Investigating the role of mitochondria at the preimplantation stages of human embryonic development

by

Razan S. Jawdat

A thesis submitted for the degree of Doctor of Philosophy at University College London

April 2018

Institute for Women’s Health University College London
I, Razan S. Jawdat, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.
Acknowledgements

It is my pleasure to thank those who made this project possible and helped me through my PhD. First of all, I would like to thank my primary supervisor, Dr. Sioban SenGupta. Sioban, you have always been there for me and guided me with patience throughout my PhD. You believed in me, listened to me and supported me to pursue an area of research that was of great interest to me. Seeing you at Chenies Mews every day made me feel at home since you treated me like a daughter. It was a great pleasure to be your PhD student. Thank you very much for making me a better scientist. Without your help, guidance and continuous support, I would not have been at this stage now.

I would like to thank my secondary supervisor, Professor Nick Lane. Professor Lane, I have always been a fan of your books, and I feel delighted that I got the chance to do this project under your guidance. Your passion in this field made me grow a great interest in mitochondria. I would like to thank you for your warm welcome to this area and all your support and guidance throughout my PhD. I would also like to thank Professor Joy Delhanty for updating me with the latest papers in the field. It is always nice to feel that someone remembers me when they read an article related to my research.

I would like to thank everyone who worked in the PGD team when I started my PhD; Roy Pascal Naja, Seema Dhanjal, Harita Gheveria, Pinar Tulay, Aisha Elaimi Thalia Mamas and Leoni Xanthapoulou. Thank you all for showing me some of the lab techniques when I joined this lab. Roy, thank you for helping with the real-time PCR analysis. I have learned so much from you and Seema, and you will always be my friends. Seema and Harita, I am happy that I met you and even though I had
more lab space and PCR machines to use when the clinical PGD has closed down, I missed you. Pinar, I was lucky to meet you at CM in your last year, you have been welcoming, and I'm glad we became friends. Leoni, thank you very much for always being in touch and for providing me with the repeated miscarriage group DNA samples. I am pleased that I got the chance to meet when you were writing up. Special thanks to Helen O’Neill and Xuhui Sun, even though I felt lonely when everyone left CM you made it a beautiful place, and I wish you all the best in your future there. Helen, you will always be a friend to remember, I enjoyed your company in the office especially when you speak to me in Arabic. I will really miss you. Sun, even though you are so quiet, I enjoyed your company. Your help with some of the statistical analysis was beneficial, and I wish your RNA expression experiments go well.

There are many other colleagues at UCL who I cannot thank enough and helped me get to this stage: Tony Brooks, Kerra Pearce, Dr Mark Kristiansen at ICH and Will Kotiadis. Tony, you were always there when I needed your help with the MiSeq. You are brilliant, and I am glad to have learned a lot from you. Kerra, I have known you from my master's, and you never failed me. I am so happy things worked out for you at the end at ICH. Special thanks to Mark for loading the samples for genotyping. Will for having an interest in my project and reading some parts of it. I was lucky to have some charming MSc student to do parts of this project. All my MSc students Ayman Haj Ali, Zosia Litwinska, Ellen Dewhurst, Caroline Picard and Okan Atilan.

I would also like to thank Professor Walther Parson and Dr Christopher Philips for guiding me in analysing of the mitochondrial haplotypes and AIM respectively. I had help in the analysis in some parts of this project from Dr Simon Andrews at the
Babraham for the initial data clean-up, Dr Yoan Diekmann for assistance in the analysis of nuclear SNPs data, Aviva Petrie for guidance in the complicated statistical analysis.

Special thanks to the CRGH: I would like to thank, Dr Paul Serhal, Kalliopi Loutradi, Rabi Odia for thawing the verified embryonic samples and providing me with them. I would also like to thank Dr Serdar Coskun and Dr Wafa Qubbaj for supporting my scholarship from KFSH&RC. I hope to come back and help set up the mtDNA analysis when I am again at KFSH&RC.

Finally, I would like to thank my family for always believing in my work and for all the support and encouragement they provided. Dad, you supported me throughout my studies and still believed in me and my siblings. I hope your investment in me paid off and I at times I felt like giving up, but I kept going to make you proud as I promised. Mom, you are amazing and thank you for being there for me. I would also like to thank my friends; Souraya, Amna, Lena, Charalambos, Martina, Haifa, Bana, Njoud, Zia Antonia, May, Sarah and Omar thank you all for being my family in London and always there for me whenever I needed you. You gave me beautiful memories in London that I will never forget.
To my family
& everyone who believed in me
Abstract

Mitochondria are the major energy producers in cells in the form of ATP. Proteins required for mitochondrial function are encoded by both mitochondrial (mtDNA) and nuclear DNA (nDNA) necessitating compatibility between the two genomes. Good quality oocytes containing an optimal number of mitochondria and sufficient levels of ATP produce higher quality blastocysts. Recent data suggests an elevated level of mtDNA may be associated with aneuploidy in blastocysts.

In this study, preimplantation embryo development was investigated in relation to mtDNA template number and aneuploidy. ATP levels were measured in blastocysts and linked to aneuploidy. To investigate possible nDNA/mtDNA mismatch, DNA from couples with repeated miscarriage or repeated implantation failure (RM/RIF) was compared with DNA from couples with no history of infertility but who had preimplantation genetic diagnosis (PGD) for monogenic disorders. DNA from both groups was genotyped using SNP arrays. Sequencing of the mitochondrial genome and a set of 53 nuclear-encoded genes important in mitochondrial function was also performed. Further sequencing of the selected nuclear genes in embryos from the PGD group was used to identify variants in the nuclear and mitochondrial genomes associated with poor embryo development.

