication genes is required for the recovery of mtDNA copy number. Alternatively, regulation at the post-transcriptional level that controls the stability and translation of mtDNA transcripts may have a role in explaining the differences observed between multiple myeloma and GBM cells<sup>379</sup>.

### 7.7. Bi-directional communication between mitochondria and the nucleus

One of the key observations from Chapter 4 was that, during the late stages of mtDNA depletion, there was up-regulation of tumor proliferation genes, such as *FGFR3* and *MAPK*. These results corresponded with the reduced ability of U266 cells to proliferate, which was also seen for cells recovered for 14 days in the absence of 2-DG after 14 days of mtDNA depletion. Based on these findings, it is possible that there exists a threshold level for mtDNA copy number. Below this threshold level, there may be insufficient mtDNA copies to provide feedback signals to the nucleus (**Figure 7.1**). The relationship between the mitochondria and nucleus is likely that the nucleus produces mitochondrial biogenesis, transcription and replication factors to positively replicate the mtDNA. However, in the event where mtDNA copy number is not replenished sufficiently to be above the critical threshold level, which could arise from an accumulation of mtDNA variants during recovery, synchrony between the mitochondria and nucleus would be disrupted. This is because the nucleus would not receive enough signals that would otherwise inform of the successful replication of mtDNA. The nucleus would then respond by increasing the expression of tumorigenic genes (**Figure 7.1**). Ultimately, this nuclear response would promote a positive feedback mechanism that enhances both the malignant characteristics of the tumor cells, as well as further stimulation of the Warburg effect.



**Figure 7.1.** Model of bi-directional communication between the nucleus and mitochondria. It was hypothesised that a feedback mechanism between the nucleus and the mitochondrion positively regulates cell proliferation. A threshold level for mtDNA copy number exists whereby, below the critical limit (identified at day 14 of mtDNA depletion), there would be disrupted synchrony in the communication between the nucleus and mitochondria. This would lead to the nucleus responding by increasing its expression for key tumorigenic genes, which would further drive both tumor development and the Warburg effect.

Support for the Warburg effect within U266 and HSR-GBM1 cells lies in the observation that these cells are able to survive even during prolonged mtDNA depletion (**Figure 4.4B**, **Figure 6.2**), as the Warburg effect produces energy from the cytoplasm. However, the Warburg effect does not occur exclusively within these cells, as respirometry results demonstrated the presence of functional OXPHOS in both cell types (**Table 4.2**).

Evidence exists that can be used to support the hypothesis that there may be bi-directional communication between the mitochondrion and the nucleus for maintaining cell proliferation. Some have suggested that communication between the two compartments is essential for controlling the expression of nuclear-encoded proteins for example, the subunits comprising the cytochrome c oxidase complex<sup>380</sup>. The interactions between the two genomes are therefore important for ensuring the correct synthesis and assembly of the protein complexes that comprise the ETC, which would then allow for energy production via OXPHOS as well as other biological processes to take place, inclusive of cellular proliferation.

The few studies that have investigated for molecules that interact with both the nucleus and the mitochondrion emerged following the discovery that peptides degraded from mitochondrial proteins are exported from the mitochondrion as a mechanism of quality control for mitochondrial proteins<sup>381,382</sup>. More than one peptide efflux pathway has been described in the mitochondrion<sup>383</sup>. In addition, several molecules have been proposed to provide interaction between the mitochondrion and the nucleus. Apart from MondoA-Mlx, mentioned in Chapter 4, others include nuclear-encoded SIR2, a histone deacetylase whose function is dependent on the presence of NAD<sup>384</sup>, and mitochondrial protein dihydroorotate dehydrogenase. Dihydroorotate dehydrogenase senses the abundance of free nucleotides to control synthesis of DNA and RNA, with the dual role of detecting the ratio of ubiquinone/ubiquinol and NADH levels since it catalyses the oxidation of dihydroorotate to orotate via ubiquinone, required for the *de novo* synthesis of pyrimidines in the mitochondrion<sup>385</sup>.

## **7.8. Further evidence to support the role of mtDNA in the tumor biology of GBM**

Recently, a study by Dickinson *et al.* have published evidence to support the hypothesis that mtDNA are crucial initiators of GBM, as discussed in Chapter 6. MtDNA copy number in GBM cell lines, HSR-GBM1, GBM L1 and GBM L2 were examined during differentiation into astrocytes<sup>386</sup>. Compared to hNSCs, the GBM cells were found not to effectively increase their mtDNA copy number after 28 days of differentiation into astrocytes. Alongside this, respirometry results demonstrated that oxygen consumption rates were higher in hNSC-derived astrocytes than astrocytes differentiated from HSR-GBM1 cells. Undifferentiated HSR-GBM1 cells also had higher lactate production, which declined with the duration of differentiation when compared with hNSCs<sup>386</sup>. On the other hand, ATP production during differentiation was greater in hNSCs than for HSR-GBM1 cells. These observations together demonstrated the importance of mtDNA copy number in the tumor biology of GBM (Figure 7.1).

Furthermore, mtDNA depletion and recovery of HSR-GBM1 cells in BALB/c nude mice, demonstrated an inverse correlation between the rate of tumor growth and the extent of mtDNA depletion<sup>386</sup>. Reduced tumor growth during recovery after mtDNA depletion could be related to the acquisition of other GBM variants during mtDNA replenishment, as described in Chapter 5. It is also likely that mtDNA copy number could be responsible for the changes to gene expression patterns identified in Chapter 6, since depletion of GBM cells using ddC did not disrupt mitochondrial integrity, as shown by stable expression of the OPA-1

mitochondrial protein<sup>386</sup>. These findings strongly support the concept that mtDNA acts as initiators of GBM tumor establishment.

### **7.9. Summary**

Overall, the conclusions from this study are:

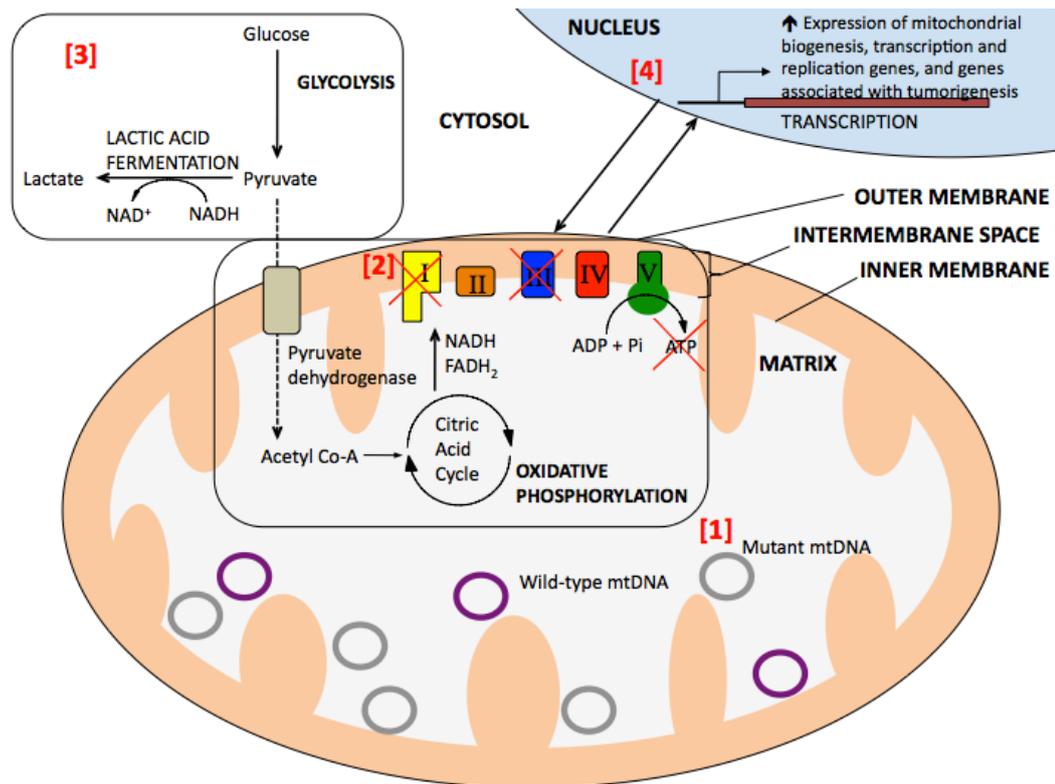
[1] The abnormal accumulation of mtDNA variants in cancer cells can result in the formation of dysfunctional protein complexes that comprise the ETC, especially if non-synonymous amino acid mutations translate as polypeptides that disrupt critical regions of the protein complex (**Figure 7.2; 1**). Conserved regions of the protein complex are more likely to carry greater deleterious effects.

[2] Impaired protein function of the ETC complexes leads to disruption in the sequential transfer of electrons, and therefore the inability to generate an electrochemical gradient that could otherwise be used to produce ATP molecules via ATP synthase (Complex V) (**Figure 7.2; 2**). This causes the cancer cell to respire predominately via glycolysis.

[3] Overtime, glycolytic cancer cells establish low mtDNA copy numbers, which correlate with tumor grade (**Figure 7.2; 3**). This is likely to occur, as mtDNA encodes genes essential for OXPHOS function, and due to the reduced demand for OXPHOS in predominately glycolytic cells, there would be a reduced need for efficient mtDNA replication. During the decline of mtDNA copy number, there may also be the acquisition of further mtDNA variants.

[4] The presence of low mtDNA copy number increases the expression levels of nuclear-encoded genes associated with mitochondrial biogenesis, transcription and replication as a compensatory effect (**Figure 7.2; 4**). The inability of mtDNA to expand its copy number in response to the increased expression of these genes, which could arise as a consequence of the accumulation of mtDNA variants, results in the increase in the expression of tumorigenic genes from the nucleus. This occurs due to disruption in the synchrony between the mitochondria and nuclear compartments. The up-regulation in the expression of tumor biology genes would positively enhance tumor cell malignancy and further promote the Warburg effect and thus cellular proliferation rather than differentiation.

In conclusion, mtDNA variants and mtDNA copy number act together in controlling the tumor biology of multiple myeloma and GBM cells. This arises as cells are forced to generate energy through aerobic glycolysis, which promotes proliferation of naïve cells rather than the differentiation of naïve cells into mature, differentiated cells.



**Figure 7.2.** Proposed model for the role of mtDNA in the tumor biology of multiple myeloma and GBM cells. The mitochondrion is positioned as the central organelle (orange) and the nucleus is shown in blue. The two biochemical pathways for energy production are presented within the boxes. [1] Mutant mtDNA copies in the mitochondrion arise as a consequence of the abnormal accumulation of mtDNA variants in cancer cells. [2] The presence of mutant mtDNA disrupts the function of proteins that comprise the ETC, which leads to defective OXPHOS in these cancer cells. [3] Since glycolysis predominates as the primary method for ATP generation, there is a reduced need for these cancer cells to efficiently replicate their mtDNA, resulting in a decline in mtDNA copy number and possibly the acquisition of further mtDNA variants. [4] The presence of low mtDNA copy number increases the expression levels of nuclear-encoded mitochondrial genes that assist in the biogenesis, transcription and replication of mtDNA. However, the failure of mtDNA to respond to the increased expression of these genes causes disruption to the synchrony between the nucleus and mitochondria. This promotes tumor malignancy by the up-regulation of key tumorigenic genes, which enhances the Warburg effect and increases cell proliferation.

### **7.10. Further work**

To further confirm that mtDNA is essential for tumor establishment, it would be beneficial to perform functional studies using immunodeficient mice models. MtDNA depleted HSR-GBM1 cells have already been examined for mtDNA variants during tumor formation in BALB/c nude mice. However, it is important to do this work with multiple myeloma cells, as no *in vivo* experiments have yet been done on these cells. This is especially important, as there is currently no effective method to mimic the bone marrow microenvironment using *in vitro* techniques. Furthermore, the microenvironment of the bone marrow, where multiple myeloma cells are naturally found, consists of various interacting cells that are capable of secreting factors to promote the survival and proliferation of multiple myeloma cells.

# References

1. Petruzzella V, Papa S. Mutations in human nuclear genes encoding for subunits of mitochondrial respiratory complex I: the NDUFS4 gene. *Gene*. 2002;286(1):149-154.
2. Garesse R, Vallejo CG. Animal mitochondrial biogenesis and function: a regulatory cross-talk between two genomes. *Gene*. 2001;263(1-2):1-16.
3. Rehling P, Brandner K, Pfanner N. Mitochondrial import and the twin-pore translocase. *Nat Rev Mol Cell Biol*. 2004;5(7):519-530.
4. Fujiki M, Verner K. Coupling of cytosolic protein synthesis and mitochondrial protein import in yeast. Evidence for cotranslational import in vivo. *J Biol Chem*. 1993;268(3):1914-1920.
5. Bhagwat SV, Biswas G, Anandatheerthavarada HK, Addya S, Pandak W, Avadhani NG. Dual targeting property of the N-terminal signal sequence of P4501A1. Targeting of heterologous proteins to endoplasmic reticulum and mitochondria. *J Biol Chem*. 1999;274(34):24014-24022.
6. Wrobel L, Trojanowska A, Sztolszterer ME, Chacinska A. Mitochondrial protein import: Mia40 facilitates Tim22 translocation into the inner membrane of mitochondria. *Mol Biol Cell*. 2013;24(5):543-554.
7. Mayer A, Lill R, Neupert W. Translocation and insertion of precursor proteins into isolated outer membranes of mitochondria. *J Cell Biol*. 1993;121(6):1233-1243.
8. Diekert K, Kispal G, Guiard B, Lill R. An internal targeting signal directing proteins into the mitochondrial intermembrane space. *Proc Natl Acad Sci U S A*. 1999;96(21):11752-11757.
9. Muro C, Grigoriev SM, Pietkiewicz D, Kinnally KW, Campo ML. Comparison of the TIM and TOM channel activities of the mitochondrial protein import complexes. *Biophys J*. 2003;84(5):2981-2989.
10. Eilers M, Oppliger W, Schatz G. Both ATP and an energized inner membrane are required to import a purified precursor protein into mitochondria. *EMBO J*. 1987;6(4):1073-1077.
11. Wu Z, Puigserver P, Andersson U, et al. Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1. *Cell*. 1999;98(1):115-124.
12. Kong X, Wang R, Xue Y, et al. Sirtuin 3, a new target of PGC-1alpha, plays an important role in the suppression of ROS and mitochondrial biogenesis. *PLoS One*. 2010;5(7):e11707.
13. Hondares E, Pineda-Torra I, Iglesias R, Staels B, Villarroya F, Giralt M. PPARdelta, but not PPARalpha, activates PGC-1alpha gene transcription in muscle. *Biochem Biophys Res Commun*. 2007;354(4):1021-1027.
14. Gleyzer N, Vercauteren K, Scarpulla RC. Control of mitochondrial transcription specificity factors (TFB1M and TFB2M) by nuclear respiratory factors (NRF-1 and NRF-2) and PGC-1 family coactivators. *Mol Cell Biol*. 2005;25(4):1354-1366.
15. Ekstrand MI, Falkenberg M, Rantanen A, et al. Mitochondrial transcription factor A regulates mtDNA copy number in mammals. *Hum Mol Genet*. 2004;13(9):935-944.
16. Cline SD. Mitochondrial DNA damage and its consequences for mitochondrial gene expression. *Biochim Biophys Acta*. 2012;1819(9-10):979-991.
17. Copeland WC. The mitochondrial DNA polymerase in health and disease. *Subcell Biochem*. 2010;50:211-222.