Our data showed that arrested embryos that were euploid had more mtDNA than arrested embryos that were aneuploid. Aneuploidy in blastocysts resulted in variability in ATP levels. When aneuploidy was present in more than ten chromosomes, ATP was almost undetectable.

From the sequencing analysis, significantly more couples with RM/RIF (5/12) had partners from the same mtDNA haplogroup compared with the PGD group of
couples (1/11). Yet, SNP data analysed by identity by state (IBS) showed no significant differences between partners in couples based on their nDNA even when the HLA region was considered separately.

Within the PGD group, the presence of the T haplogroup in the male partner was associated with a smaller percentage of embryos developing to blastocysts. Analysis of embryos from these couples suggested a link with SNPs in nuclear genes (specifically COQ9 and PPARGC1a) encoding mitochondrial proteins which may contribute to poor embryo development due to a disturbance in the electron transport chain or mitochondrial biogenesis respectively.

Determining the mitochondrial haplogroups of both parents is a useful tool to investigate potential mismatch between the nuclear and mitochondrial genomes in embryos which may influence their development.
# Table of Contents

Acknowledgements ............................................................................................................ 3

Abstract ................................................................................................................................. 7

List of Tables .......................................................................................................................... 15

Chapter 1  Introduction ........................................................................................................ 22
1.1 Fertilisation ...................................................................................................................... 23
1.2 Mitochondria .................................................................................................................... 25
  1.2.1 Function of Mitochondria in the eukaryotic cell ......................................................... 27
  1.2.2 The Mitochondrial Genetic System ............................................................................ 32
  1.2.3 Mito-nuclear interaction ............................................................................................ 40
  1.2.4 Mitochondrial disorders ............................................................................................ 47
  1.2.5 Mitochondria in early mammalian development ....................................................... 50
  1.2.6 Fertility and mitochondrial DNA ................................................................................. 53
1.3 Repeated miscarriage and repeated implantation failure ............................................... 70
  1.3.1 Genetic factors ........................................................................................................... 70
  1.3.2 Major histocompatibility complex (MHC) and the human leucocyte antigen (HLA) .... 72
1.4 Association studies ......................................................................................................... 73
  1.4.1 Genome Wide Association Studies (GWAS) ............................................................... 74
  1.4.2 SNP-genotyping technology ....................................................................................... 76
1.5 Hypothesis ....................................................................................................................... 78
1.6 Aims ................................................................................................................................. 78
  1.6.1 Aim 1: To assess the effect of mtDNA copy number and ATP with preimplantation embryo quality by: .............................................................. 78
  1.6.2 Aim 2: To investigate the effect of parental ethnicity on parental fertility and embryo quality 79
  1.6.3 Aim 3: To examine the effect of mtDNA/nDNA mismatch on human embryos quality and fertility 79

Chapter 2  Materials & Methods ....................................................................................... 81
Examine the effect of mtDNA/nDNA mismatch on human embryos quality and fertility .... 84
2.1 General laboratory practice ............................................................................................. 85
2.2 Sample collection and patient information ................................................................... 86
  2.2.1 Human embryos from PGD and PGS cycles ............................................................. 86
  2.2.2 Genomic samples from fertile and infertile patients .................................................. 88
2.3 Sample preparation ........................................................................................................ 89
  2.3.1 Single cell isolation and embryo tubing ..................................................................... 89
  2.3.2 Lymphocyte separation using the Ficoll-Paque plus method ...................................... 90
  2.3.3 Single cell lysis .......................................................................................................... 91
  2.3.4 Whole genome amplification ..................................................................................... 93
2.4 Sample processing and analysis .................................................................................... 94
  2.4.1 General methods ....................................................................................................... 94

Chapter 3  Assessment of mtDNA copy number and ATP with preimplantation embryo quality 107
3.1 Aims Summary ................................................................................................................. 107
3.2 Materials and methods ................................................................................................. 108
  3.2.1 Assessment of mitochondrial DNA copy number by qPCR ...................................... 110
  3.2.2 Assessment of ATP content in human preimplantation embryos ............................. 114
3.3 Results ............................................................................................................................ 116
3.3.1 Quantification of mtDNA by Resolight and TaqMan assays ......................................... 117
3.3.2 Relative quantification by TaqMan copy number assays ........................................... 130
3.3.3 Aneuploidy detection by GenetiSure α-CGH .............................................................. 131
3.3.4 Analysis of ATP levels with aneuploidy in embryos .................................................. 136
3.4 Discussion ...................................................................................................................... 137
3.4.1 Absolute quantification of mtDNA copy number ....................................................... 138
3.4.2 Relative Quantification of mtDNA by TaqMan copy number assays ......................... 141
3.4.3 ATP ........................................................................................................................ 142
3.5 Results summary ......................................................................................................... 144
3.6 Limitations and future work ......................................................................................... 145

Chapter 4  The effect of parental ethnicity on fertility and embryo quality ......................... 147
4.1 Aims summary ............................................................................................................... 147
4.2 Materials and methods ............................................................................................... 149
4.2.1 Ethnicity determined by mtDNA haplogroup ........................................................... 151
4.2.2 Ethnicity determined by nuclear genes .................................................................... 156
4.2.3 Identity by State (IBS) ............................................................................................. 160
4.2.4 Aneuploidy comparison ............................................................................................ 161
4.3 Results .......................................................................................................................... 162
4.3.1 Ethnicity determined by the mitochondrial genome ............................................... 162
4.3.2 The effect of mtDNA haplogroup T and embryo mtDNA number ......................... 173
4.3.3 Ethnicity determined by nuclear genes .................................................................... 174
4.3.4 Nuclear genome template difference between individuals within the fertile and the RM/RIF couples ........................................................................................................... 178
4.3.5 Differences between partners in couples in nuclear encoded mitochondrial SNPs and the MHC region .................................................................................................... 180
4.3.6 Karyomapping analysis with BlueFuse ..................................................................... 181
4.3.7 Comparison between HumanKaryomap SNP chip & α-CGH for aneuploidy detection .... 183
4.4 Discussion ...................................................................................................................... 188
4.4.1 Couples within the fertile group were more likely to be distantly related compared to the infertile group .................................................................................................... 189
4.4.2 Mitochondrial haplogroup T is associated with poor blastocyst rate ....................... 190
4.4.3 Nuclear genotype differences between partners (IBS) had no significant effect on fertility 195
4.4.4 Ancestry genotype differences between partners (IBS) had no significant effect on fertility ................................................................. 196
4.4.5 Immunological factors had no effect on fertility outcome .................................... 197
4.5 Results summary ......................................................................................................... 198
4.6 Limitations and future work ......................................................................................... 199

Chapter 5  Incompatibilities between the mitochondrial and nuclear genomes that affect preimplantation embryo development .......................................................... 200
5.1 Aims summary ............................................................................................................... 200
5.2 Materials & Methods .................................................................................................. 201
5.2.1 Samples .................................................................................................................. 201
5.2.2 Nuclear encoded mitochondrial genes selection and assay design ......................... 202
5.2.3 Nuclear genome sequencing and haplotyping ......................................................... 205
5.3 Results .......................................................................................................................... 209
5.3.1 Candidate genes selection for mitochondrial-nuclear mismatch investigation .......... 209
5.3.2 Targeted next generation sequencing ..................................................................... 211
5.3.3 Analysis of variants compared between blastocysts & arrested or degenerated embryos 215
5.3.4 Analysis of SNPs and indels between the fertile and RM group ........................................... 224
5.3.5 mtDNA haplogroup specific comparison based on nuclear encoded mitochondrial genes associated SNPs in genomics ........................................................................................................ 225
5.4 Discussion .................................................................................................................................. 226
  5.4.1 Mitochondrial-nuclear mismatch investigation in embryos ............................................... 226
  5.4.2 Selection of candidate nuclear genes .................................................................................. 227
  5.4.3 Nuclear DNA variations associated with poor embryo development ............................... 231
  5.4.4 How do those SNPs interact with the mitochondrial genome? ........................................ 233
5.5 Results summary ....................................................................................................................... 237
5.6 Limitations and future work ..................................................................................................... 238

Chapter 6 General discussion ........................................................................................................ 239
  6.1 Conclusion ............................................................................................................................... 249

Chapter 7 Appendix ...................................................................................................................... 250
  7.1 Appendix A ............................................................................................................................. 250
    7.1.1 Validation of GenetiSure kit for aneuploidy screening .................................................. 250
  7.2 Appendix B ............................................................................................................................. 254
    7.2.1 Calculation used to dilute the standards (oligonucleotides) used in this study ............. 254
    7.2.2 Calculation used to dilute the ATP standards ............................................................... 255
    7.2.3 Embryonic aneuploidy status, morphology & mitochondrial haplogroups .................. 256
    7.2.4 Validation of mtDNA template quantification technique on small number of cells: ...... 259
    7.2.5 Validation of standard curves used for mtDNA template quantification: ..................... 260
  7.3 Appendix C ............................................................................................................................. 262
    7.3.1 Identification of mitochondrial haplogroups by EMPOP ............................................. 262
    7.3.2 List of AIM available on the Karyomap SNP chip ......................................................... 263
    7.3.3 Script used in the IBS analysis ...................................................................................... 264
    7.3.4 The probe IDs of the identified mitochondrial SNPs showing differences between fertile and RM/RIF groups analysed by the CoreExome SNP chip ....................................................... 265
  7.4 Appendix D ............................................................................................................................. 266
    7.4.1 Script to filter out valid SNPs for the analysis .............................................................. 266

Chapter 8 Publications ..................................................................................................................... 281
  8.1 Published abstracts .................................................................................................................. 281
  8.2 Manuscripts in preparation ..................................................................................................... 282

Chapter 9 References ..................................................................................................................... 283
  9.1 Websites ............................................................................................................................... 283
  9.2 Books .................................................................................................................................... 283
  9.3 Papers .................................................................................................................................... 284
List of figures