18. Tynismaa H, Sembongi H, Bokori-Brown M, et al. Twinkle helicase is essential for mtDNA maintenance and regulates mtDNA copy number. *Hum Mol Genet.* 2004;13(24):3219-3227.
19. Asin-Cayuela J, Schwend T, Farge G, Gustafsson CM. The human mitochondrial transcription termination factor (mTERF) is fully active in vitro in the non-phosphorylated form. *J Biol Chem.* 2005;280(27):25499-25505.
20. Fisher RP, Lisowsky T, Parisi MA, Clayton DA. DNA wrapping and bending by a mitochondrial high mobility group-like transcriptional activator protein. *J Biol Chem.* 1992;267(5):3358-3367.
21. Robberson DL, Kasamatsu H, Vinograd J. Replication of mitochondrial DNA. Circular replicative intermediates in mouse L cells. *Proc Natl Acad Sci U S A.* 1972;69(3):737-741.
22. Moraes CT. What regulates mitochondrial DNA copy number in animal cells? *Trends Genet.* 2001;17(4):199-205.
23. Rowntree RK, Lee JT. Mapping of DNA replication origins to noncoding genes of the X-inactivation center. *Mol Cell Biol.* 2006;26(10):3707-3717.
24. Pohjoismaki JL, Holmes JB, Wood SR, et al. Mammalian mitochondrial DNA replication intermediates are essentially duplex but contain extensive tracts of RNA/DNA hybrid. *J Mol Biol.* 2010;397(5):1144-1155.
25. Holt IJ, Lorimer HE, Jacobs HT. Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA. *Cell.* 2000;100(5):515-524.
26. Elenbaas B, Spirio L, Koerner F, et al. Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. *Genes Dev.* 2001;15(1):50-65.
27. Hao X, Du M, Bishop AE, Talbot IC. Imbalance between proliferation and apoptosis in the development of colorectal carcinoma. *Virchows Arch.* 1998;433(6):523-527.
28. Mattern J, Volm M. Imbalance of cell proliferation and apoptosis during progression of lung carcinomas. *Anticancer Res.* 2004;24(6):4243-4246.
29. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell.* 2000;100(1):57-70.
30. Bartek J, Lukas J. Mammalian G1- and S-phase checkpoints in response to DNA damage. *Curr Opin Cell Biol.* 2001;13(6):738-747.
31. DiPaola RS. To Arrest or Not To G2-M Cell-Cycle Arrest : Commentary re: A. K. Tyagi et al., Silibinin Strongly Synergizes Human Prostate Carcinoma DU145 Cells to Doxorubicin-induced Growth Inhibition, G2-M Arrest, and Apoptosis. *Clin. Cancer Res.*, 8: 3512-3519, 2002. *Clinical Cancer Research.* 2002;8(11):3311-3314.
32. Gorbsky GJ. Cell cycle checkpoints: arresting progress in mitosis. *Bioessays.* 1997;19(3):193-197.
33. Musacchio A, Salmon ED. The spindle-assembly checkpoint in space and time. *Nat Rev Mol Cell Biol.* 2007;8(5):379-393.
34. Gandarillas A. The mysterious human epidermal cell cycle, or an oncogene-induced differentiation checkpoint. *Cell Cycle.* 2012;11(24):4507-4516.
35. Yang G, Rosen DG, Liu G, et al. CXCR2 promotes ovarian cancer growth through dysregulated cell cycle, diminished apoptosis, and enhanced angiogenesis. *Clin Cancer Res.* 2010;16(15):3875-3886.

36. Konigsberg R, Rogelsperger O, Jager W, et al. Cell cycle dysregulation influences survival in high risk breast cancer patients. *Cancer Invest.* 2008;26(7):734-740.
37. Rivlin N, Brosh R, Oren M, Rotter V. Mutations in the p53 Tumor Suppressor Gene: Important Milestones at the Various Steps of Tumorigenesis. *Genes Cancer.* 2011;2(4):466-474.
38. Matozaki T, Sakamoto C, Suzuki T, et al. p53 gene mutations in human gastric cancer: wild-type p53 but not mutant p53 suppresses growth of human gastric cancer cells. *Cancer Res.* 1992;52(16):4335-4341.
39. Lane DP. p53 and human cancers. *Br Med Bull.* 1994;50(3):582-599.
40. Cormio A, Guerra F, Cormio G, et al. The PGC-1alpha-dependent pathway of mitochondrial biogenesis is upregulated in type I endometrial cancer. *Biochem Biophys Res Commun.* 2009;390(4):1182-1185.
41. Chen IC, Chiang WF, Liu SY, Chen PF, Chiang HC. Role of SIRT3 in the regulation of redox balance during oral carcinogenesis. *Mol Cancer.* 2013;12:68.
42. Mandal S, Guptan P, Owusu-Ansah E, Banerjee U. Mitochondrial regulation of cell cycle progression during development as revealed by the tenured mutation in *Drosophila*. *Dev Cell.* 2005;9(6):843-854.
43. Owusu-Ansah E, Yavari A, Mandal S, Banerjee U. Distinct mitochondrial retrograde signals control the G1-S cell cycle checkpoint. *Nat Genet.* 2008;40(3):356-361.
44. Dohi T, Beltrami E, Wall NR, Plescia J, Altieri DC. Mitochondrial survivin inhibits apoptosis and promotes tumorigenesis. *J Clin Invest.* 2004;114(8):1117-1127.
45. Singh KK, Ayyasamy V, Owens KM, Koul MS, Vujcic M. Mutations in mitochondrial DNA polymerase-gamma promote breast tumorigenesis. *J Hum Genet.* 2009;54(9):516-524.
46. Li F, Wang Y, Zeller KI, et al. Myc stimulates nuclearly encoded mitochondrial genes and mitochondrial biogenesis. *Mol Cell Biol.* 2005;25(14):6225-6234.
47. Tatuch Y, Christodoulou J, Feigenbaum A, et al. Heteroplasmic mtDNA mutation (T----G) at 8993 can cause Leigh disease when the percentage of abnormal mtDNA is high. *Am J Hum Genet.* 1992;50(4):852-858.
48. Folmes CD, Martinez-Fernandez A, Perales-Clemente E, et al. Disease-causing mitochondrial heteroplasmy segregated within induced pluripotent stem cell clones derived from a patient with MELAS. *Stem Cells.* 2013;31(7):1298-1308.
49. Jones JB, Song JJ, Hempen PM, Parmigiani G, Hruban RH, Kern SE. Detection of mitochondrial DNA mutations in pancreatic cancer offers a "mass"-ive advantage over detection of nuclear DNA mutations. *Cancer Res.* 2001;61(4):1299-1304.
50. Coller HA, Khrapko K, Bodyak ND, Nekhaeva E, Herrero-Jimenez P, Thilly WG. High frequency of homoplasmic mitochondrial DNA mutations in human tumors can be explained without selection. *Nat Genet.* 2001;28(2):147-150.
51. He Y, Wu J, Dressman DC, et al. Heteroplasmic mitochondrial DNA mutations in normal and tumour cells. *Nature.* 2010;464(7288):610-614.

52. Jazin EE, Cavelier L, Eriksson I, Orelund L, Gyllensten U. Human brain contains high levels of heteroplasmy in the noncoding regions of mitochondrial DNA. *Proc Natl Acad Sci U S A*. 1996;93(22):12382-12387.
53. Hu X, Cong Y, Xu C, Feng J, Jiang Y, Jin H. Variants of the mitochondrial displacement loop in patients with myelodysplastic syndromes. *Chinese Journal of Clinical Oncology*. 2008;5(5):343-348.
54. Vanecek T, Vorel F, Sip M. Mitochondrial DNA D-loop hypervariable regions: Czech population data. *Int J Legal Med*. 2004;118(1):14-18.
55. Lehocky I, Baldovic M, Kadasi L, Metspalu E. A database of mitochondrial DNA hypervariable regions I and II sequences of individuals from Slovakia. *Forensic Sci Int Genet*. 2008;2(4):e53-59.
56. Nur Haslindawaty AR, Panneerchelvam S, Edinur HA, Norazmi MN, Zafarina Z. Sequence polymorphisms of mtDNA HV1, HV2, and HV3 regions in the Malay population of Peninsular Malaysia. *Int J Legal Med*. 2010;124(5):415-426.
57. Howell N. Origin, cellular expression, and cybrid transmission of mitochondrial CAP-R, PYR-IND, and OLI-R mutant phenotypes. *Somatic Cell Genet*. 1983;9(1):1-24.
58. Petros JA, Baumann AK, Ruiz-Pesini E, et al. mtDNA mutations increase tumorigenicity in prostate cancer. *Proc Natl Acad Sci U S A*. 2005;102(3):719-724.
59. Hayashi JI, Yonekawa H, Tagashira Y. Nuclear but not mitochondrial genome involvement in 3-methylcholanthrene-induced expression of tumorigenicity in mouse somatic cells. *Cancer Res*. 1989;49(17):4715-4720.
60. Berg JM TJ, Stryer L. Chapter 16: Glycolysis and Gluconeogenesis. *Biochemistry*. New York: W H Freeman; 2002.
61. Mahalingam B, Cuesta-Munoz A, Davis EA, Matschinsky FM, Harrison RW, Weber IT. Structural model of human glucokinase in complex with glucose and ATP: implications for the mutants that cause hypo- and hyperglycemia. *Diabetes*. 1999;48(9):1698-1705.
62. Rozovsky S, Jogl G, Tong L, McDermott AE. Solution-state NMR investigations of triosephosphate isomerase active site loop motion: ligand release in relation to active site loop dynamics. *J Mol Biol*. 2001;310(1):271-280.
63. Turner N, Pearson A, Sharpe R, et al. FGFR1 amplification drives endocrine therapy resistance and is a therapeutic target in breast cancer. *Cancer Res*. 2010;70(5):2085-2094.
64. Fan J, Hitosugi T, Chung TW, et al. Tyrosine phosphorylation of lactate dehydrogenase A is important for NADH/NAD(+) redox homeostasis in cancer cells. *Mol Cell Biol*. 2011;31(24):4938-4950.
65. Littlefield JW, Sanadi DR. Role of coenzyme A and diphosphopyridine nucleotide in the oxidation of pyruvate. *J Biol Chem*. 1952;199(1):65-70.
66. Gottschalk G. The stereospecificity of the citrate synthase in sulfate-reducing and photosynthetic bacteria. *Eur J Biochem*. 1968;5(3):346-351.
67. Chinopoulos C. Which way does the citric acid cycle turn during hypoxia? The critical role of alpha-ketoglutarate dehydrogenase complex. *J Neurosci Res*. 2013;91(8):1030-1043.
68. Halabe Bucay A. The biological significance of cancer: mitochondria as a cause of cancer and the inhibition of glycolysis with citrate as a cancer treatment. *Med Hypotheses*. 2007;69(4):826-828.

69. Tornheim K, Lowenstein JM. Control of phosphofructokinase from rat skeletal muscle. Effects of fructose diphosphate, AMP, ATP, and citrate. *J Biol Chem.* 1976;251(23):7322-7328.
70. Velichko MG, Trebukhina RV, Ostrovskii Iu M. [Features of pyruvate and lactate metabolism in tumor-bearing rats following citrate administration]. *Vopr Med Khim.* 1981;27(1):68-72.
71. Pelicano H, Martin DS, Xu RH, Huang P. Glycolysis inhibition for anticancer treatment. *Oncogene.* 2006;25(34):4633-4646.
72. Kaplon J, Zheng L, Meissl K, et al. A key role for mitochondrial gatekeeper pyruvate dehydrogenase in oncogene-induced senescence. *Nature.* 2013;498(7452):109-112.
73. Tsui KH, Chung LC, Wang SW, Feng TH, Chang PL, Juang HH. Hypoxia upregulates the gene expression of mitochondrial aconitase in prostate carcinoma cells. *J Mol Endocrinol.* 2013;51(1):131-141.
74. Smeitink JA, Loeffen JL, Triepels RH, Smeets RJ, Trijbels JM, van den Heuvel LP. Nuclear genes of human complex I of the mitochondrial electron transport chain: state of the art. *Hum Mol Genet.* 1998;7(10):1573-1579.
75. Korla K, Mitra CK. Modelling the Krebs cycle and oxidative phosphorylation. *J Biomol Struct Dyn.* 2013.
76. Mourier A, Larsson NG. Tracing the trail of protons through complex I of the mitochondrial respiratory chain. *PLoS Biol.* 2011;9(8):e1001129.
77. Hong S, Pedersen PL. ATP synthase and the actions of inhibitors utilized to study its roles in human health, disease, and other scientific areas. *Microbiol Mol Biol Rev.* 2008;72(4):590-641, Table of Contents.
78. Wycherley G, Kane MT, Hynes AC. Oxidative phosphorylation and the tricarboxylic acid cycle are essential for normal development of mouse ovarian follicles. *Hum Reprod.* 2005;20(10):2757-2763.
79. Loeffen J, Smeitink J, Triepels R, et al. The first nuclear-encoded complex I mutation in a patient with Leigh syndrome. *Am J Hum Genet.* 1998;63(6):1598-1608.
80. Keeney PM, Xie J, Capaldi RA, Bennett JP, Jr. Parkinson's disease brain mitochondrial complex I has oxidatively damaged subunits and is functionally impaired and misassembled. *J Neurosci.* 2006;26(19):5256-5264.
81. Taylor RW, Birch-Machin MA, Schaefer J, et al. Deficiency of complex II of the mitochondrial respiratory chain in late-onset optic atrophy and ataxia. *Ann Neurol.* 1996;39(2):224-232.
82. Mancuso M, Filosto M, Stevens JC, et al. Mitochondrial myopathy and complex III deficiency in a patient with a new stop-codon mutation (G339X) in the cytochrome b gene. *J Neurol Sci.* 2003;209(1-2):61-63.
83. Bruno C, Santorelli FM, Assereto S, et al. Progressive exercise intolerance associated with a new muscle-restricted nonsense mutation (G142X) in the mitochondrial cytochrome b gene. *Muscle Nerve.* 2003;28(4):508-511.
84. Rahman S, Brown RM, Chong WK, Wilson CJ, Brown GK. A SURF1 gene mutation presenting as isolated leukodystrophy. *Ann Neurol.* 2001;49(6):797-800.
85. Budde SM, van den Heuvel LP, Janssen AJ, et al. Combined enzymatic complex I and III deficiency associated with mutations in the nuclear encoded NDUF54 gene. *Biochem Biophys Res Commun.* 2000;275(1):63-68.

86. Acin-Perez R, Bayona-Bafaluy MP, Fernandez-Silva P, et al. Respiratory complex III is required to maintain complex I in mammalian mitochondria. *Mol Cell*. 2004;13(6):805-815.
87. Lanciano P, Lee DW, Yang H, Darrouzet E, Daldal F. Intermonomer electron transfer between the low-potential b hemes of cytochrome bc(1). *Biochemistry*. 2011;50(10):1651-1663.
88. Cooley JW, Lee DW, Daldal F. Across membrane communication between the Q(o) and Q(i) active sites of cytochrome bc(1). *Biochemistry*. 2009;48(9):1888-1899.
89. Smith PM, Fox JL, Winge DR. Biogenesis of the cytochrome bc(1) complex and role of assembly factors. *Biochim Biophys Acta*. 2012;1817(2):276-286.
90. Darrouzet E, Valkova-Valchanova M, Moser CC, Dutton PL, Daldal F. Uncovering the [2Fe2S] domain movement in cytochrome bc1 and its implications for energy conversion. *Proc Natl Acad Sci U S A*. 2000;97(9):4567-4572.
91. Darrouzet E, Daldal F. Protein-protein interactions between cytochrome b and the Fe-S protein subunits during QH2 oxidation and large-scale domain movement in the bc1 complex. *Biochemistry*. 2003;42(6):1499-1507.
92. Cooley JW, Ohnishi T, Daldal F. Binding dynamics at the quinone reduction (Qi) site influence the equilibrium interactions of the iron sulfur protein and hydroquinone oxidation (Qo) site of the cytochrome bc1 complex. *Biochemistry*. 2005;44(31):10520-10532.
93. Yin Y, Yang S, Yu L, Yu CA. Reaction mechanism of superoxide generation during ubiquinol oxidation by the cytochrome bc1 complex. *J Biol Chem*. 2010;285(22):17038-17045.
94. Stiburek L, Zeman J. Assembly factors and ATP-dependent proteases in cytochrome c oxidase biogenesis. *Biochim Biophys Acta*. 2010;1797(6-7):1149-1158.
95. Hornig-Do HT, Tatsuta T, Buckermann A, et al. Nonsense mutations in the COX1 subunit impair the stability of respiratory chain complexes rather than their assembly. *EMBO J*. 2012;31(5):1293-1307.
96. Fontanesi F, Barrientos, A. Chapter 15: Mitochondrial Cytochrome c Oxidase Assembly in Health and Human Diseases. In: Wong L-JC, ed. *Mitochondrial Disorders Caused by Nuclear Genes*: Springer New York; 2013:p.239-259.
97. Herrmann JM, Funes S. Biogenesis of cytochrome oxidase-sophisticated assembly lines in the mitochondrial inner membrane. *Gene*. 2005;354:43-52.
98. Stiburek L, Hansikova H, Tesarova M, Cerna L, Zeman J. Biogenesis of eukaryotic cytochrome c oxidase. *Physiol Res*. 2006;55 Suppl 2:S27-41.
99. Nijtmans LG, Taanman JW, Muijsers AO, Speijer D, Van den Bogert C. Assembly of cytochrome-c oxidase in cultured human cells. *Eur J Biochem*. 1998;254(2):389-394.
100. Tsukihara T, Aoyama H, Yamashita E, et al. The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 Å. *Science*. 1996;272(5265):1136-1144.
101. Wong LJ. Pathogenic mitochondrial DNA mutations in protein-coding genes. *Muscle Nerve*. 2007;36(3):279-293.