Figure 1–1: Diagram illustrating the main steps of preimplantation human embryo development........24
Figure 1–2: The mitochondrion inner structure. .................................................................26
Figure 1–3: The mitochondrial electron transport chain ................................................28
Figure 1–4: ATP generation processes inside the mitochondria and the production of ROS.........29
Figure 1–5: Mitochondria and apoptosis ........................................................................31
Figure 1–6: Mitochondrial haplogroups and major human population migration ................39
Figure 1–7: Scheme summarizing the various modes of mito-nuclear interactions ...............41
Figure 1–8: Co-adaptation is required for functional respiration ......................................42
Figure 1–9: Free-radicals signal mitochondrial biogenesis, triggering apoptosis or increasing capacity. .................................................................43
Figure 1–10: Human mitochondrial genome. .................................................................45
Figure 1–11: Graphical representation of the mitochondrial and nuclear encoded mitochondrial proteins involved in the OXPHOS system ................................................48
Figure 1–12: Proposed mechanisms for the mtDNA genetic bottleneck .........................53
Figure 1–13: Schematic diagram of the effect of oxidative stress on fertility .....................61
Figure 1–14: Graphical representation of an indirect association between a disease phenotype and marker locus .............................................................................75
Figure 2–1: The work flow of the project to achieve the project aims ................................83
Figure 2–2: Diagram illustrating the embryonic disaggregation scheme on a petri-dish .......90
Figure 2–3: Lymphocyte separation using the Ficoll-Paque plus method .......................91
Figure 2–4: Multiple displacement amplification .........................................................93
Figure 2–5: Example of sample electropherogram showing pre-capture analysis of amplified DNA library using the Bioanalyzer and a DNA 1000 Assay ...........................................98
Figure 2–6: Schematic diagram of embryo analysis for chromosomal imbalances using a-CGH. ....101
Figure 3–1: The work flow for assessing mtDNA template number and ATP levels in embryonic samples ..................................................................................109
Figure 3–2: The synthetic oligonucleotide used for quantification of mtDNA .................110
Figure 3–3: Serial dilution of synthetic oligonucleotide ...............................................111
Figure 3–4: Diagram illustrating the principle for ATP measurement in this procedure .......115
Figure 3–5: Absolute quantification of the normal oligonucleotide dilutions by Resolight and HRM.117
Figure 3–6: Standard curve of the normal oligonucleotide dilution series by Resolight and HRM ....118
Figure 3–7: Absolute quantification of normal oligonucleotide after performing a 7 cycle PCR ......120
Figure 3–8: HRM of the absolute quantification of normal oligonucleotides after 7 cycle PCR ......120
Figure 3–9: Gel electrophoresis confirming the primer-dimer formation when 2 sets of PCR reactions were performed ................................................................................121
Figure 3–10: Quantification of embryonic samples lysed by PK showing differences in amplification based on cell numbers ........................................................................122
Figure 3–11: Absolute quantification of mtDNA in embryonic samples lysed by ALB........................................123
Figure 3–12: mtDNA copy number comparison between embryos based on morphology.................................130
Figure 3–13: Relative quantification comparison between embryos based on maternal age.............................131
Figure 3–14: Array-CGH traces from cytoreports, generated by the CytoGenomics Software. An example of a euploid and aneuploid embryonic samples.............................................................132
Figure 3–15: Array-CGH traces from cytoreports, generated by the CytoGenomics Software. An example of aneuploid ≤5 and multiple aneuploid >5 embryonic samples. .........................................................133
Figure 3–16: Comparison between euploid and aneuploid embryos based on mtDNA copy number. 
.........................................................................................................................................................134
Figure 3–17: Comparison between euploid and aneuploid embryos based on mtDNA copy number excluding embryos with losses and gains on RNAase P.................................................................135
Figure 3–18: Comparison of mtDNA (normalised to RNAse P) template number between euploid and aneuploid arrested and blastocysts embryonic samples.........................................................135
Figure 3–19: Comparison of mtDNA (normalised to RNAse P) template number between euploid and aneuploid arrested and blastocysts embryonic samples.........................................................136
Figure 3–20: Correlation of ATP levels with aneuploidy in embryos.................................................................137
Figure 4–1: An illustration of the workflow applied for sequencing the mitochondrial genome of parental genomic samples as well as some of their embryos.................................................................150
Figure 4–2: The five overlapping amplicons used for LR-PCR amplification of the mtDNA genome.152
Figure 4–3: A test data set example of the triangular matrix of allele differences. ..............................................161
Figure 4–4: Gel electrophoresis (0.8%) showing the size of the PCR products following mtDNA sequencing. .............................................................................................................................................162
Figure 4–5: Electropherogram of the LR-MITO three sets of PCR Amplicon on an embryonic sample. 
.........................................................................................................................................................163
Figure 4–6: Example of Electropherogram of the MTL1-MTL2 two sets of PCR Amplicon on genomic and an embryonic sample.........................................................................................................164
Figure 4–7: The QScore distribution for the second mitochondrial genome sequencing run. .................165
Figure 4–8: Haplotype distance comparison between the fertile and RM/RIF group.........................................168
Figure 4–9: Mitochondrial haplogroup effect on fertilisation and blastocyst rate within the fertile group. ...............................................................................................................................................170
Figure 4–10: The difference in depth of read from 5 different embryonic samples as shown on IGV. 
.........................................................................................................................................................173
Figure 4–11: Principle component analysis (PCA) clustering for cases(RM/RIF) and controls (PGD) based on 7 populations using 229 SNPs.................................................................175
Figure 4–12: A comparison of IBS between the fertile (PGD) and the infertile(RM/RIF) group of patients against all the SNPs present on the chip.................................................................179
Figure 4–13: Mitochondrial SNPs differences between couples within the fertile and infertile groups. 
.........................................................................................................................................................180
Figure 4–14: IBS difference comparison between the PGD (fertile) and the infertile (RM/RIF) groups based on the nuclear encoded mitochondrial selected genes and MHC SNPs. ..........................181

Figure 4–15: Comparison between (A) Karyomapping and (B) Array CGH by GenetiSure. ....................187

Figure 5–1: An illustration of the workflow applied for sequencing the targeted nuclear encoded mitochondrial genes of parental DNA samples as well as their embryos...........................................204

Figure 5–2: Target Enrichment Analysis pipeline ......................................................................................208

Figure 5–3: A graphic representation of the mitochondrial and nuclear encoded mitochondrial proteins including the genes that were focused on. ........................................................................................................209

Figure 5–4: The selected nuclear genes for mitochondrial-nuclear mismatch investigation......................210

Figure 5–5: The pre-capture analysis of amplified DNA library amplicons using the Bioanalyser ..........212

Figure 5–6: The pre-capture analysis of amplified DNA library amplicons using the TapeStation. .........213

Figure 5–7: Post capture analysis of amplified indexed library DNA using the TapeStation. ...............214

Figure 5–8: Analysis of SNPs associated with embryos arrested at the cleavage stage in Family 3 and Family 4. .................................................................................................................................218

Figure 5–9: Pedigree showing the analysis of SNPs associated with the arrested embryos in Family-9..............................................................................................................................................220

Figure 5–10: The associated SNPs with family 10 as displayed on IGV .....................................................221

Figure 5–11: Nuclear SNPs comparisons between male carriers for mitochondrial haplogroup T with carriers for M mitochondrial haplogroups. ................................................................................225

Figure 5–12: Diagram explaining the Hydrophobicity and Threonine content effect on the energy release. ........................................................................................................................................228

Figure 5–13: A proposed schematic model for the interaction network of NDUFC2, NDUFA1, and mtDNA-encoded complex I subunits ..................................................................................230

Figure 5–14: Schematic diagram for mtDNA–nDNA mismatch theory..................................................236

Figure 7–1: Standard curve used in mtDNA template analysis for samples lysed by PK .........................259

Figure 7–2: Standard curve for testing the distribution accuracy of the standards dilution used for mtDNA template analysis ........................................................................................................261

Figure 7–3: Mitochondrial haplogroups assignment using EMPOP. ........................................................262
List of Tables

Table 1-1: Comparison between the human mitochondrial and nuclear genomes. ..................................................34
Table 2-1: Techniques used for sample collection, preparation, processing and analysis to fulfil all aims of this thesis. .........................................................................................................................84
Table 2-2: Laboratory activities were carried out in specific designated areas......................................................85
Table 2-3: Components of hybridisation master mix ..........................................................................................103
Table 2-4 Washing conditions following a-CGH hybridisation .......................................................................104
Table 2-5: QC metric thresholds used by Agilient for experimental samples analysed by Genetisure. ..............106
Table 3-1: Samples used for mtDNA quantification by Resolight and TaqMan assays & ATP assessment: .................................................................................................................................116
Table 3-2: Quantification of mtDNA template of the synthetic oligonucleotide by a single round of PCR. .................................................................................................................................118
Table 3-3: Absolute quantification of normal oligonucleotide after performing a 7 cycle PCR. ......121
Table 3-4: Absolute quantification of mtDNA in single cells, 2 cells and clump of embryonic cells. ...127
Table 3-5: Absolute quantification of mtDNA in whole embryos. .................................................................129
Table 3-6: The number of euploid and aneuploid embryos. ..............................................................................132
Table 4-1: Long Range PCR primers for mtDNA amplification .......................................................................152
Table 4-2: Results of all runs used for mitochondrial genome sequencing. .........................................................164
Table 4-3: comparison of mtDNA haplotypes between the fertile and the RM/RIF group. .........................167
Table 4-4: The maternal and paternal mitochondrial haplotypes with fertilisation and blastocyst formation for fertile couples......................................................................................................................169
Table 4-5: Correlation between blastocyst rate and mtDNA haplotype .............................................................170
Table 4-6: The SNPs making up the mitochondrial haplogroup T........................................................................171
Table 4-7: Ethnicities determined by nuclear genes compared to the mitochondrial haplogroups. ...176
Table 4-8: Mitochondrial SNPs that are different between the fertile and RM group. .................................180
Table 4-9: Karyomapping SNP call rates. ...............................................................................................................182
Table 4-10: Comparison between HumanKaryomap SNP chip and a-CGH. ......................................................184
Table 5-1: The characteristics of the sequencing assay designed by Agilent .....................................................203
Table 5-2: Details of all the runs performed for sequencing the nuclear encoded genes ...............................215
Table 5-3: Summary of the genes associated with poor embryo development in Family-3. ..........................216
Table 5-4: Summary of the genes associated with poor embryo development in Family-9. .......................219
Table 5-5: Summary of the genes associated with poor embryo development in Family-10. ..........................221
Table 5-6: The indels associated with the poorly developing embryos ..........................................................223
Table 5-7: Comparison of variants between the fertile and RM/RIF group......................................................224
Table 7-1: A comparison between BlueGnome and Genetisure array results. .................................................251
Table 7-2: Embryo results using different techniques (FISH, array-CGH and karyomapping) .................253
Table 7-3: Embryonic aneuploidy status, morphology & mitochondrial haplogroups. ...............................256