102. da Fonseca RR, Johnson WE, O'Brien SJ, Ramos MJ, Antunes A. The adaptive evolution of the mammalian mitochondrial genome. *BMC Genomics*. 2008;9:119.
103. Warburg O. On the origin of cancer cells. *Science*. 1956;123(3191):309-314.
104. Lunt SY, Vander Heiden MG. Aerobic glycolysis: meeting the metabolic requirements of cell proliferation. *Annu Rev Cell Dev Biol*. 2011;27:441-464.
105. Vazquez A, Liu J, Zhou Y, Oltvai ZN. Catabolic efficiency of aerobic glycolysis: the Warburg effect revisited. *BMC Syst Biol*. 2010;4:58.
106. Pedersen PL. Warburg, me and Hexokinase 2: Multiple discoveries of key molecular events underlying one of cancers' most common phenotypes, the "Warburg Effect", i.e., elevated glycolysis in the presence of oxygen. *J Bioenerg Biomembr*. 2007;39(3):211-222.
107. Gottfried E, Kunz-Schughart LA, Ebner S, et al. Tumor-derived lactic acid modulates dendritic cell activation and antigen expression. *Blood*. 2006;107(5):2013-2021.
108. Fischer K, Hoffmann P, Voelkl S, et al. Inhibitory effect of tumor cell-derived lactic acid on human T cells. *Blood*. 2007;109(9):3812-3819.
109. Moreno-Sanchez R, Rodriguez-Enriquez S, Marin-Hernandez A, Saavedra E. Energy metabolism in tumor cells. *FEBS J*. 2007;274(6):1393-1418.
110. Suganuma K, Miwa H, Imai N, et al. Energy metabolism of leukemia cells: glycolysis versus oxidative phosphorylation. *Leuk Lymphoma*. 2010;51(11):2112-2119.
111. Semenza GL, Roth PH, Fang HM, Wang GL. Transcriptional regulation of genes encoding glycolytic enzymes by hypoxia-inducible factor 1. *J Biol Chem*. 1994;269(38):23757-23763.
112. Shim H, Dolde C, Lewis BC, et al. c-Myc transactivation of LDH-A: implications for tumor metabolism and growth. *Proc Natl Acad Sci U S A*. 1997;94(13):6658-6663.
113. Gogvadze V, Orrenius S, Zhivotovsky B. Mitochondria in cancer cells: what is so special about them? *Trends Cell Biol*. 2008;18(4):165-173.
114. Fogal V, Richardson AD, Karmali PP, Scheffler IE, Smith JW, Ruoslahti E. Mitochondrial p32 protein is a critical regulator of tumor metabolism via maintenance of oxidative phosphorylation. *Mol Cell Biol*. 2010;30(6):1303-1318.
115. Xu W, Barrientos T, Andrews NC. Iron and copper in mitochondrial diseases. *Cell Metab*. 2013;17(3):319-328.
116. Thelander L, Graslund A, Thelander M. Continual presence of oxygen and iron required for mammalian ribonucleotide reduction: possible regulation mechanism. *Biochem Biophys Res Commun*. 1983;110(3):859-865.
117. Thelander M, Graslund A, Thelander L. Subunit M2 of mammalian ribonucleotide reductase. Characterization of a homogeneous protein isolated from M2-overproducing mouse cells. *J Biol Chem*. 1985;260(5):2737-2741.
118. Horowitz MP, Greenamyre JT. Mitochondrial iron metabolism and its role in neurodegeneration. *J Alzheimers Dis*. 2010;20 Suppl 2:S551-568.
119. Marchi S, Lupini L, Patergnani S, et al. Downregulation of the mitochondrial calcium uniporter by cancer-related miR-25. *Curr Biol*. 2013;23(1):58-63.

120. Shokolenko I, Venediktova N, Bochkareva A, Wilson GL, Alexeyev MF. Oxidative stress induces degradation of mitochondrial DNA. *Nucleic Acids Res.* 2009;37(8):2539-2548.
121. Jackson AL, Loeb LA. The contribution of endogenous sources of DNA damage to the multiple mutations in cancer. *Mutat Res.* 2001;477(1-2):7-21.
122. Burdon RH. Superoxide and hydrogen peroxide in relation to mammalian cell proliferation. *Free Radic Biol Med.* 1995;18(4):775-794.
123. Covarrubias D, Bai RK, Wong LJ, Leal SM. Mitochondrial DNA variant interactions modify breast cancer risk. *J Hum Genet.* 2008;53(10):924-928.
124. Bai RK, Leal SM, Covarrubias D, Liu A, Wong LJ. Mitochondrial genetic background modifies breast cancer risk. *Cancer Res.* 2007;67(10):4687-4694.
125. Booker LM, Habermacher GM, Jessie BC, et al. North American white mitochondrial haplogroups in prostate and renal cancer. *J Urol.* 2006;175(2):468-472; discussion 472-463.
126. Ely B, Wilson JL, Jackson F, Jackson BA. African-American mitochondrial DNAs often match mtDNAs found in multiple African ethnic groups. *BMC Biol.* 2006;4:34.
127. Li XY, Guo YB, Su M, Cheng L, Lu ZH, Tian DP. Association of mitochondrial haplogroup D and risk of esophageal cancer in Taihang Mountain and Chaoshan areas in China. *Mitochondrion.* 2011;11(1):27-32.
128. Aikhionbare FO, Khan M, Carey D, Okoli J, Go R. Is cumulative frequency of mitochondrial DNA variants a biomarker for colorectal tumor progression? *Mol Cancer.* 2004;3:30.
129. Zhu W, Qin W, Bradley P, Wessel A, Puckett CL, Sauter ER. Mitochondrial DNA mutations in breast cancer tissue and in matched nipple aspirate fluid. *Carcinogenesis.* 2005;26(1):145-152.
130. Gasparre G, Porcelli AM, Bonora E, et al. Disruptive mitochondrial DNA mutations in complex I subunits are markers of oncocytic phenotype in thyroid tumors. *Proc Natl Acad Sci U S A.* 2007;104(21):9001-9006.
131. Zhai K, Chang L, Zhang Q, Liu B, Wu Y. Mitochondrial C150T polymorphism increases the risk of cervical cancer and HPV infection. *Mitochondrion.* 2011;11(4):559-563.
132. Santoro A, Salvioli S, Raule N, et al. Mitochondrial DNA involvement in human longevity. *Biochim Biophys Acta.* 2006;1757(9-10):1388-1399.
133. Polyak K, Li Y, Zhu H, et al. Somatic mutations of the mitochondrial genome in human colorectal tumours. *Nat Genet.* 1998;20(3):291-293.
134. Allegra E, Garozzo A, Lombardo N, De Clemente M, Carey TE. Mutations and polymorphisms in mitochondrial DNA in head and neck cancer cell lines. *Acta Otorhinolaryngol Ital.* 2006;26(4):185-190.
135. Fliss MS, Usadel H, Caballero OL, et al. Facile detection of mitochondrial DNA mutations in tumors and bodily fluids. *Science.* 2000;287(5460):2017-2019.
136. Dai JG, Xiao YB, Min JX, Zhang GQ, Yao K, Zhou RJ. Mitochondrial DNA 4977 BP deletion mutations in lung carcinoma. *Indian J Cancer.* 2006;43(1):20-25.
137. Maximo V, Soares P, Lima J, Cameselle-Teijeiro J, Sobrinho-Simoes M. Mitochondrial DNA somatic mutations (point mutations and large deletions) and mitochondrial DNA variants in human thyroid pathology: a study with emphasis on Hurthle cell tumors. *Am J Pathol.* 2002;160(5):1857-1865.

138. Ye C, Shu XO, Wen W, et al. Quantitative analysis of mitochondrial DNA 4977-bp deletion in sporadic breast cancer and benign breast diseases. *Breast Cancer Res Treat.* 2008;108(3):427-434.
139. Shieh DB, Chou WP, Wei YH, Wong TY, Jin YT. Mitochondrial DNA 4,977-bp deletion in paired oral cancer and precancerous lesions revealed by laser microdissection and real-time quantitative PCR. *Ann N Y Acad Sci.* 2004;1011:154-167.
140. Wu CW, Yin PH, Hung WY, et al. Mitochondrial DNA mutations and mitochondrial DNA depletion in gastric cancer. *Genes Chromosomes Cancer.* 2005;44(1):19-28.
141. Mithani SK, Taube JM, Zhou S, et al. Mitochondrial mutations are a late event in the progression of head and neck squamous cell cancer. *Clin Cancer Res.* 2007;13(15 Pt 1):4331-4335.
142. Zhou S, Kachhap S, Sun W, et al. Frequency and phenotypic implications of mitochondrial DNA mutations in human squamous cell cancers of the head and neck. *Proc Natl Acad Sci U S A.* 2007;104(18):7540-7545.
143. Lee HC, Hsu LS, Yin PH, Lee LM, Chi CW. Heteroplasmic mutation of mitochondrial DNA D-loop and 4977-bp deletion in human cancer cells during mitochondrial DNA depletion. *Mitochondrion.* 2007;7(1-2):157-163.
144. Kyle RA, Therneau TM, Rajkumar SV, et al. Prevalence of monoclonal gammopathy of undetermined significance. *N Engl J Med.* 2006;354(13):1362-1369.
145. Kyle RA, Therneau TM, Rajkumar SV, et al. A long-term study of prognosis in monoclonal gammopathy of undetermined significance. *N Engl J Med.* 2002;346(8):564-569.
146. Kariyawan CC, Hughes DA, Jayatilake MM, Mehta AB. Multiple myeloma: causes and consequences of delay in diagnosis. *QJM.* 2007;100(10):635-640.
147. Alsina M, Boyce B, Devlin RD, et al. Development of an in vivo model of human multiple myeloma bone disease. *Blood.* 1996;87(4):1495-1501.
148. Morgan GJ, Walker BA, Davies FE. The genetic architecture of multiple myeloma. *Nat Rev Cancer.* 2012;12(5):335-348.
149. Attalmanan M, Levinson SS. Understanding and identifying monoclonal gammopathies. *Clin Chem.* 2000;46(8 Pt 2):1230-1238.
150. Hose D, Reme T, Hielscher T, et al. Proliferation is a central independent prognostic factor and target for personalized and risk-adapted treatment in multiple myeloma. *Haematologica.* 2011;96(1):87-95.
151. Jia H, Liu C, Ge F, et al. Identification of ubiquitinated proteins from human multiple myeloma U266 cells by proteomics. *Biomed Environ Sci.* 2011;24(4):422-430.
152. Egan JB, Shi CX, Tembe W, et al. Whole-genome sequencing of multiple myeloma from diagnosis to plasma cell leukemia reveals genomic initiating events, evolution, and clonal tides. *Blood.* 2012;120(5):1060-1066.
153. Dalton WS. Targeting the mitochondria: an exciting new approach to myeloma therapy. Commentary re: N. J. Bahlis et al., Feasibility and correlates of arsenic trioxide combined with ascorbic acid-mediated depletion of intracellular glutathione for the treatment of relapsed/refractory multiple myeloma. *Clin. Cancer Res.*, 8: 3658-3668, 2002. *Clin Cancer Res.* 2002;8(12):3643-3645.

154. Mari M, Morales A, Colell A, Garcia-Ruiz C, Fernandez-Checa JC. Mitochondrial glutathione, a key survival antioxidant. *Antioxid Redox Signal*. 2009;11(11):2685-2700.
155. Dvorakova K, Waltmire CN, Payne CM, Tome ME, Briehl MM, Dorr RT. Induction of mitochondrial changes in myeloma cells by imexon. *Blood*. 2001;97(11):3544-3551.
156. Schwartzbaum JA, Fisher JL, Aldape KD, Wrensch M. Epidemiology and molecular pathology of glioma. *Nat Clin Pract Neurol*. 2006;2(9):494-503; quiz 491 p following 516.
157. Furnari FB, Fenton T, Bachoo RM, et al. Malignant astrocytic glioma: genetics, biology, and paths to treatment. *Genes Dev*. 2007;21(21):2683-2710.
158. Krex D, Klink B, Hartmann C, et al. Long-term survival with glioblastoma multiforme. *Brain*. 2007;130(Pt 10):2596-2606.
159. Griguer CE, Oliva CR. Bioenergetics pathways and therapeutic resistance in gliomas: emerging role of mitochondria. *Curr Pharm Des*. 2011;17(23):2421-2427.
160. Sizoo EM, Braam L, Postma TJ, et al. Symptoms and problems in the end-of-life phase of high-grade glioma patients. *Neuro Oncol*. 2010;12(11):1162-1166.
161. Miller CR, Dunham CP, Scheithauer BW, Perry A. Significance of necrosis in grading of oligodendroglial neoplasms: a clinicopathologic and genetic study of newly diagnosed high-grade gliomas. *J Clin Oncol*. 2006;24(34):5419-5426.
162. Prestegarden L, Enger P. Cancer stem cells in the central nervous system-a critical review. *Cancer Res*. 2010;70(21):8255-8258.
163. Pallini R, Ricci-Vitiani L, Montano N, et al. Expression of the stem cell marker CD133 in recurrent glioblastoma and its value for prognosis. *Cancer*. 2011;117(1):162-174.
164. Gao X, Wang H, Yang JJ, Liu X, Liu ZR. Pyruvate kinase M2 regulates gene transcription by acting as a protein kinase. *Mol Cell*. 2012;45(5):598-609.
165. Mukherjee J, Phillips JJ, Zheng S, Wiencke J, Ronen SM, Pieper RO. Pyruvate kinase M2 expression, but not pyruvate kinase activity, is up-regulated in a grade-specific manner in human glioma. *PLoS One*. 2013;8(2):e57610.
166. Janiszewska M, Suva ML, Riggi N, et al. Imp2 controls oxidative phosphorylation and is crucial for preserving glioblastoma cancer stem cells. *Genes Dev*. 2012;26(17):1926-1944.
167. Konishi T, Shimada Y, Nagao T, Okabe H, Konoshima T. Antiproliferative sesquiterpene lactones from the roots of *Inula helenium*. *Biol Pharm Bull*. 2002;25(10):1370-1372.
168. Khan M, Yi F, Rasul A, et al. Alantolactone induces apoptosis in glioblastoma cells via GSH depletion, ROS generation, and mitochondrial dysfunction. *IUBMB Life*. 2012;64(9):783-794.
169. Nilsson K, Bennich H, Johansson SG, Ponten J. Established immunoglobulin producing myeloma (IgE) and lymphoblastoid (IgG) cell lines from an IgE myeloma patient. *Clin Exp Immunol*. 1970;7(4):477-489.
170. Sambrook J, Fritsch, E.F., Maniatis, T. *Molecular Cloning: A laboratory manual*, Volume 1. Cold Spring Harbor Laboratory. (1989).
171. Facucho-Oliveira JM, Alderson J, Spikings EC, Egginton S, St John JC. Mitochondrial DNA replication during differentiation of murine embryonic stem cells. *J Cell Sci*. 2007;120(Pt 22):4025-4034.

172. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 2001;29(9):e45.
173. Hals IK, Rokstad AM, Strand BL, Oberholzer J, Grill V. Alginate microencapsulation of human islets does not increase susceptibility to acute hypoxia. *J Diabetes Res.* 2013;2013:374925.
174. Calabrese R, Capriotti E, Fariselli P, Martelli PL, Casadio R. Functional annotations improve the predictive score of human disease-related mutations in proteins. *Hum Mutat.* 2009;30(8):1237-1244.
175. Li B, Krishnan VG, Mort ME, et al. Automated inference of molecular mechanisms of disease from amino acid substitutions. *Bioinformatics.* 2009;25(21):2744-2750.
176. Thusberg J, Olatubosun A, Vihinen M. Performance of mutation pathogenicity prediction methods on missense variants. *Hum Mutat.* 2011;32(4):358-368.
177. Camilli F, Borrmann A, Gholizadeh S, te Beek TA, Kuipers RK, Venselaar H. The future of HOPE: what can and cannot be predicted about the molecular effects of a disease causing point mutation in a protein? Vol. 17; 2011.
178. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature.* 2001;409(6822):860-921.
179. Zhang J, Chiodini R, Badr A, Zhang G. The impact of next-generation sequencing on genomics. *J Genet Genomics.* 2011;38(3):95-109.
180. Liu L, Li Y, Li S, et al. Comparison of next-generation sequencing systems. *J Biomed Biotechnol.* 2012;2012:251364.
181. Loman NJ, Misra RV, Dallman TJ, et al. Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol.* 2012;30(5):434-439.
182. Clark MJ, Homer N, O'Connor BD, et al. U87MG decoded: the genomic sequence of a cytogenetically aberrant human cancer cell line. *PLoS Genet.* 2010;6(1):e1000832.
183. Berger MF, Hodis E, Heffernan TP, et al. Melanoma genome sequencing reveals frequent PREX2 mutations. *Nature.* 2012;485(7399):502-506.
184. Ley TJ, Mardis ER, Ding L, et al. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature.* 2008;456(7218):66-72.
185. Mardis ER, Ding L, Dooling DJ, et al. Recurring mutations found by sequencing an acute myeloid leukemia genome. *N Engl J Med.* 2009;361(11):1058-1066.
186. Puente XS, Pinyol M, Quesada V, et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature.* 2011;475(7354):101-105.
187. Chapman MA, Lawrence MS, Keats JJ, et al. Initial genome sequencing and analysis of multiple myeloma. *Nature.* 2011;471(7339):467-472.
188. Drexler HG, Matsuo Y. Malignant hematopoietic cell lines: in vitro models for the study of multiple myeloma and plasma cell leukemia. *Leuk Res.* 2000;24(8):681-703.
189. Turilova VI, Pendina AA, Smirnova TD. [Metaphase chromosomes of human multiple myeloma cells of line U-266: karyotype and morpho-functional characteristics of nucleolar organizer regions]. *Tsitologiya.* 2003;45(6):606-613.
190. Hellman L, Josephson S, Jernberg H, Nilsson K, Pettersson U. Immunoglobulin synthesis in the human myeloma cell line U-266; expression of

- two immunoglobulin heavy chain isotypes (epsilon and alpha) after long-term cultivation in vitro. *Eur J Immunol.* 1988;18(6):905-910.
191. Rothberg JM, Hinz W, Rearick TM, et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature.* 2011;475(7356):348-352.
192. Elliott AM, Radecki J, Moghis B, Li X, Kammesheidt A. Rapid detection of the ACMG/ACOG-recommended 23 CFTR disease-causing mutations using ion torrent semiconductor sequencing. *J Biomol Tech.* 2012;23(1):24-30.
193. Momozawa Y, Mni M, Nakamura K, et al. Resequencing of positional candidates identifies low frequency IL23R coding variants protecting against inflammatory bowel disease. *Nat Genet.* 2011;43(1):43-47.
194. Wheeler DA, Srinivasan M, Egholm M, et al. The complete genome of an individual by massively parallel DNA sequencing. *Nature.* 2008;452(7189):872-876.
195. Rohde H, Qin J, Cui Y, et al. Open-source genomic analysis of Shiga-toxin-producing *E. coli* O104:H4. *N Engl J Med.* 2011;365(8):718-724.
196. Losada L, Varga JJ, Hostetler J, et al. Genome sequencing and analysis of *Yersinia pestis* KIM D27, an avirulent strain exempt from select agent regulation. *PLoS One.* 2011;6(4):e19054.
197. Krypuy M, Newnham GM, Thomas DM, Conron M, Dobrovic A. High resolution melting analysis for the rapid and sensitive detection of mutations in clinical samples: KRAS codon 12 and 13 mutations in non-small cell lung cancer. *BMC Cancer.* 2006;6:295.
198. Hu M, Chilton NB, Gasser RB. Long PCR-based amplification of the entire mitochondrial genome from single parasitic nematodes. *Mol Cell Probes.* 2002;16(4):261-267.
199. Hu M, Jex AR, Campbell BE, Gasser RB. Long PCR amplification of the entire mitochondrial genome from individual helminths for direct sequencing. *Nat Protoc.* 2007;2(10):2339-2344.
200. Meyer M, Stenzel U, Myles S, Prufer K, Hofreiter M. Targeted high-throughput sequencing of tagged nucleic acid samples. *Nucleic Acids Res.* 2007;35(15):e97.
201. Sosa MX, Sivakumar IK, Maragh S, et al. Next-generation sequencing of human mitochondrial reference genomes uncovers high heteroplasmy frequency. *PLoS Comput Biol.* 2012;8(10):e1002737.
202. Gunnarsdottir ED, Li M, Bauchet M, Finstermeier K, Stoneking M. High-throughput sequencing of complete human mtDNA genomes from the Philippines. *Genome Res.* 2011;21(1):1-11.
203. Kirino Y, Yasukawa T, Ohta S, et al. Codon-specific translational defect caused by a wobble modification deficiency in mutant tRNA from a human mitochondrial disease. *Proc Natl Acad Sci U S A.* 2004;101(42):15070-15075.
204. de Wit HM, Westeneng HJ, van Engelen BG, Mudde AH. MIDD or MELAS : that's not the question MIDD evolving into MELAS : a severe phenotype of the m.3243A>G mutation due to paternal co-inheritance of type 2 diabetes and a high heteroplasmy level. *Neth J Med.* 2012;70(10):460-462.
205. Cheung WC, Van Ness B. The bone marrow stromal microenvironment influences myeloma therapeutic response in vitro. *Leukemia.* 2001;15(2):264-271.

206. Zhou Q, Yao Y, Ericson SG. The protein tyrosine phosphatase CD45 is required for interleukin 6 signaling in U266 myeloma cells. *Int J Hematol.* 2004;79(1):63-73.
207. Moreau P, Robillard N, Avet-Loiseau H, et al. Patients with CD45 negative multiple myeloma receiving high-dose therapy have a shorter survival than those with CD45 positive multiple myeloma. *Haematologica.* 2004;89(5):547-551.
208. Mahmoud MS, Ishikawa H, Fujii R, Kawano MM. Induction of CD45 expression and proliferation in U-266 myeloma cell line by interleukin-6. *Blood.* 1998;92(10):3887-3897.
209. Hofmann S, Jaksch M, Bezold R, et al. Population genetics and disease susceptibility: characterization of central European haplogroups by mtDNA gene mutations, correlation with D loop variants and association with disease. *Hum Mol Genet.* 1997;6(11):1835-1846.
210. Yao YG, Kong QP, Salas A, Bandelt HJ. Pseudomitochondrial genome haunts disease studies. *J Med Genet.* 2008;45(12):769-772.
211. Ozawa T, Tanaka M, Ino H, et al. Distinct clustering of point mutations in mitochondrial DNA among patients with mitochondrial encephalomyopathies and with Parkinson's disease. *Biochem Biophys Res Commun.* 1991;176(2):938-946.
212. Rajkumar R, Banerjee J, Gunturi HB, Trivedi R, Kashyap VK. Phylogeny and antiquity of M macrohaplogroup inferred from complete mt DNA sequence of Indian specific lineages. *BMC Evol Biol.* 2005;5:26.
213. Zhao F, Guan M, Zhou X, et al. Leber's hereditary optic neuropathy is associated with mitochondrial ND6 T14502C mutation. *Biochem Biophys Res Commun.* 2009;389(3):466-472.
214. Mkaouar-Rebai E, Chamkha I, Kammoun F, et al. Two new mutations in the MT-TW gene leading to the disruption of the secondary structure of the tRNA(Trp) in patients with Leigh syndrome. *Mol Genet Metab.* 2009;97(3):179-184.
215. Puomila A, Hamalainen P, Kivioja S, et al. Epidemiology and penetrance of Leber hereditary optic neuropathy in Finland. *Eur J Hum Genet.* 2007;15(10):1079-1089.
216. Cooper DN, Krawczak M. Cytosine methylation and the fate of CpG dinucleotides in vertebrate genomes. *Hum Genet.* 1989;83(2):181-188.
217. Chang DD, Clayton DA. Priming of human mitochondrial DNA replication occurs at the light-strand promoter. *Proc Natl Acad Sci U S A.* 1985;82(2):351-355.
218. Montoya J, Christianson T, Levens D, Rabinowitz M, Attardi G. Identification of initiation sites for heavy-strand and light-strand transcription in human mitochondrial DNA. *Proc Natl Acad Sci U S A.* 1982;79(23):7195-7199.
219. Lodeiro MF, Uchida A, Bestwick M, et al. Transcription from the second heavy-strand promoter of human mtDNA is repressed by transcription factor A in vitro. *Proc Natl Acad Sci U S A.* 2012;109(17):6513-6518.
220. Moraes CT, Andreetta F, Bonilla E, Shanske S, DiMauro S, Schon EA. Replication-competent human mitochondrial DNA lacking the heavy-strand promoter region. *Mol Cell Biol.* 1991;11(3):1631-1637.
221. Wanrooij S, Miralles Fuste J, Stewart JB, et al. In vivo mutagenesis reveals that OriL is essential for mitochondrial DNA replication. *EMBO Rep.* 2012;13(12):1130-1137.

222. Park JS, Sharma LK, Li H, et al. A heteroplasmic, not homoplasmic, mitochondrial DNA mutation promotes tumorigenesis via alteration in reactive oxygen species generation and apoptosis. *Hum Mol Genet.* 2009;18(9):1578-1589.
223. Bozic I, Antal T, Ohtsuki H, et al. Accumulation of driver and passenger mutations during tumor progression. *Proc Natl Acad Sci U S A.* 2010;107(43):18545-18550.
224. Larman TC, DePalma SR, Hadjipanayis AG, et al. Spectrum of somatic mitochondrial mutations in five cancers. *Proc Natl Acad Sci U S A.* 2012;109(35):14087-14091.
225. Kumar S, Rajkumar SV, Kimlinger T, Greipp PR, Witzig TE. CD45 expression by bone marrow plasma cells in multiple myeloma: clinical and biological correlations. *Leukemia.* 2005;19(8):1466-1470.
226. Asosingh K, Willems A, Van Riet I, Van Camp B, Vanderkerken K. Delayed in vivo disease progression is associated with high proportions of CD45+ myeloma cells in the 5T2MM murine model. *Cancer Res.* 2003;63(12):3019-3020.
227. Asosingh K, De Raeve H, Menu E, et al. Angiogenic switch during 5T2MM murine myeloma tumorigenesis: role of CD45 heterogeneity. *Blood.* 2004;103(8):3131-3137.
228. Kim D, Park CY, Medeiros BC, Weissman IL. CD19-CD45 low/- CD38 high/CD138+ plasma cells enrich for human tumorigenic myeloma cells. *Leukemia.* 2012;26(12):2530-2537.
229. Asosingh K, De Raeve H, Croucher P, et al. In vivo homing and differentiation characteristics of mature (CD45-) and immature (CD45+) 5T multiple myeloma cells. *Exp Hematol.* 2001;29(1):77-84.
230. Danielson ML, Lill MA. New computational method for prediction of interacting protein loop regions. *Proteins.* 2010;78(7):1748-1759.
231. Betts MJ, Russell, R. B. . Amino Acid Properties and Consequences of Substitutions, in *Bioinformatics for Geneticists: John Wiley & Sons, Ltd, Chichester, UK.; (2003).*
232. Bolsover SR, Hyams, J.S, Shephard, E.A., White, H.A., Wiedemann, C.G. *Protein Structure in Cell Biology: A Short Course, Second Edition. (2004).*
233. Khalimonchuk O, Bestwick M, Meunier B, Watts TC, Winge DR. Formation of the redox cofactor centers during Cox1 maturation in yeast cytochrome oxidase. *Mol Cell Biol.* 2010;30(4):1004-1017.
234. Tsukihara T, Aoyama H, Yamashita E, et al. Structures of metal sites of oxidized bovine heart cytochrome c oxidase at 2.8 Å. *Science.* 1995;269(5227):1069-1074.
235. Cardol P, Lapaille M, Minet P, Franck F, Matagne RF, Remacle C. ND3 and ND4L subunits of mitochondrial complex I, both nucleus encoded in *Chlamydomonas reinhardtii*, are required for activity and assembly of the enzyme. *Eukaryot Cell.* 2006;5(9):1460-1467.
236. Fisher N, Rich PR. A motif for quinone binding sites in respiratory and photosynthetic systems. *J Mol Biol.* 2000;296(4):1153-1162.
237. Damm F, Bunke T, Thol F, et al. Prognostic implications and molecular associations of NADH dehydrogenase subunit 4 (ND4) mutations in acute myeloid leukemia. *Leukemia.* 2012;26(2):289-295.
238. Pulkes T, Eunson L, Patterson V, et al. The mitochondrial DNA G13513A transition in ND5 is associated with a LHON/MELAS overlap

syndrome and may be a frequent cause of MELAS. *Ann Neurol.* 1999;46(6):916-919.

239. Majamaa K, Turkka J, Karppa M, Winqvist S, Hassinen IE. The common MELAS mutation A3243G in mitochondrial DNA among young patients with an occipital brain infarct. *Neurology.* 1997;49(5):1331-1334.

240. Cooper GM. Chapter 10: The Mechanism of Oxidative Phosphorylation. *The Cell: A Molecular Approach* 2nd edition. Sunderland (MA): Sinauer Associates.; 2000.

241. Raje N, Kumar S, Hideshima T, et al. Seliciclib (CYC202 or R-roscovitine), a small-molecule cyclin-dependent kinase inhibitor, mediates activity via down-regulation of Mcl-1 in multiple myeloma. *Blood.* 2005;106(3):1042-1047.

242. Yew TL, Hung YT, Li HY, et al. Enhancement of wound healing by human multipotent stromal cell conditioned medium: the paracrine factors and p38 MAPK activation. *Cell Transplant.* 2011;20(5):693-706.

243. Kudo M, Jono H, Shinriki S, et al. Antitumor effect of humanized anti-interleukin-6 receptor antibody (tocilizumab) on glioma cell proliferation. Laboratory investigation. *J Neurosurg.* 2009;111(2):219-225.

244. Hashizume M, Mihara M. Influence of humanized anti-IL-6R antibody, tocilizumab on the activity of soluble gp130, natural inhibitor of IL-6 signaling. *Rheumatol Int.* 2009;29(4):397-401.

245. Chang KT, Tsai CM, Chiou YC, Chiu CH, Jeng KS, Huang CY. IL-6 induces neuroendocrine dedifferentiation and cell proliferation in non-small cell lung cancer cells. *Am J Physiol Lung Cell Mol Physiol.* 2005;289(3):L446-453.

246. Yadav V, Zhang X, Liu J, et al. Reactivation of mitogen-activated protein kinase (MAPK) pathway by FGF receptor 3 (FGFR3)/Ras mediates resistance to vemurafenib in human B-RAF V600E mutant melanoma. *J Biol Chem.* 2012;287(33):28087-28098.

247. Loffler D, Brocke-Heidrich K, Pfeifer G, et al. Interleukin-6 dependent survival of multiple myeloma cells involves the Stat3-mediated induction of microRNA-21 through a highly conserved enhancer. *Blood.* 2007;110(4):1330-1333.

248. Wen XY, Stewart AK, Sooknanan RR, et al. Identification of c-myc promoter-binding protein and X-box binding protein 1 as interleukin-6 target genes in human multiple myeloma cells. *Int J Oncol.* 1999;15(1):173-178.

249. Martinon F, Chen X, Lee AH, Glimcher LH. TLR activation of the transcription factor XBPI regulates innate immune responses in macrophages. *Nat Immunol.* 2010;11(5):411-418.

250. Zhang B, Potyagaylo V, Fenton RG. IL-6-independent expression of Mcl-1 in human multiple myeloma. *Oncogene.* 2003;22(12):1848-1859.

251. Cavarretta IT, Neuwirt H, Zaki MH, et al. Mcl-1 is regulated by IL-6 and mediates the survival activity of the cytokine in a model of late stage prostate carcinoma. *Adv Exp Med Biol.* 2008;617:547-555.

252. Schwickart M, Huang X, Lill JR, et al. Deubiquitinase USP9X stabilizes MCL1 and promotes tumour cell survival. *Nature.* 2010;463(7277):103-107.

253. Liu H, Ma Y, Cole SM, et al. Serine phosphorylation of STAT3 is essential for Mcl-1 expression and macrophage survival. *Blood.* 2003;102(1):344-352.