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>-dF/Dt</td>
<td>Negative derivative of the fluorescence with respect to temperature</td>
</tr>
<tr>
<td>^φm</td>
<td>Mitochondrial membrane potential (MMP)</td>
</tr>
<tr>
<td>°C</td>
<td>Degree Celsius</td>
</tr>
<tr>
<td>△Cq</td>
<td>Difference in cycle of quantification</td>
</tr>
<tr>
<td>1-IBS</td>
<td>Genetic difference</td>
</tr>
<tr>
<td>A</td>
<td>Adenine</td>
</tr>
<tr>
<td>a-CGH</td>
<td>Array comparative genomic hybridization</td>
</tr>
<tr>
<td>ABF</td>
<td>Abnormally fertilized embryo</td>
</tr>
<tr>
<td>aCGH</td>
<td>Array comparative genomic hybridization</td>
</tr>
<tr>
<td>ADO</td>
<td>Allele drop out</td>
</tr>
<tr>
<td>ADP</td>
<td>Adenosine Diphosphate</td>
</tr>
<tr>
<td>AIM</td>
<td>Ancestry Informative Marker</td>
</tr>
<tr>
<td>ALB</td>
<td>Alkaline Lysis Buffer</td>
</tr>
<tr>
<td>ART</td>
<td>Assisted Reproductive Technology</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine Triphosphate</td>
</tr>
<tr>
<td>ATPase6</td>
<td>Adenosine triphosphate synthase subunit 6</td>
</tr>
<tr>
<td>ATPase8</td>
<td>Adenosine triphosphate synthase subunit 8</td>
</tr>
<tr>
<td>B</td>
<td>Blastomere</td>
</tr>
<tr>
<td>BAF</td>
<td>B-Allele Frequency</td>
</tr>
<tr>
<td>BAM</td>
<td>Binary Alignment/Map</td>
</tr>
<tr>
<td>Bc1-2</td>
<td>B-cell CLL/lymphoma 2</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic local alignment search tool</td>
</tr>
<tr>
<td>bp</td>
<td>base-pair(s)</td>
</tr>
<tr>
<td>BrdU</td>
<td>Bromodeoxyuridine</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>BWA</td>
<td>Burrows-Wheeler Aligner</td>
</tr>
<tr>
<td>C</td>
<td>Cytosine</td>
</tr>
<tr>
<td>CA^2+</td>
<td>Calcium ion</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary deoxyribonucleic acid</td>
</tr>
<tr>
<td>chr</td>
<td>Chromosome</td>
</tr>
<tr>
<td>Chr</td>
<td>Chromosome</td>
</tr>
<tr>
<td>CI</td>
<td>Complex I</td>
</tr>
<tr>
<td>CII</td>
<td>Complex II</td>
</tr>
<tr>
<td>CIII</td>
<td>Complex III</td>
</tr>
<tr>
<td>CIV</td>
<td>Complex IV</td>
</tr>
<tr>
<td>CO₂</td>
<td>Carbon dioxide</td>
</tr>
<tr>
<td>conc.</td>
<td>Concentration</td>
</tr>
<tr>
<td>COQ</td>
<td>Coenzyme Q</td>
</tr>
<tr>
<td>COX</td>
<td>Cytochrome C oxidase</td>
</tr>
<tr>
<td>COXI</td>
<td>Cytochrome C oxidase subunit 1</td>
</tr>
<tr>
<td>COXII</td>
<td>Cytochrome C oxidase subunit 2</td>
</tr>
<tr>
<td>COXIII</td>
<td>Cytochrome C oxidase subunit 3</td>
</tr>
<tr>
<td>CPEO</td>
<td>Chronic Progressive External Ophthalmoplegia</td>
</tr>
<tr>
<td>Cq</td>
<td>Cycle of quantification</td>
</tr>
<tr>
<td>CRGH</td>
<td>Centre for Reproductive and Genetics Health</td>
</tr>
</tbody>
</table>
CT  Cytoplasmic transfer
CV  Complex V
Cy  Cyanine
Cyt Cytochrome
D-Loop Displacement loop
DG Degenerated
dH2O Deionised water
DLB Denaturing lysis buffer
DMSO Dimethyl sulfoxide
DNA Deoxyribonucleic acid
dNTPs Deoxyribonucleotide triphosphate
dsDNA Double stranded deoxyribonucleic acid
DTT Dithiothreitol
dUTP Deoxyuridine triphosphate
E Embryo
EDTA Ethylenediaminotetra acetic acid
EGA Embryonic genome activation
EMPOP Mitochondrial DNA sequence database
ESC Embryonic stem cell
ESHRE European Society of Human Reproduction and Embryology
ETC Electron transport chain
FISH Fluorescence in situ hybridization
FRET Fluorescence resonance energy transfer
G Guanine nucleotide base
GATK genome analysis toolkit
GV Germinal vesicle
GWAS Genome wide association studies
H-strand Heavy strand
Hap Haplotype
HapMap Haplotype map
HFEA Human Fertilisation and Embryology Authority
HGDP-CEPH Human genome diversity cell line
HLA Human leukocyte antigen
HRM High resolution melting
HS High Sensitivity
HSP Heavy strand promoter
HYD Hydrophobicity
Hp protons
IBS Identical by state
ICH Institute for Child's Health
ICM Inner cell mass
ICSI Intra-cytoplasmic sperm injection
IFWH Institute for Women's Health
IGV Integrative genomic viewer
IMM Inner mitochondrial membrane
Indels Small insertions and deletion
IUPAC International Union of Pure and Applied Chemistry
IVF In vitro Fertilisation
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>Kilo</td>
</tr>
<tr>
<td>Kb</td>
<td>Kilo base</td>
</tr>
<tr>
<td>KDa</td>
<td>Kilo Dalton</td>
</tr>
<tr>
<td>L</td>
<td>Litre</td>
</tr>
<tr>
<td>L-strand</td>
<td>Light strand</td>
</tr>
<tr>
<td>LC</td>
<td>Light cycler</td>
</tr>
<tr>
<td>LD</td>
<td>Linkage disequilibrium</td>
</tr>
<tr>
<td>LHON</td>
<td>Leber's hereditary optic neuropathy</td>
</tr>
<tr>
<td>LR-PCR</td>
<td>Long range PCR</td>
</tr>
<tr>
<td>LSP</td>
<td>Light strand promoter</td>
</tr>
<tr>
<td>M</td>
<td>Molar</td>
</tr>
<tr>
<td>Mb</td>
<td>Mega base (pair)s</td>
</tr>
<tr>
<td>MDA</td>
<td>Multiple displacement amplification</td>
</tr>
<tr>
<td>MELAS</td>
<td>Mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes</td>
</tr>
<tr>
<td>MERRF</td>
<td>Myoclonic epilepsy and ragged-red fibres</td>
</tr>
<tr>
<td>MFN1</td>
<td>Mitofusion 1</td>
</tr>
<tr>
<td>Mg²⁺</td>
<td>Magnesium ion</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>Magnesium chloride</td>
</tr>
<tr>
<td>MgSO₄</td>
<td>Magnesium sulfate</td>
</tr>
<tr>
<td>MHC</td>
<td>Major Histocompatibility Complex</td>
</tr>
<tr>
<td>MI</td>
<td>First meiotic division</td>
</tr>
<tr>
<td>MIDD</td>
<td>Maternally-inherited diabetes and deafness</td>
</tr>
<tr>
<td>MII</td>
<td>Metaphase of the second meiotic division</td>