254. Bhattacharya S, Ray RM, Johnson LR. STAT3-mediated transcription of Bcl-2, Mcl-1 and c-IAP2 prevents apoptosis in polyamine-depleted cells. *Biochem J.* 2005;392(Pt 2):335-344.
255. Acehan D, Jiang X, Morgan DG, Heuser JE, Wang X, Akey CW. Three-dimensional structure of the apoptosome: implications for assembly, procaspase-9 binding, and activation. *Mol Cell.* 2002;9(2):423-432.
256. Hosgood HD, 3rd, Baris D, Zhang Y, et al. Caspase polymorphisms and genetic susceptibility to multiple myeloma. *Hematol Oncol.* 2008;26(3):148-151.
257. O'Donnell MA, Perez-Jimenez E, Oberst A, et al. Caspase 8 inhibits programmed necrosis by processing CYLD. *Nat Cell Biol.* 2011;13(12):1437-1442.
258. Hakem A, El Ghamrasni S, Maire G, et al. Caspase-8 is essential for maintaining chromosomal stability and suppressing B-cell lymphomagenesis. *Blood.* 2012;119(15):3495-3502.
259. Hong NA, Flannery M, Hsieh SN, Cado D, Pedersen R, Winoto A. Mice lacking Dad1, the defender against apoptotic death-1, express abnormal N-linked glycoproteins and undergo increased embryonic apoptosis. *Dev Biol.* 2000;220(1):76-84.
260. Apte SS, Mattei MG, Seldin MF, Olsen BR. The highly conserved defender against the death 1 (DAD1) gene maps to human chromosome 14q11-q12 and mouse chromosome 14 and has plant and nematode homologs. *FEBS Lett.* 1995;363(3):304-306.
261. Feng JY, Johnson AA, Johnson KA, Anderson KS. Insights into the molecular mechanism of mitochondrial toxicity by AIDS drugs. *The Journal of biological chemistry.* 2001;276(26):23832-23837.
262. Lewis W. Pharmacogenomics, toxicogenomics, and DNA polymerase gamma. *J Infect Dis.* 2007;195(10):1399-1401.
263. Hartl DL, Ruvolo, M. Chapter 6: Molecular Biology of DNA Replication and Recombination. *Genetics* 8th edition. Burlington, MA: Jones & Bartlett Learning; 2012:p.212.
264. van der Vliet PC, Kwant MM. Role of DNA polymerase gamma in adenovirus DNA replication. Mechanism of inhibition by 2',3'-dideoxynucleoside 5'-triphosphates. *Biochemistry.* 1981;20(9):2628-2632.
265. Hanes JW, Johnson KA. Exonuclease removal of dideoxycytidine (zalcitabine) by the human mitochondrial DNA polymerase. *Antimicrobial agents and chemotherapy.* 2008;52(1):253-258.
266. Sottnik JL, Lori JC, Rose BJ, Thamm DH. Glycolysis inhibition by 2-deoxy-D-glucose reverts the metastatic phenotype in vitro and in vivo. *Clin Exp Metastasis.* 2011;28(8):865-875.
267. Barban S, Schulze HO. The effects of 2-deoxyglucose on the growth and metabolism of cultured human cells. *J Biol Chem.* 1961;236:1887-1890.
268. Dwarakanath BS, Singh D, Banerji AK, et al. Clinical studies for improving radiotherapy with 2-deoxy-D-glucose: present status and future prospects. *J Cancer Res Ther.* 2009;5 Suppl 1:S21-26.
269. Cheng G, Zielonka J, Dranka BP, et al. Mitochondria-targeted drugs synergize with 2-deoxyglucose to trigger breast cancer cell death. *Cancer Res.* 2012;72(10):2634-2644.
270. Birkus G, Hitchcock MJ, Cihlar T. Assessment of mitochondrial toxicity in human cells treated with tenofovir: comparison with other nucleoside reverse

- transcriptase inhibitors. *Antimicrobial agents and chemotherapy*. 2002;46(3):716-723.
271. Chandel NS, Schumacker PT. Cells depleted of mitochondrial DNA (rho0) yield insight into physiological mechanisms. *FEBS letters*. 1999;454(3):173-176.
272. Lund KC, Peterson LL, Wallace KB. Absence of a universal mechanism of mitochondrial toxicity by nucleoside analogs. *Antimicrob Agents Chemother*. 2007;51(7):2531-2539.
273. Moro L, Arbini AA, Marra E, Greco M. Mitochondrial DNA depletion reduces PARP-1 levels and promotes progression of the neoplastic phenotype in prostate carcinoma. *Cell Oncol*. 2008;30(4):307-322.
274. Higuchi M, Kudo T, Suzuki S, et al. Mitochondrial DNA determines androgen dependence in prostate cancer cell lines. *Oncogene*. 2006;25(10):1437-1445.
275. Dickinson A, Yeung KY, Donoghue J, et al. The regulation of mitochondrial DNA copy number in glioblastoma cells. *Cell Death Differ*. 2013;20(12):1644-1653.
276. Verschoor ML, Wilson LA, Verschoor CP, Singh G. Ets-1 regulates energy metabolism in cancer cells. *PLoS One*. 2010;5(10):e13565.
277. Anna Korzynska MZ. A Method of Estimation of the Cell Doubling Time on Basis of the Cell Culture Monitoring Data. *Biocybernetics and Biomedical Engineering*. 2008;28(4):75-82.
278. Meierhofer D, Mayr JA, Foetschl U, et al. Decrease of mitochondrial DNA content and energy metabolism in renal cell carcinoma. *Carcinogenesis*. 2004;25(6):1005-1010.
279. Simonnet H, Alazard N, Pfeiffer K, et al. Low mitochondrial respiratory chain content correlates with tumor aggressiveness in renal cell carcinoma. *Carcinogenesis*. 2002;23(5):759-768.
280. Correia RL, Oba-Shinjo SM, Uno M, Huang N, Marie SK. Mitochondrial DNA depletion and its correlation with TFAM, TFB1M, TFB2M and POLG in human diffusely infiltrating astrocytomas. *Mitochondrion*. 2011;11(1):48-53.
281. Munshi NC, Hideshima T, Carrasco D, et al. Identification of genes modulated in multiple myeloma using genetically identical twin samples. *Blood*. 2004;103(5):1799-1806.
282. Jelluma N, Yang X, Stokoe D, Evan GI, Dansen TB, Haas-Kogan DA. Glucose withdrawal induces oxidative stress followed by apoptosis in glioblastoma cells but not in normal human astrocytes. *Mol Cancer Res*. 2006;4(5):319-330.
283. Wanders RJ, Ruiten JP, L IJ, Waterham HR, Houten SM. The enzymology of mitochondrial fatty acid beta-oxidation and its application to follow-up analysis of positive neonatal screening results. *J Inherit Metab Dis*. 2010;33(5):479-494.
284. Cipolat S, Martins de Brito O, Dal Zilio B, Scorrano L. OPA1 requires mitofusin 1 to promote mitochondrial fusion. *Proc Natl Acad Sci U S A*. 2004;101(45):15927-15932.
285. Facucho-Oliveira JM, St John JC. The relationship between pluripotency and mitochondrial DNA proliferation during early embryo development and embryonic stem cell differentiation. *Stem cell reviews*. 2009;5(2):140-158.

286. Kelly RD, Sumer H, McKenzie M, et al. The effects of nuclear reprogramming on mitochondrial DNA replication. *Stem cell reviews*. 2013;9(1):1-15.
287. Kelly RD, Mahmud A, McKenzie M, Trounce IA, St John JC. Mitochondrial DNA copy number is regulated in a tissue specific manner by DNA methylation of the nuclear-encoded DNA polymerase gamma A. *Nucleic acids research*. 2012;40(20):10124-10138.
288. St John JC, Ramalho-Santos J, Gray HL, et al. The expression of mitochondrial DNA transcription factors during early cardiomyocyte in vitro differentiation from human embryonic stem cells. *Cloning and stem cells*. 2005;7(3):141-153.
289. Wu CW, Yin PH, Hung WY, et al. Mitochondrial DNA mutations and mitochondrial DNA depletion in gastric cancer. *Genes, chromosomes & cancer*. 2005;44(1):19-28.
290. Dey R, Moraes CT. Lack of oxidative phosphorylation and low mitochondrial membrane potential decrease susceptibility to apoptosis and do not modulate the protective effect of Bcl-x(L) in osteosarcoma cells. *The Journal of biological chemistry*. 2000;275(10):7087-7094.
291. Amuthan G, Biswas G, Ananatheerthavarada HK, Vijayasathy C, Shephard HM, Avadhani NG. Mitochondrial stress-induced calcium signaling, phenotypic changes and invasive behavior in human lung carcinoma A549 cells. *Oncogene*. 2002;21(51):7839-7849.
292. Alberio S, Mineri R, Tiranti V, Zeviani M. Depletion of mtDNA: syndromes and genes. *Mitochondrion*. 2007;7(1-2):6-12.
293. Seyfried TN, Shelton LM. Cancer as a metabolic disease. *Nutr Metab (Lond)*. 2010;7:7.
294. Traven A, Wong JM, Xu D, Sopta M, Ingles CJ. Interorganellar communication. Altered nuclear gene expression profiles in a yeast mitochondrial dna mutant. *The Journal of biological chemistry*. 2001;276(6):4020-4027.
295. Bulst S, Holinski-Feder E, Payne B, et al. In vitro supplementation with deoxynucleoside monophosphates rescues mitochondrial DNA depletion. *Molecular genetics and metabolism*. 2012;107(1-2):95-103.
296. Poteet E, Choudhury GR, Winters A, et al. Reversing the warburg effect as a treatment for glioblastoma. *J Biol Chem*. 2013;288(13):9153-9164.
297. Peterson CW, Stoltzman CA, Sighinolfi MP, Han KS, Ayer DE. Glucose controls nuclear accumulation, promoter binding, and transcriptional activity of the MondoA-Mlx heterodimer. *Mol Cell Biol*. 2010;30(12):2887-2895.
298. Stoltzman CA, Peterson CW, Breen KT, Muoio DM, Billin AN, Ayer DE. Glucose sensing by MondoA:Mlx complexes: a role for hexokinases and direct regulation of thioredoxin-interacting protein expression. *Proc Natl Acad Sci U S A*. 2008;105(19):6912-6917.
299. Sans CL, Satterwhite DJ, Stoltzman CA, Breen KT, Ayer DE. MondoA-Mlx heterodimers are candidate sensors of cellular energy status: mitochondrial localization and direct regulation of glycolysis. *Mol Cell Biol*. 2006;26(13):4863-4871.
300. Xiao Y, Decker PA, Rice T, et al. SSBP2 variants are associated with survival in glioblastoma patients. *Clin Cancer Res*. 2012;18(11):3154-3162.
301. Gan HK, Kaye AH, Luwor RB. The EGFRvIII variant in glioblastoma multiforme. *J Clin Neurosci*. 2009;16(6):748-754.

302. Maletinska L, Blakely EA, Bjornstad KA, Deen DF, Knoff LJ, Forte TM. Human glioblastoma cell lines: levels of low-density lipoprotein receptor and low-density lipoprotein receptor-related protein. *Cancer Res.* 2000;60(8):2300-2303.
303. Galli R, Binda E, Orfanelli U, et al. Isolation and characterization of tumorigenic, stem-like neural precursors from human glioblastoma. *Cancer Res.* 2004;64(19):7011-7021.
304. Ruano Y, Ribalta T, de Lope AR, et al. Worse outcome in primary glioblastoma multiforme with concurrent epidermal growth factor receptor and p53 alteration. *Am J Clin Pathol.* 2009;131(2):257-263.
305. Pandita A, Aldape KD, Zadeh G, Guha A, James CD. Contrasting in vivo and in vitro fates of glioblastoma cell subpopulations with amplified EGFR. *Genes Chromosomes Cancer.* 2004;39(1):29-36.
306. Zhu TS, Costello MA, Talsma CE, et al. Endothelial cells create a stem cell niche in glioblastoma by providing NOTCH ligands that nurture self-renewal of cancer stem-like cells. *Cancer Res.* 2011;71(18):6061-6072.
307. He J, Liu Y, Xie X, et al. Identification of cell surface glycoprotein markers for glioblastoma-derived stem-like cells using a lectin microarray and LC-MS/MS approach. *J Proteome Res.* 2010;9(5):2565-2572.
308. Gore A, Li Z, Fung HL, et al. Somatic coding mutations in human induced pluripotent stem cells. *Nature.* 2011;471(7336):63-67.
309. Fearnley IM, Walker JE. Conservation of sequences of subunits of mitochondrial complex I and their relationships with other proteins. *Biochim Biophys Acta.* 1992;1140(2):105-134.
310. Shoffner JM, Lott MT, Lezza AM, Seibel P, Ballinger SW, Wallace DC. Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a mitochondrial DNA tRNA(Lys) mutation. *Cell.* 1990;61(6):931-937.
311. Goto Y, Nonaka I, Horai S. A mutation in the tRNA(Leu)(UUR) gene associated with the MELAS subgroup of mitochondrial encephalomyopathies. *Nature.* 1990;348(6302):651-653.
312. Moslemi AR, Darin N. Molecular genetic and clinical aspects of mitochondrial disorders in childhood. *Mitochondrion.* 2007;7(4):241-252.
313. Bandelt HJ, Quintana-Murci L, Salas A, Macaulay V. The fingerprint of phantom mutations in mitochondrial DNA data. *Am J Hum Genet.* 2002;71(5):1150-1160.
314. David A, Pike TJB, Sjana L, Bauer and Elizabeth (Blake) Kipp. MtDNA Haplotype T Phylogeny Based On Full Mitochondrial Sequence. *Journal of Genetic Genealogy.* 2010;6(1):1-24.
315. Mambo E, Gao X, Cohen Y, Guo Z, Talalay P, Sidransky D. Electrophile and oxidant damage of mitochondrial DNA leading to rapid evolution of homoplasmic mutations. *Proc Natl Acad Sci U S A.* 2003;100(4):1838-1843.
316. Bai Y, Hu P, Park JS, et al. Genetic and functional analysis of mitochondrial DNA-encoded complex I genes. *Ann N Y Acad Sci.* 2004;1011:272-283.
317. Calabrese C, Iommarini L, Kurelac I, et al. Respiratory complex I is essential to induce a Warburg profile in mitochondria-defective tumor cells. *Cancer & Metabolism.* 2013;1(1):1-15.
318. Bai Y, Attardi G. The mtDNA-encoded ND6 subunit of mitochondrial NADH dehydrogenase is essential for the assembly of the membrane arm and the respiratory function of the enzyme. *EMBO J.* 1998;17(16):4848-4858.

319. Lopez JV, Culver M, Stephens JC, Johnson WE, O'Brien SJ. Rates of nuclear and cytoplasmic mitochondrial DNA sequence divergence in mammals. *Molecular Biology and Evolution*. 1997;14(3):277-286.
320. Ishikawa K, Hayashi J. A novel function of mtDNA: its involvement in metastasis. *Ann N Y Acad Sci*. 2010;1201:40-43.
321. Khaidakov M, Shmookler Reis RJ. Possibility of selection against mtDNA mutations in tumors. *Mol Cancer*. 2005;4:36.
322. Zhou Y, Shingu T, Feng L, et al. Metabolic alterations in highly tumorigenic glioblastoma cells: preference for hypoxia and high dependency on glycolysis. *J Biol Chem*. 2011;286(37):32843-32853.
323. Venugopal C, Wang XS, Manoranjan B, et al. GBM secretome induces transient transformation of human neural precursor cells. *J Neurooncol*. 2012;109(3):457-466.
324. Yu M, Zhou Y, Shi Y, et al. Reduced mitochondrial DNA copy number is correlated with tumor progression and prognosis in Chinese breast cancer patients. *IUBMB Life*. 2007;59(7):450-457.
325. Wang Y, Liu VW, Xue WC, Cheung AN, Ngan HY. Association of decreased mitochondrial DNA content with ovarian cancer progression. *Br J Cancer*. 2006;95(8):1087-1091.
326. Enriquez JA, Fernandez-Silva P, Perez-Martos A, Lopez-Perez MJ, Montoya J. The synthesis of mRNA in isolated mitochondria can be maintained for several hours and is inhibited by high levels of ATP. *Eur J Biochem*. 1996;237(3):601-610.
327. Wolburg H, Noell S, Fallier-Becker P, Mack AF, Wolburg-Buchholz K. The disturbed blood-brain barrier in human glioblastoma. *Mol Aspects Med*. 2012;33(5-6):579-589.
328. Abbott NJ, Ronnback L, Hansson E. Astrocyte-endothelial interactions at the blood-brain barrier. *Nat Rev Neurosci*. 2006;7(1):41-53.
329. McCaslin AF, Chen BR, Radosevich AJ, Cauli B, Hillman EM. In vivo 3D morphology of astrocyte-vasculature interactions in the somatosensory cortex: implications for neurovascular coupling. *J Cereb Blood Flow Metab*. 2011;31(3):795-806.
330. Tchaicha JH, Reyes SB, Shin J, Hossain MG, Lang FF, McCarty JH. Glioblastoma angiogenesis and tumor cell invasiveness are differentially regulated by beta8 integrin. *Cancer Res*. 2011;71(20):6371-6381.
331. Giese A, Kluwe L, Laube B, Meissner H, Berens ME, Westphal M. Migration of human glioma cells on myelin. *Neurosurgery*. 1996;38(4):755-764.
332. Gilbertson RJ, Rich JN. Making a tumour's bed: glioblastoma stem cells and the vascular niche. *Nat Rev Cancer*. 2007;7(10):733-736.
333. Rampazzo E, Persano L, Pistollato F, et al. Wnt activation promotes neuronal differentiation of glioblastoma. *Cell Death Dis*. 2013;4:e500.
334. Campos B, Wan F, Farhadi M, et al. Differentiation therapy exerts antitumor effects on stem-like glioma cells. *Clin Cancer Res*. 2010;16(10):2715-2728.
335. Genc S, Kurnaz IA, Ozilgen M. Astrocyte-neuron lactate shuttle may boost more ATP supply to the neuron under hypoxic conditions--in silico study supported by in vitro expression data. *BMC Syst Biol*. 2011;5:162.
336. Payne VA, Arden C, Wu C, Lange AJ, Agius L. Dual role of phosphofructokinase-2/fructose biphosphatase-2 in regulating the