</tr>
<tr>
<td>MILS</td>
<td>Maternally-inherited Leigh syndrome</td>
</tr>
<tr>
<td>MIM</td>
<td>Mitochondrial inner membrane</td>
</tr>
<tr>
<td>min</td>
<td>Minutes</td>
</tr>
<tr>
<td>mito-nuclear</td>
<td>Mitochondrial nuclear</td>
</tr>
<tr>
<td>ml</td>
<td>Millilitres</td>
</tr>
<tr>
<td>mm</td>
<td>Millimetres</td>
</tr>
<tr>
<td>mM</td>
<td>Milli molar</td>
</tr>
<tr>
<td>µM</td>
<td>Micromolar</td>
</tr>
<tr>
<td>MMP</td>
<td>Mitochondrial membrane potential</td>
</tr>
<tr>
<td>Mp</td>
<td>Mega bases</td>
</tr>
<tr>
<td>mPT pore</td>
<td>Mitochondrial permeability transition pore</td>
</tr>
<tr>
<td>MRB</td>
<td>Midi reaction buffer</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger ribonucleic acid</td>
</tr>
<tr>
<td>mtDNA</td>
<td>Mitochondrial deoxyribonucleic acid</td>
</tr>
<tr>
<td>MTERF</td>
<td>mitochondrial transcription termination factor</td>
</tr>
<tr>
<td>mtUPR</td>
<td>mitochondrial Unfolded Protein Response</td>
</tr>
<tr>
<td>n</td>
<td>Number</td>
</tr>
<tr>
<td>N/A</td>
<td>Not available</td>
</tr>
<tr>
<td>Na</td>
<td>Sodium</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium chloride</td>
</tr>
<tr>
<td>NADH</td>
<td>Nicotinamide adenine dinucleotide, reduced</td>
</tr>
<tr>
<td>NaOH</td>
<td>Sodium hydroxide</td>
</tr>
<tr>
<td>NARP</td>
<td>Neurogenic weakness, ataxia and retinitis pigmentosa</td>
</tr>
<tr>
<td>NC</td>
<td>No-call</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Qscore</td>
<td>Quality score</td>
</tr>
<tr>
<td>$r^2$</td>
<td>Coefficient of determination in statistics</td>
</tr>
<tr>
<td>Ramp</td>
<td>Rate of amplification</td>
</tr>
<tr>
<td>rBAM</td>
<td>reordered BAM files</td>
</tr>
<tr>
<td>RC</td>
<td>Respiratory chain</td>
</tr>
<tr>
<td>rCRS</td>
<td>Revised Cambridge Reference Sequence</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomised controlled trial</td>
</tr>
<tr>
<td>REC</td>
<td>Research ethics committee</td>
</tr>
<tr>
<td>RefSeq</td>
<td>Reference Sequence</td>
</tr>
<tr>
<td>RFLP</td>
<td>Restriction fragment length polymorphism</td>
</tr>
<tr>
<td>RIF</td>
<td>Repeated implantation failure</td>
</tr>
<tr>
<td>RLU</td>
<td>Relative light units</td>
</tr>
<tr>
<td>RM</td>
<td>Repeated miscarriage</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RNASEH1</td>
<td>Ribonuclease H1</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive oxygen species</td>
</tr>
<tr>
<td>RPL</td>
<td>Recurrent pregnancy loss</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolution per minute</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions per minute</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal ribonucleic acid</td>
</tr>
<tr>
<td>RSB</td>
<td>Re-Suspension Buffer</td>
</tr>
<tr>
<td>RT</td>
<td>Room temperature</td>
</tr>
<tr>
<td>S</td>
<td>Seconds</td>
</tr>
<tr>
<td>SAMtools</td>
<td>Sequence alignment/map.</td>
</tr>
<tr>
<td>SAP</td>
<td>Shrimp alkaline nuclease</td>
</tr>
<tr>
<td>SC</td>
<td>Single cell</td>
</tr>
<tr>
<td>SCNT</td>
<td>Somatic cell nuclear transfer</td>
</tr>
<tr>
<td>SDH</td>
<td>Succinate dehydrogenase</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulphate</td>
</tr>
<tr>
<td>Ser</td>
<td>Serine</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>SOD</td>
<td>Superoxide dismutase</td>
</tr>
<tr>
<td>SOD</td>
<td>Superoxide dismutase</td>
</tr>
<tr>
<td>ssDNA</td>
<td>Single stranded deoxyribonucleic acid</td>
</tr>
<tr>
<td>STC</td>
<td>Ser/Thr content</td>
</tr>
<tr>
<td>STRs</td>
<td>Short tandem repeats</td>
</tr>
<tr>
<td>Y</td>
<td>Tyrosine nucleotide base</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris-borate-EDTA</td>
</tr>
<tr>
<td>TC</td>
<td>Threonine content</td>
</tr>
<tr>
<td>TCA</td>
<td>Tricarboxylic acid</td>
</tr>
<tr>
<td>TE</td>
<td>Trophoderm</td>
</tr>
<tr>
<td>TE buffer</td>
<td>Tris/EDTA buffer</td>
</tr>
<tr>
<td>TFAM</td>
<td>Mitochondrial transcription factor A</td>
</tr>
<tr>
<td>TFB1M</td>
<td>Mitochondrial transcription factor B1</td>
</tr>
<tr>
<td>TFB2M</td>
<td>Mitochondrial transcription factor B2</td>
</tr>
<tr>
<td>Thr</td>
<td>Threonine</td>
</tr>
<tr>
<td>$Tm$</td>
<td>Melting temperature</td>
</tr>
<tr>
<td>Tris</td>
<td>Trisaminomethane</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>tRNA</td>
<td>Transfer ribonucleic acid</td>
</tr>
<tr>
<td>UCH</td>
<td>University College Hospital</td>
</tr>
<tr>
<td>UCL</td>
<td>University College London</td>
</tr>
<tr>
<td>UQCRB</td>
<td>Ubiquinol-cytochrome c reductase binding protein</td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated region</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>V</td>
<td>Volts</td>
</tr>
<tr>
<td>v</td>
<td>version</td>
</tr>
<tr>
<td>VCF</td>
<td>Variant Call File</td>
</tr>
<tr>
<td>VDAC</td>
<td>Voltage-dependent anion channel</td>
</tr>
<tr>
<td>VEP</td>
<td>Variant effect predictor</td>
</tr>
<tr>
<td>vs.</td>
<td>Versus</td>
</tr>
<tr>
<td>VWR</td>
<td>chemical and laboratory scientific supplies company</td>
</tr>
<tr>
<td>w/v</td>
<td>Mass to volume</td>
</tr>
<tr>
<td>WGA</td>
<td>Whole Genome Amplification</td>
</tr>
<tr>
<td>Xq</td>
<td>q arm of chromosome X</td>
</tr>
<tr>
<td>YBP</td>
<td>Years before present</td>
</tr>
<tr>
<td>ZP</td>
<td>Zona pellucida</td>
</tr>
<tr>
<td>β-Me</td>
<td>Beta-mercaptoethanol</td>
</tr>
<tr>
<td>λ</td>
<td>Lambda</td>
</tr>
<tr>
<td>μ</td>
<td>Micro</td>
</tr>
</tbody>
</table>
Chapter 1 Introduction