- compartmentation and expression of glucokinase in hepatocytes. *Diabetes*. 2005;54(7):1949-1957.
337. Gallagher CN, Carpenter KL, Grice P, et al. The human brain utilizes lactate via the tricarboxylic acid cycle: a <sup>13</sup>C-labelled microdialysis and high-resolution nuclear magnetic resonance study. *Brain*. 2009;132(Pt 10):2839-2849.
338. Belanger M, Allaman I, Magistretti PJ. Brain energy metabolism: focus on astrocyte-neuron metabolic cooperation. *Cell Metab*. 2011;14(6):724-738.
339. Kann O, Kovacs R. Mitochondria and neuronal activity. *Am J Physiol Cell Physiol*. 2007;292(2):C641-657.
340. Riester M, Stephan-Otto Attolini C, Downey RJ, Singer S, Michor F. A differentiation-based phylogeny of cancer subtypes. *PLoS Comput Biol*. 2010;6(5):e1000777.
341. Shimizu T, Okayama A, Inoue T, Takeda K. Analysis of gene expression during staurosporine-induced neuronal differentiation of human prostate cancer cells. *Oncol Rep*. 2005;14(2):441-448.
342. Drabsch Y, Robert RG, Gonda TJ. MYB suppresses differentiation and apoptosis of human breast cancer cells. *Breast Cancer Res*. 2010;12(4):R55.
343. Hall CN, Klein-Flugge MC, Howarth C, Attwell D. Oxidative phosphorylation, not glycolysis, powers presynaptic and postsynaptic mechanisms underlying brain information processing. *J Neurosci*. 2012;32(26):8940-8951.
344. Mehrabian Z, Liu LI, Fiskum G, Rapoport SI, Chandrasekaran K. Regulation of mitochondrial gene expression by energy demand in neural cells. *J Neurochem*. 2005;93(4):850-860.
345. Lantos PL, Cox DJ. The origin of experimental brain tumours: a sequential study. *Experientia*. 1976;32(11):1467-1468.
346. Friedmann-Morvinski D, Bushong EA, Ke E, et al. Dedifferentiation of Neurons and Astrocytes by Oncogenes Can Induce Gliomas in Mice. *Science*. 2012;338(6110):1080-1084.
347. Almeida A, Almeida J, Bolanos JP, Moncada S. Different responses of astrocytes and neurons to nitric oxide: the role of glycolytically generated ATP in astrocyte protection. *Proc Natl Acad Sci U S A*. 2001;98(26):15294-15299.
348. Herrero-Mendez A, Almeida A, Fernandez E, Maestre C, Moncada S, Bolanos JP. The bioenergetic and antioxidant status of neurons is controlled by continuous degradation of a key glycolytic enzyme by APC/C-Cdh1. *Nat Cell Biol*. 2009;11(6):747-752.
349. Vilchez D, Ros S, Cifuentes D, et al. Mechanism suppressing glycogen synthesis in neurons and its demise in progressive myoclonus epilepsy. *Nat Neurosci*. 2007;10(11):1407-1413.
350. Chen Y, Cairns R, Papandreou I, Koong A, Denko NC. Oxygen consumption can regulate the growth of tumors, a new perspective on the Warburg effect. *PLoS One*. 2009;4(9):e7033.
351. Dipaolo JA, Moore GE. The influence of prolonged oxygen change on the formation of spontaneous and induced mouse cancer. *Cancer Res*. 1959;19:1175-1180.
352. Buchler P, Reber HA, Lavey RS, et al. Tumor hypoxia correlates with metastatic tumor growth of pancreatic cancer in an orthotopic murine model. *J Surg Res*. 2004;120(2):295-303.
353. Ward RJ, Dirks PB. Cancer stem cells: at the headwaters of tumor development. *Annu Rev Pathol*. 2007;2:175-189.

354. O'Flaherty JD, Barr M, Fennell D, et al. The cancer stem-cell hypothesis: its emerging role in lung cancer biology and its relevance for future therapy. *J Thorac Oncol.* 2012;7(12):1880-1890.
355. Reya T, Morrison SJ, Clarke MF, Weissman IL. Stem cells, cancer, and cancer stem cells. *Nature.* 2001;414(6859):105-111.
356. Vogelstein B, Fearon ER, Hamilton SR, et al. Genetic alterations during colorectal-tumor development. *N Engl J Med.* 1988;319(9):525-532.
357. Schwitalla S, Fingerle AA, Cammareri P, et al. Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. *Cell.* 2013;152(1-2):25-38.
358. Prigione A, Fauler B, Lurz R, Lehrach H, Adjaye J. The senescence-related mitochondrial/oxidative stress pathway is repressed in human induced pluripotent stem cells. *Stem Cells.* 2010;28(4):721-733.
359. Suhr ST, Chang EA, Tjong J, et al. Mitochondrial rejuvenation after induced pluripotency. *PLoS One.* 2010;5(11):e14095.
360. Smith NG, Webster MT, Ellegren H. Deterministic mutation rate variation in the human genome. *Genome Res.* 2002;12(9):1350-1356.
361. Pesole G, Gissi C, De Chirico A, Saccone C. Nucleotide substitution rate of mammalian mitochondrial genomes. *J Mol Evol.* 1999;48(4):427-434.
362. Paumard P, Vaillier J, Couly B, et al. The ATP synthase is involved in generating mitochondrial cristae morphology. *EMBO J.* 2002;21(3):221-230.
363. Nakada K, Inoue K, Hayashi J. Interaction theory of mammalian mitochondria. *Biochem Biophys Res Commun.* 2001;288(4):743-746.
364. Moghadam AA, Ebrahimie E, Taghavi SM, et al. How the nucleus and mitochondria communicate in energy production during stress: nuclear MtATP6, an early-stress responsive gene, regulates the mitochondrial F(1)F(0)-ATP synthase complex. *Mol Biotechnol.* 2013;54(3):756-769.
365. Spees JL, Olson SD, Whitney MJ, Prockop DJ. Mitochondrial transfer between cells can rescue aerobic respiration. *Proc Natl Acad Sci U S A.* 2006;103(5):1283-1288.
366. Thyssen G, Svab Z, Maliga P. Cell-to-cell movement of plastids in plants. *Proc Natl Acad Sci U S A.* 2012;109(7):2439-2443.
367. Kirches E, Krause G, Warich-Kirches M, et al. High frequency of mitochondrial DNA mutations in glioblastoma multiforme identified by direct sequence comparison to blood samples. *Int J Cancer.* 2001;93(4):534-538.
368. Kirches E, Krause G, Weis S, Mawrin C, Dietzmann K. Comparison between mitochondrial DNA sequences in low grade astrocytomas and corresponding blood samples. *Mol Pathol.* 2002;55(3):204-206.
369. Kiebish MA, Seyfried TN. Absence of pathogenic mitochondrial DNA mutations in mouse brain tumors. *BMC Cancer.* 2005;5:102.
370. Mangiola A, Saulnier N, De Bonis P, et al. Gene expression profile of glioblastoma peritumoral tissue: an ex vivo study. *PLoS One.* 2013;8(3):e57145.
371. Trimboli AJ, Cantemir-Stone CZ, Li F, et al. Pten in stromal fibroblasts suppresses mammary epithelial tumours. *Nature.* 2009;461(7267):1084-1091.
372. Bronisz A, Godlewski J, Wallace JA, et al. Reprogramming of the tumour microenvironment by stromal PTEN-regulated miR-320. *Nat Cell Biol.* 2012;14(2):159-167.
373. Pavlides S, Whitaker-Menezes D, Castello-Cros R, et al. The reverse Warburg effect: aerobic glycolysis in cancer associated fibroblasts and the tumor stroma. *Cell Cycle.* 2009;8(23):3984-4001.

374. Bonuccelli G, Whitaker-Menezes D, Castello-Cros R, et al. The reverse Warburg effect: glycolysis inhibitors prevent the tumor promoting effects of caveolin-1 deficient cancer associated fibroblasts. *Cell Cycle*. 2010;9(10):1960-1971.
375. Seidel-Rogol BL, Shadel GS. Modulation of mitochondrial transcription in response to mtDNA depletion and repletion in HeLa cells. *Nucleic Acids Res*. 2002;30(9):1929-1934.
376. Fisher RP, Lisowsky T, Breen GA, Clayton DA. A rapid, efficient method for purifying DNA-binding proteins. Denaturation-renaturation chromatography of human and yeast mitochondrial extracts. *J Biol Chem*. 1991;266(14):9153-9160.
377. Rubio-Cosials A, Sidow JF, Jimenez-Menendez N, et al. Human mitochondrial transcription factor A induces a U-turn structure in the light strand promoter. *Nat Struct Mol Biol*. 2011;18(11):1281-1289.
378. Dairaghi DJ, Shadel GS, Clayton DA. Human mitochondrial transcription factor A and promoter spacing integrity are required for transcription initiation. *Biochim Biophys Acta*. 1995;1271(1):127-134.
379. Rorbach J, Minczuk M. The post-transcriptional life of mammalian mitochondrial RNA. *Biochem J*. 2012;444(3):357-373.
380. Poyton RO, McEwen JE. Crosstalk between nuclear and mitochondrial genomes. *Annu Rev Biochem*. 1996;65:563-607.
381. Augustin S, Nolden M, Muller S, Hardt O, Arnold I, Langer T. Characterization of peptides released from mitochondria: evidence for constant proteolysis and peptide efflux. *The Journal of biological chemistry*. 2005;280(4):2691-2699.
382. Arnold I, Wagner-Ecker M, Ansorge W, Langer T. Evidence for a novel mitochondria-to-nucleus signalling pathway in respiring cells lacking i-AAA protease and the ABC-transporter Mdl1. *Gene*. 2006;367:74-88.
383. Young L, Leonhard K, Tatsuta T, Trowsdale J, Langer T. Role of the ABC transporter Mdl1 in peptide export from mitochondria. *Science*. 2001;291(5511):2135-2138.
384. Parsons XH, Garcia SN, Pillus L, Kadonaga JT. Histone deacetylation by Sir2 generates a transcriptionally repressed nucleoprotein complex. *Proceedings of the National Academy of Sciences of the United States of America*. 2003;100(4):1609-1614.
385. Olgun A, Akman S. Mitochondrial DNA-deficient models and aging. *Ann N Y Acad Sci*. 2007;1100:241-245.
386. Dickinson A, Yeung KY, Donoghue J, et al. The regulation of mitochondrial DNA copy number in glioblastoma cells. *Cell Death Differ*. 2013.

# **Appendices**

### **i. Further information on sequence quality from the Ion Torrent**

To describe in detail the quality of the sequencing results obtained from the Ion Torrent, the data from one of the patient multiple myeloma samples (sample S219, sequenced in **Table 3.6**) was included as a representative example.

#### **Quality reports on the raw sequencing data**

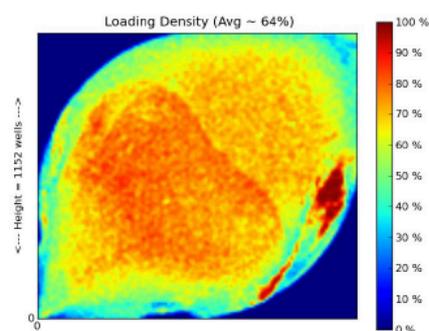
All sequencing runs performed using the Ion Torrent were automatically analysed by the pre-installed Ion Torrent Suite Software. The quality reports that were produced at the end of each run provide information on the loading density of the Ion chip, in the form of a heat map. This provides an indication of how many of the microwells on the Ion chip were successfully occupied by the Ion Sphere Particles (ISPs), and their distributions across the Ion chip. Quantitative data was also provided to inform the user of how successful the emulsion PCR was in amplifying the DNA template on these ISPs.

Using the S219 myeloma sample as an example, typically there was an even loading of the ISPs across the Ion chip (**Figure A1**). The proportion of microwells containing an ISP, was usually above 60%, and of these ISPs the proportion of 'live ISPs' were generally always above 80%, since template-positive ISPs were enriched for prior to the sequencing reaction (section 2.18.10). The percentage of live ISPs is therefore representative of the enrichment efficiency, and is measured by the presence of a key sequence 'TCAG' attached to the ends of successfully amplified DNA template, located on the 3' adaptor itself. Further information on these template-positive ISPs was also provided to inform users of the final percentage of library reads that were

included in the sequencing output (FASTQ) file. The filtering process to produce the final percentage of acceptable library reads was automated on the Ion Torrent Suite Software, and was based on the exclusion of low quality DNA templates, which yielded low sequencing signal or polyclonal reads.

	Count	Percentage
Total Addressable Wells	1,262,519	
▸ Wells with ISPs	828,251	66%
▸ Live ISPs	739,924	89%
▸ Test Fragment ISPs	7,796	1%
▸ Library ISPs	732,128	99%

	Count	Percentage
Library ISPs / Percent Enrichment	732,128	89%
▸ Filtered: Too short	20,382	3%
▸ Filtered: Keypass failure	61,493	8%
▸ Filtered: Mixed / Polyclonal	37,920	5%
▸ Filtered: Low Signal	13,029	2%
▸ Filtered: Poor Signal Profile	241,645	33%
▸ Filtered: 3' Adapter trim	382	<1%
▸ Filtered: 3' Quality trim	0	<1%
▸ <b>Final Library Reads</b>	<b>356,977</b>	<b>49%</b>



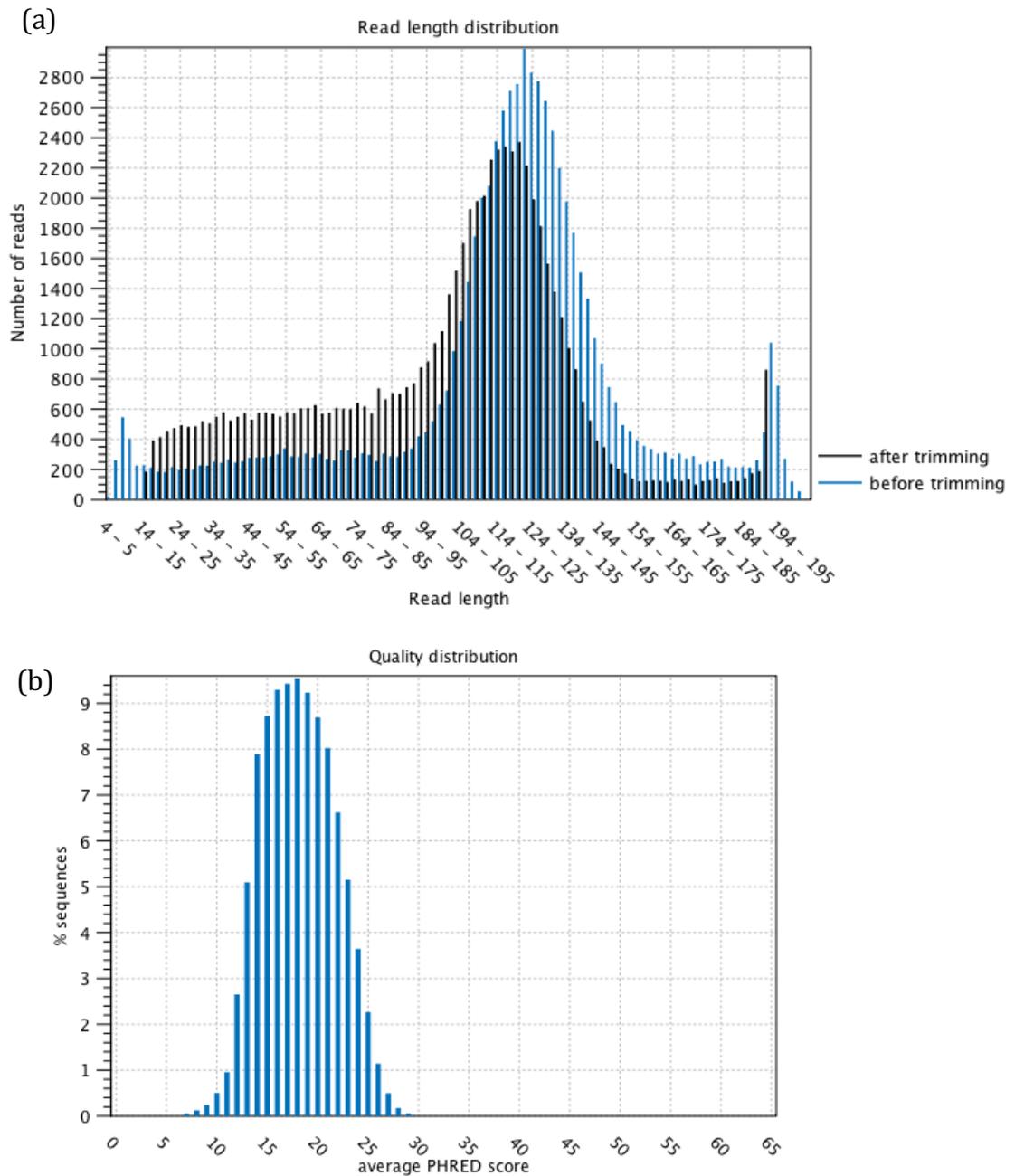
**Figure A1:** *Quality report on the raw sequencing data obtained from the Ion Torrent Suite Software.*

Any sequencing runs that reported uneven loading of the microwells on the Ion chip, or even unsatisfactory statistics on, for example, the percentage of live ISPs, were re-run on the Ion Torrent.

### Sequence analysis before and after trimming

As mentioned in Section 2.19, all raw sequencing reads from the Ion Torrent were subjected to filtering using the CLC Genomics Workbench software. The purpose of this was to select for reads above a Phred quality score of 15 and to discard short reads below a length of 15bp. All sequences were also trimmed on each end by 1bp to prevent reporting nucleotides that may be ambiguous. An

example showing changes to the distribution of read lengths before and after quality-based trimming is shown in **Figure A2a**, with the typical distribution of Phred scores before trimming shown in **Figure A2b**. Quantitative data to support **Figure A2a** are presented in **Table (i)**.



**Figure A2:** (a) Read length distribution for the S219 multiple myeloma sample, before and after trimming. (b) Distribution of the Phred quality scores before trimming. Images taken from CLC Genomics Workbench analysis software.

**Table (i):** Detailed sequencing results before and after trimming (sample S219)

Trim	Input reads	No trim	Trimmed	Nothing left or Discarded
Trim on quality	68,916	26,444	42,462	10
Ambiguity trim	68,906	68,906	0	0
Trim ends	68,906	0	68,863	43
Filter on length	68,863	66,088	0	2,775

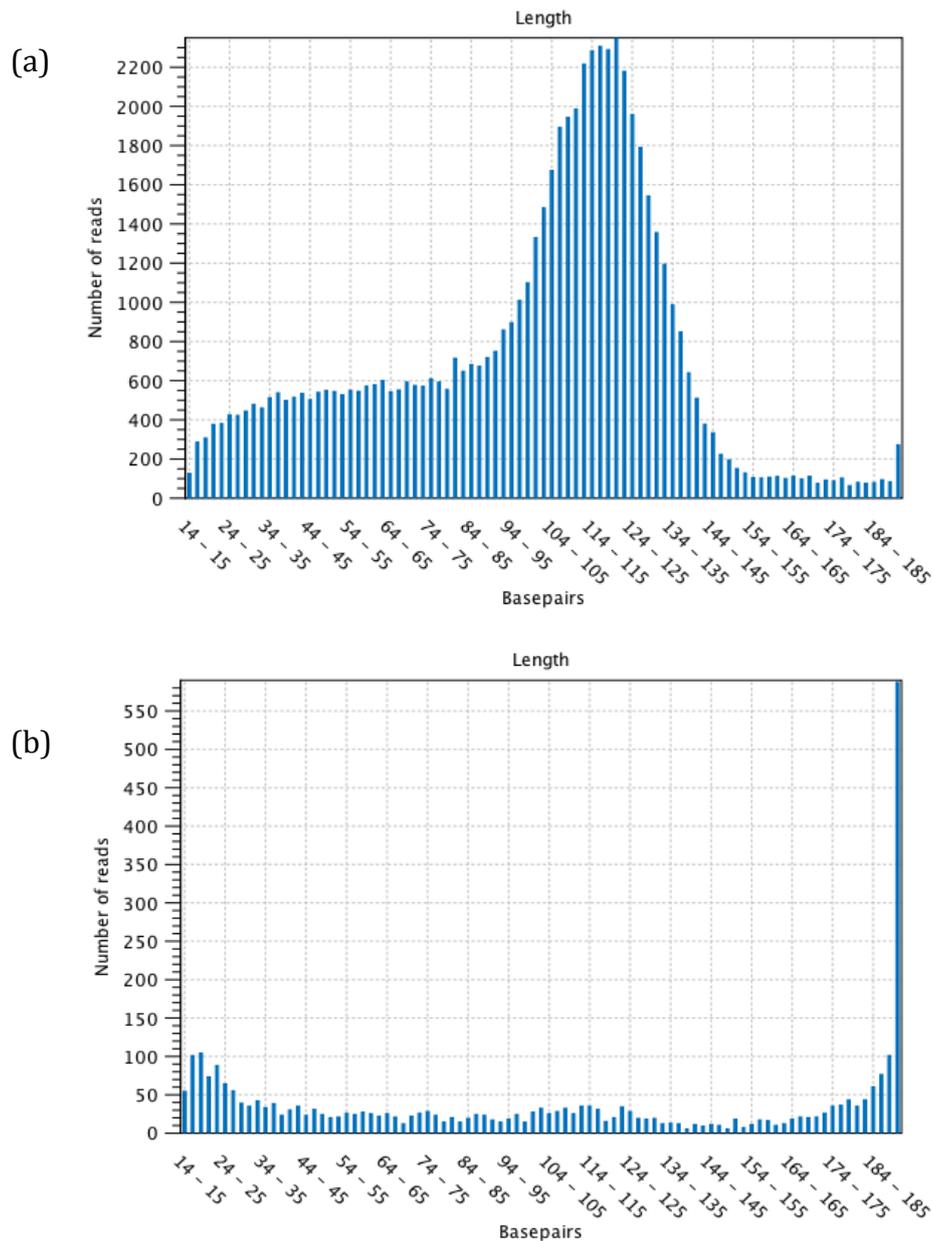
Values from **Table (i)** demonstrated that by applying analysis parameters that filter the raw sequencing data based on their Phred quality scores, low quality reads were discarded to improve the reliability and accuracy of the data prior to SNP analysis. From the example above, which is representative of the trimming outcomes from other samples, approximately a third of all reads (26,444) already met the quality criteria for read inclusion. Results also showed that the data retrieved did not yield any ambiguous nucleotides, indicating precision in the sequencing chemistry used by the Ion Torrent, which relies on the detection of pH changes in the microwells of the Ion chip. Finally, it can be seen that in total, 2,775 reads were completely discarded upon filtering reads of less than 15bp.

Overall, the profile of read length distribution after trimming shows little change compared to the profile of read length distributions before trimming (**Figure A2a**). The exception is that there are a greater number of shorter reads (between 15bp and 100bp) accompanied with a slight fall in the number of longer reads (from 140bp to 200bp) after the trimming process.

### **Sequence analysis of the mapped data**

All trimmed reads were mapped to the human mitochondrial Cambridge Reference Sequence (NC\_012920.1). Using the myeloma S219 data as a

representative sample, almost all (95.12%) of the trimmed reads were successfully aligned to the reference genome (**Figure A3a**). Those that were not mapped (4.88%) represented the proportion of trimmed reads that did not share at least 80% identity with the reference sequence (**Figure A3b**), which was part of the inclusion criteria for reads as detailed in section 2.19.

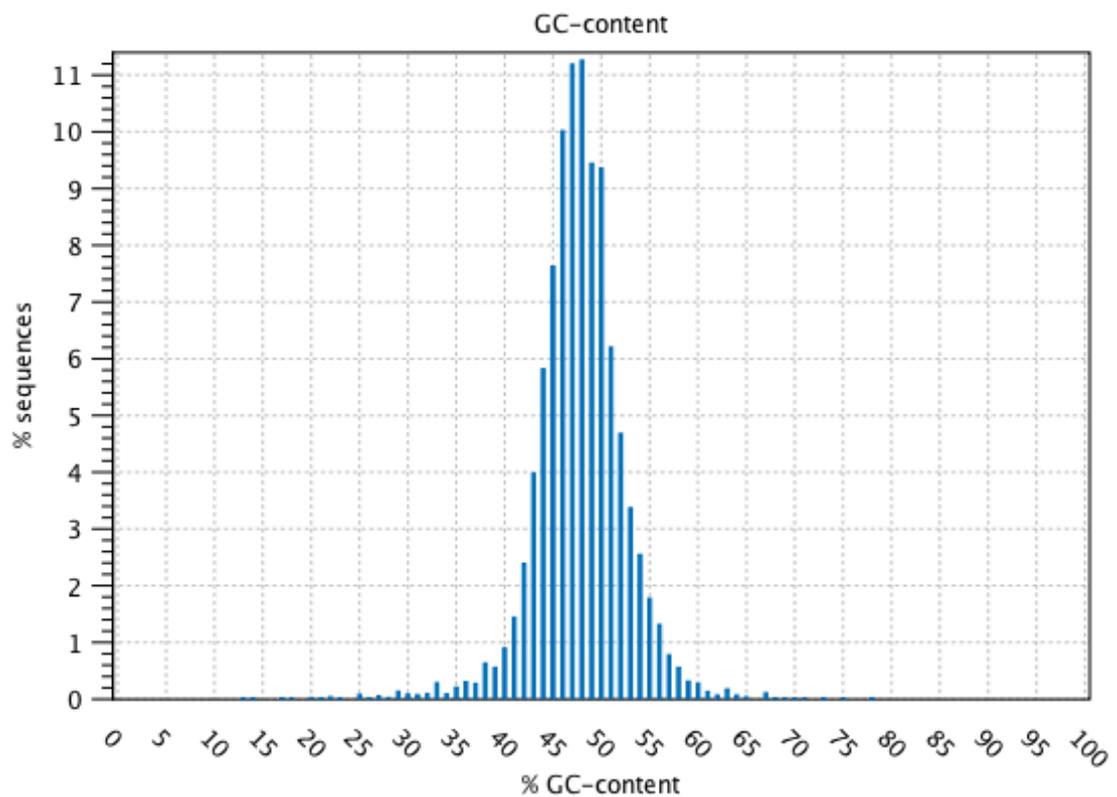


**Figure A3:** *Distribution of trimmed read lengths for the S219 sample (a) successfully mapped or (b) not mapped to the human mitochondrial reference sequence.*



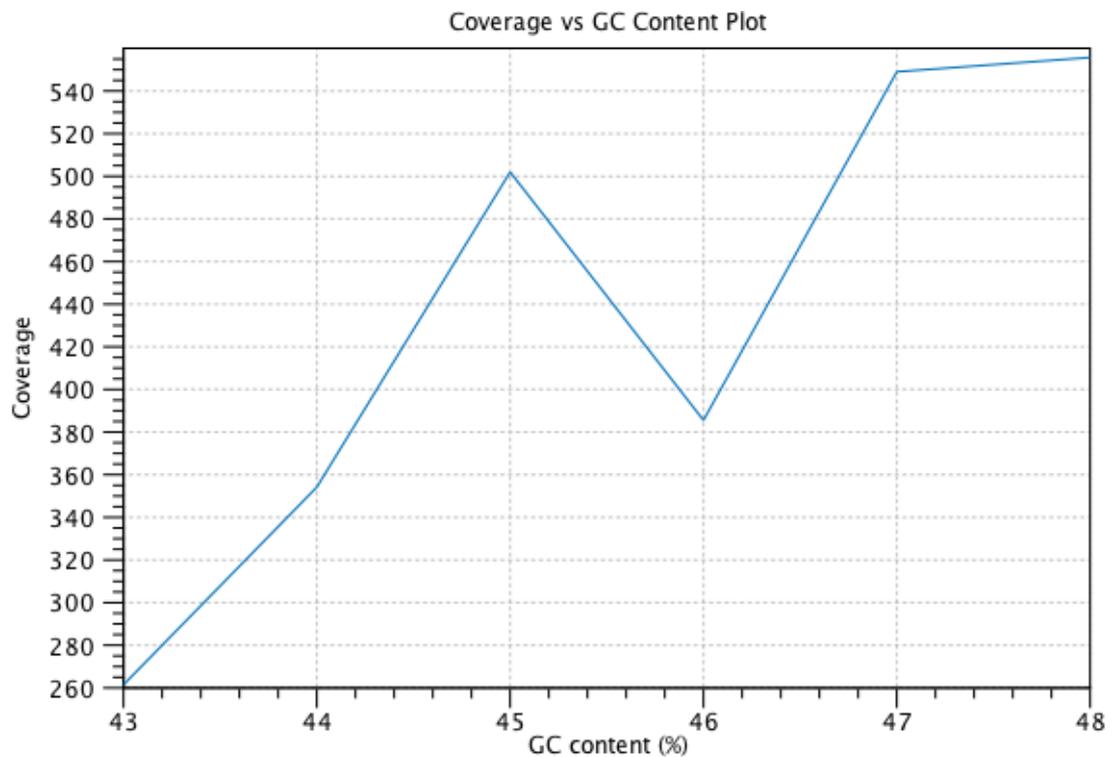
For much of the mtDNA sequence, coverage remains consistent (around 600x) for the S219 sample (**Figure A4a**). However, there were also a few regions of low coverage (approximately 100x) (**Figure A4b**). Examining this particular region further, it was found that the mtDNA region of low coverage spanned approximately 54bp and comprised of a low (33.3%) GC content. Based on this observation, the relationship between GC content and sequence coverage was investigated to determine if there was a correlation between the two variables.

Firstly, the sequence reads were analysed for their GC composition (**Figure A5**). Results revealed that the majority of reads had a GC content of between 45-50%.



**Figure A5:** Examining the relationship between the percentage numbers of sequence reads carrying a particular GC composition (%), using the S219 myeloma data as an example.

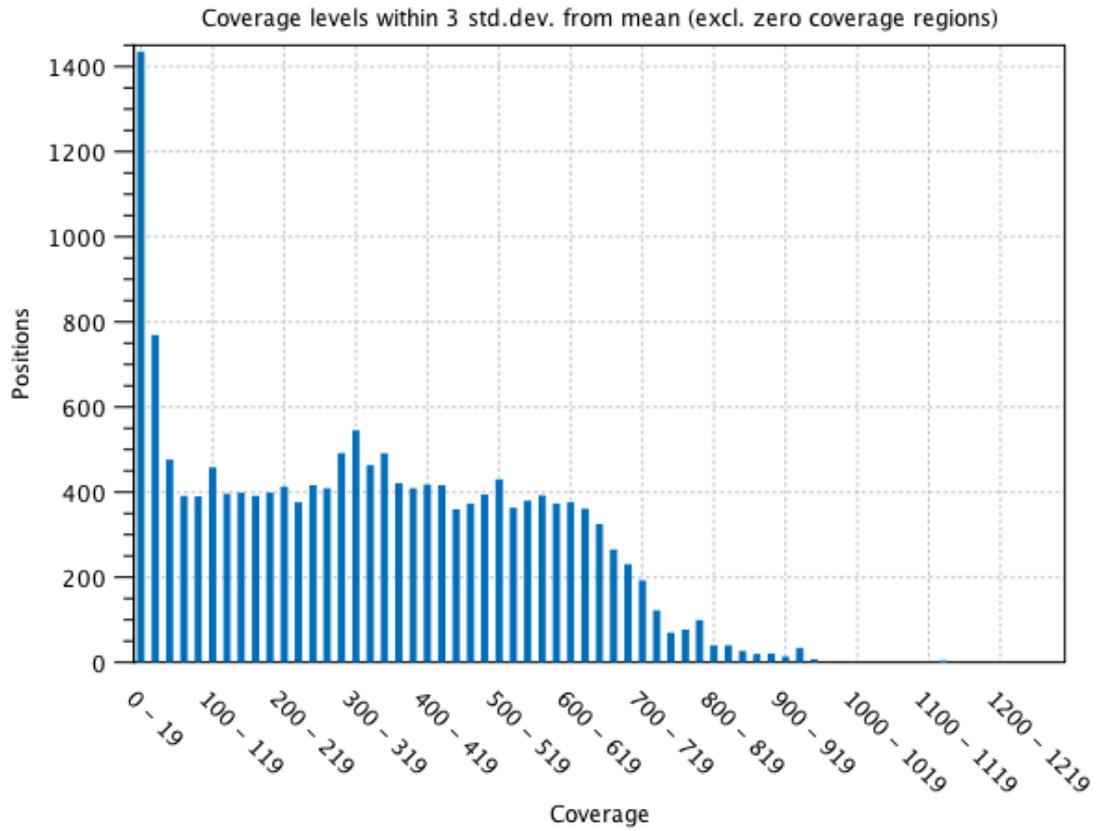
From analysing the influence of GC content to sequence coverage, it was determined that coverage levels were variable especially in sequence regions where the percentage GC content was 46% (**Figure A6**). This corresponded with the ‘peak’ seen from **Figure A5**, which meant that coverage was variable for most of the sequence reads and that there was the absence of a strong correlation between the two factors. This observation was also true for many other Ion Torrent DNA samples.



**Figure A6:** Examining the relationship between sequence coverage and the percentage GC content in the S219 myeloma sample.

Finally, investigating the levels of coverage in the S219 sample revealed that there was no bias in the coverage levels (of up to 650x) for most positions across the entire mtDNA sequence (**Figure A7**). High levels of coverage (over 650x) occurred less frequently across the mtDNA, which meant that for the majority of

nucleotide positions on the mtDNA there is an almost equal chance of the sequence coverage ranging from 60x to 650x.



**Figure A7:** Overall profile for sequence coverage across the mtDNA of the S219 myeloma sample.

## **ii. Chemical composition of laboratory reagents**

Formulations of laboratory chemicals are listed below, where details have been available.

### **Cell culture**

#### **Dulbecco's Modified Eagle Medium (DMEM), high glucose**

##### Amino acids:

0.4mM Glycine  
0.398mM L-Arginine hydrochloride  
0.201mM L-Cystine 2HCl  
4mM L-Glutamine  
0.2mM L-Histidine hydrochloride-H<sub>2</sub>O  
0.802mM L-Isoleucine  
0.802mM L-Leucine  
0.798mM L-lysine hydrochloride  
0.201mM L-Methionine  
0.4mM L-Phenylalanine  
0.4mM L-Serine  
0.798mM L-Threonine  
0.0784mM L-Tryptophan  
0.398mM L-Tyrosine disodium salt dihydrate  
0.803mM L-Valine

##### Vitamins:

0.0286mM Choline chloride  
0.00839mM D-Calcium pantothenate  
0.00907mM Folic Acid  
0.0328mM Niacinamide  
0.0196mM Pyridoxine hydrochloride  
0.00106mM Riboflavin  
0.0119mM Thiamine hydrochloride  
0.04mM i-Inositol

##### Inorganic Salts:

1.8mM Calcium Chloride (CaCl<sub>2</sub>) (anhydrate)  
0.000248mM Ferric Nitrate (Fe(NO<sub>3</sub>)<sub>3</sub>·9H<sub>2</sub>O)  
0.814mM Magnesium Sulfate (MgSO<sub>4</sub>) (anhydrate)  
5.33mM Potassium Chloride (KCl)  
44.05mM Sodium Bicarbonate (NaHCO<sub>3</sub>)  
110.34mM Sodium Chloride (NaCl)  
0.906mM Sodium Phosphate monobasic (NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O)

##### Other Components:

25mM D-Glucose (Dextrose)  
0.0399mM Phenol Red

#### **Dulbecco's Phosphate Buffered Saline (PBS), no calcium, no magnesium**

##### Inorganic Salts:

2.67mM Potassium Chloride (KCl)  
1.47mM Potassium Phosphate monobasic (KH<sub>2</sub>PO<sub>4</sub>)  
137.93mM Sodium Chloride (NaCl)

8.06mM Sodium Phosphate dibasic (Na<sub>2</sub>HPO<sub>4</sub>·7H<sub>2</sub>O)

**RPMI 1640 media**

Amino Acids:

0.133mM Glycine  
1.15mM L-Arginine  
0.379mM L-Asparagine  
0.15mM L-Aspartic Acid  
0.208mM L-Cystine 2HCl  
0.136mM L-Glutamic Acid  
0.0968mM L-Histidine  
0.153mM L-Hydroxyproline  
0.382mM L-Isoleucine  
0.382mM L-leucine  
0.219mM L-Lysine hydrochloride  
0.101mM L-Methionine  
0.0909mM L-Phenylalanine  
0.174mM L-Proline  
0.286mM L-Serine  
0.168mM L-Threonine  
0.0245mM L-Tryptophan  
0.111mM L-Tyrosine disodium salt dihydrate  
0.171mM L-Valine

Vitamins

0.00082mM Biotin  
0.0214mM Choline chloride  
0.000524mM D-Calcium pantothenate  
0.00227mM Folic Acid  
0.0082mM Niacinamide  
0.0073mM Para-Aminobenzoic Acid  
0.00485mM Pyridoxine hydrochloride  
0.000532mM Riboflavin  
0.00297mM Thiamine hydrochloride  
0.0000037mM Vitamin B12  
0.194mM i-Inositol

Inorganic Salts

0.424mM Calcium nitrate (Ca(NO<sub>3</sub>)<sub>2</sub>·4H<sub>2</sub>O)  
0.407mM Madnesium Sulfate (MgSO<sub>4</sub>) (anhyd.)  
5.33mM Potassium Chloride (KCl)  
23.81mM Sodium Bicarbonate (NaHCO<sub>3</sub>)  
103.45mM Sodium Chloride (NaCl)  
5.63mM Sodium Phosphate dibasic (Na<sub>2</sub>HPO<sub>4</sub>) anhydrous

Other Components

11.11mM D-Glucose (Dextrose)  
0.00326mM Glutathione (reduced)  
0.0133mM Phenol Red

**DMEM/F12**

Amino Acids

0.25mM Glycine

0.05mM L-Alanine  
0.699mM L-Arginine hydrochloride  
0.05mM L-Asparagine-H<sub>2</sub>O  
0.05mM L-Aspartic Acid  
0.0998mM L-Cysteine hydrochloride-H<sub>2</sub>O  
0.1mM L-Cystine 2HCl  
0.05mM L-Glutamic Acid  
2.5mM L-Glutamine  
0.15mM L-Histidine hydrochloride-H<sub>2</sub>O  
0.416mM L-Isoleucine  
0.451mM L-Leucine  
0.499mM L-Lysine hydrochloride  
0.116mM L-Methionine  
0.215mM L-Phenylalanine  
0.15mM L-Proline  
0.25mM L-Serine  
0.449mM L-Threonine  
0.0442mM L-Tryptophan  
0.214mM L-Tyrosine disodium salt dihydrate  
0.452mM L-Valine

#### Vitamins

0.0000143mM Biotin  
0.0641mM Choline chloride  
0.0047mM D-Calcium pantothenate  
0.00601mM Folic Acid  
0.0166mM Niacinamide  
0.00971mM Pyridoxine hydrochloride  
0.000582mM Riboflavin  
0.00644mM Thiamine hydrochloride  
0.000502mM Vitamin B12  
0.07mM i-Inositol

#### Inorganic Salts

1.05mM Calcium Chloride (CaCl<sub>2</sub>) (anhyd.)  
0.0000052mM Cupric sulfate (CuSO<sub>4</sub>·5H<sub>2</sub>O)  
0.000124mM Ferric Nitrate (Fe(NO<sub>3</sub>)<sub>3</sub>·9H<sub>2</sub>O)  
0.0015mM Ferric Sulfate (FeSO<sub>4</sub>·7H<sub>2</sub>O)  
0.301mM Magnesium Chloride (anhydrous)  
0.407mM Magnesium Sulfate (MgSO<sub>4</sub>) (anhyd)  
4.16mM Potassium Chloride (KCl)  
29.02mM Sodium Bicarbonate (NaHCO<sub>3</sub>)  
120.61mM Sodium Chloride (NaCl)  
0.5mM Sodium Phosphate dibasic (Na<sub>2</sub>HPO<sub>4</sub>) anhydrous  
0.453mM Sodium Phosphate monobasic (NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O)  
0.0015mM Zinc sulfate (ZnSO<sub>4</sub>·7H<sub>2</sub>O)

#### Other components

17.51mM D-Glucose (Dextrose)  
0.015mM Hypoxanthine Na  
0.00015mM Linoleic Acid  
0.00051mM Lipoic Acid  
0.0215mM Phenol Red

0.000503mM Putrescine 2HCl  
0.5mM Sodium pyruvate  
0.00151mM Thymidine

**TrypLE Express (1X), no phenol red**

Inorganic Salts:

2.67mM Potassium Chloride (KCl)  
1.47mM Potassium Phosphate monobasic (KH<sub>2</sub>PO<sub>4</sub>)  
137.93mM Sodium Chloride (NaCl)  
8.06mM Sodium Phosphate dibasic (Na<sub>2</sub>HPO<sub>4</sub>·7H<sub>2</sub>O)

Other Components:

1.1mM ethylenediamine tetraacetic acid (EDTA)  
(concentration of recombinant protease undisclosed by manufacturer)

**DNA extraction**

**Buffer AE**

10mM Tris-HCl  
0.5mM EDTA  
pH 9.0

**Gel extraction and PCR purification**

**Buffer QG**

5.5M guanidine thiocyanate (GuSCN)  
20mM Tris-HCl, pH 6.6

**Buffer PB**

5.5M guanidine hydrochloride (GuHCl)  
20mM Tris-HCl, pH 6.6

**Buffer PE**

20mM NaCl  
2mM Tris-HCl, pH 7.5  
80% ethanol

**Buffer EB**

10mM Tris-HCl, pH 8.5

**Plasmid extraction**

**Buffer P1**

50mM Tris-HCl, pH 8.0  
10mM EDTA  
50µg/ml RNase A

**Buffer P2**

0.2M NaCl

1% SDS

**Buffer P3**

4M guanidine hydrochloride (Gu HCl)

0.5M potassium acetate, pH 4.2

**Buffer PB**

5.5M guanidine hydrochloride (GuHCl)

20mM Tris-HCl, pH 6.6

**Buffer PE**

As before

**Buffer EB**

10mM Tris-HCl, pH 8.5

**Reverse transcriptase PCR****RiboSafe RNase inhibitor**Storage buffer

20mM HEPES-KOH (pH 7.6)

50mM KCl

8mM DTT

0.1mM EDTA

50% glycerol

stabilizer (details not disclosed)

**Bioscript**Storage buffer

25mM Tris-HCl (pH 7.9)

100mM NaCl

1mM EDTA

5mM DTT

Stabilizer (details not disclosed)

50% glycerol

**PCR****Reaction buffer**160mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>

670mM Tris-HCl, pH 8.8

0.1% Tween-20

## **Gel electrophoresis**

### **6X Loading Dye**

10mM Tris-hydrochloric acid (HCl) (pH 7.6),  
0.03% bromophenol blue,  
0.03% xylene cyanol FF  
60% glycerol  
60mM EDTA

### **50X TAE Buffer**

2M Tris-acetate  
50mM EDTA  
pH 8.0

### **1X TAE Buffer**

1X TAE buffer was made from the stock solution of 50X TAE, diluting 1 part of the 50X TAE buffer with 49 parts of deionized water.

Final concentration of the 1X TAE buffer solution:

40mM Tris acetate  
1mM EDTA

## **Long PCR**

### **Platinum Taq DNA Polymerase High Fidelity Storage Buffer**

20mM Tris-HCl (pH 8.0)  
0.1mM EDTA  
1mM Dithiothreitol (DTT)  
stabilizers (specifics not disclosed by Invitrogen)  
50% (v/v) glycerol

### **10X High Fidelity PCR Buffer**

600mM Tris-SO<sub>4</sub> (pH 8.9)  
180mM Ammonium Sulfate

## **Real time PCR**

### **Sensimix SYBR**

SYBR Green I dye  
dNTPs  
stabilizers  
enhancers (exact details not disclosed)

## **Respirometry (Oxygraph)**

### **10x Hank's Balanced Salt Solution (HBSS)**

#### Inorganic Salts

1.26mM Calcium Chloride (CaCl<sub>2</sub>) (anhydrous)  
0.493mM Magnesium Chloride (MgCl<sub>2</sub>·6H<sub>2</sub>O)  
0.407mM Magnesium Sulfate (MgSO<sub>4</sub>·7H<sub>2</sub>O)  
5.33mM Potassium Chloride (KCl)  
0.441mM Potassium Phosphate monobasic (KH<sub>2</sub>PO<sub>4</sub>)  
4.17mM Sodium Bicarbonate (NaHCO<sub>3</sub>)  
137.93mM Sodium Chloride (NaCl)  
0.338mM Sodium Phosphate dibasic (Na<sub>2</sub>HPO<sub>4</sub>) anhydrous

#### Other components

5.56mM D-Glucose (Dextrose)

## **DNA cloning**

### **pCR2.1, linearized**

25ng/μL in 10mM Tris-HCl

1mM EDTA, pH 8

### **10X Ligation Buffer**

60mM Tris-HCl, pH 7.5

60mM MgCl<sub>2</sub>

50mM NaCl

1mg/ml bovine serum albumin

70mM beta (β)-mercaptoethanol

1mM ATP

20mM dithiothreitol

10mM spermidine

### **S.O.C Medium**

2% Tryptone

0.5% Yeast Extract

10mM NaCl

2.5mM KCl

10mM MgCl<sub>2</sub>

10mM MgSO<sub>4</sub>

20mM glucose (dextrose)

### **5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-GAL) solution**

40mg/ml X-GAL dissolved in N, N-dimethylformamide (DMF)

### **Luria-Bertani (LB) agar plates**

1.0% Tryptone

0.5% Yeast Extract

1.0% NaCl

15g/L Agarose (pH. 7.0)

100μg/ml ampicillin

**Luria-Bertani Broth**

1.0% Tryptone  
0.5% Yeast Extract  
1.0% NaCl  
100µg/ml ampicillin, pH 7.0

**Capillary Sequencing****5x BDT buffer**

400mM Tris  
10mM MgCl<sub>2</sub>, pH 9.0

**Next Generation Sequencing****E-Gel**

0.1 to 0.3µg/ml ethidium bromide  
0.055% proclin (preservative)

**Platinum PCR SuperMix High Fidelity**

22 U/ml complexed recombinant Taq DNA polymerase  
Pyrococcus species GB-D thermostable polymerase  
Platinum Taq Antibody  
66mM Tris-SO<sub>4</sub> (pH 8.9)  
19.8mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>  
2.4mM MgSO<sub>4</sub>  
220µM dNTPs  
stabilizers

**Melt-Off Solution**

125mM NaOH  
0.1% Tween 20

**1x TEX**

10mM Tris  
1mM EDTA  
0.01% Triton X-100, pH 8

**1x SSPE**

150mM Sodium Chloride  
10mM Sodium Phosphate  
1mM Na<sub>2</sub>EDTA

**Buffer TE**

10mM Tris-HCl  
1mM EDTA.Na<sub>2</sub>  
pH 8.0

**Low TE Buffer**

10mM Tris-HCl

0.1mM EDTA

pH 8.0

**Dynabeads MyOne Streptavidin C1 beads**

Dynabeads are supplied in PBS, pH 7.4

0.01% Tween-20

0.02% NaN<sub>2</sub> (preservative)

stored at 4°C

**Western blot****4X Loading Dye**

200mM Tris-Cl pH 6.8

0.4M DTT

8% SDS

40% glycerol

stored in -20C

**1X Transfer Buffer**

20% methanol

47.8mM Tris

38.6mM Glycine

0.03% SDS

**1X SDS Running Buffer**

24.7mM Tris

58.6mM Glycine

0.2% SDS

### **iii. Publications**

1. **Yeung K.Y.**, Dickinson A., Donoghue J., Polekhina G., Grammatopoulos D.K., McKenzie M., Johns T.G., St John J.C. The Identification of Mitochondrial DNA Variants in Glioblastoma Multiforme. 2014; Jan 2;2(1):1. (PMID: 24383468)
2. **Yeung K.**, Dickinson A., St. John J. Chapter 6: The role of mitochondrial DNA in tumour stem cells (2013). In *Mitochondrial DNA, Disease and Stem Cells*. Ed St. John J. Humana Press. ISBN 978-1-62703-101-1. p1-36.
3. Dickinson A., **Yeung K.Y.**, Donoghue J., Kelly R.D.W., McKenzie M., Johns T.G., St John J.C. The regulation of mitochondrial DNA copy number in gliomaspheres. *Cell Death Differ.* 2013; Dec; 20(12):1644-53. (PMID: 23995230)
4. **Yeung K.Y.**, Khong T., McKenzie M., Spencer A., St John J.C., Mitochondrial DNA is essential for the proliferation and control of tumorigenesis in U266 Multiple Myeloma cells (Submitted)