In the late 70s, help was offered to many couples struggling to conceive with the introduction of in vitro fertilisation (IVF). Since that time, knowledge about embryo development, implantation and general fertility has grown. However, the arrest of some early embryos or the failure of implantation of normally developed embryos remains a question that has not been fully elucidated. Aneuploidy was found to be one of the main reasons behind this complex phenotype affecting gametes (mostly oocytes) and later the developing embryos.

Human fertility is affected by environmental and genetic factors. It is well-defined that age plays a significant role in reproduction. In females, fertility drops dramatically in an age-dependent manner due to the decline in the quality and quantity of the oocytes and follicle pool (Qi et al., 2014). This could be due to chromosome segregation errors occurring during female meiosis which worsen with advancing maternal age. About 25-30% of oocytes from women in their early 30s are chromosomally abnormal with a significant increase in aneuploidy up to 75-85% in the oocytes of women over 40 years old (Fragouli et al., 2011, Franasiak et al., 2014, Rodrigo et al., 2014, Wells et al., 2014). Embryos from couples undergoing preimplantation genetic diagnosis (PGD) cycles, who presumably have normal fecundity, also show chromosomal abnormalities (Delhanty, 2013). Thus, aneuploidy is not only restricted to couples with fertility problems, and it has been proposed that the susceptibility to chromosomal errors at early stages in life could be one of the reasons for low fertility in humans in comparison with other mammalian species.

Embryonic aneuploidy may lead to preimplantation embryonic developmental arrest, repeated implantation failure and repeated miscarriage. In Assisted
Reproductive Technology (ART), not all the oocytes that are collected have the same developmental potential despite being produced under similar conditions. The morphological feature of an embryo, even when monitored during embryo culture, does not provide information about its chromosomal status and is weakly correlated with its potential to implant and form a viable pregnancy (Fragouli et al., 2014). At present, the exact cellular mechanisms that affect embryo quality have not been revealed. The proper function of the energy-supplying organelles, mitochondria, have been proposed to influence the developmental potential of the oocyte and later of the embryo strongly (Wang et al., 2009).

1.1 Fertilisation

At fertilisation, the oocyte completes meiosis and the fertilised oocyte is called the zygote. Sperm and oocyte nuclei fuse resulting in syngamy. Following that, several rounds of mitosis occur whilst the fertilised egg travels through the oviduct towards the uterus. This leads to a decrease in the size of embryonic cells (called blastomeres) due to the absence of cell growth between the cell divisions. This stage of embryonic division is known as the cleavage stage (Figure 1–1). During cleavage divisions, the embryonic genome becomes transcriptionally active in a process known as embryonic genome activation (EGA). This occurs between 4 to 8-cell stage in humans (Braude et al., 1988), the 2-cell stage in mice (Bolton et al., 1984), at 4-cell stage in pigs (Jarrell et al., 1991), and around 8 to 16-cell stage in sheep (Calarco and McLaren, 1976). In order for embryos to complete this process efficiently, they require energy in the form of Adenosine Triphosphate (ATP). ATP can either be produced by fermentation through glycolysis to form lactate or aerobic respiration by the oxidative phosphorylation system (OXPHOS) in the mitochondria.
Evidence suggests that, aerobic mitochondrial metabolism contributes a major role in the supply of ATP. A second pathway, fermentation, is also active and the two pathways work in synchrony to supply all the necessary ATP (Wilding et al., 2009).

Fertilisation results in the formation of a zygote. Following that, cleavage divisions result in an increase in cell number leading to the formation of a morula and the blastocyst in later stages.

At the 8 to 16-cell stage, embryonic blastomeres begin to increase cell-cell contacts and flatten, in a process known as compaction (Pratt et al., 1982). Fluid starts to accumulate between blastomeres forming a cavity known as the blastocoel in the compacted morula. At this stage, the embryo is known as a blastocyst and is made of two cell types, those of the inner embryo, as the inner cell mass (ICM), that give rise to the fetal tissues and those of the outer layer known as the trophoblasts, which will then form the placenta (Gardner and Johnson, 1972). Once it reaches the uterus about 5 days post fertilisation, the Zona Pellucida (ZP) is lost by “hatching” and the blastocyst becomes free to implant in the wall of the uterus.

Several important processes such as parts of meiosis, fertilisation and regulation of early cleavage are controlled by the oocyte (Polanski, 1997, Acton et al., 2004). Epigenetic developmental changes of the embryonic genome which manifest later in
embryogenesis are also regulated by the oocytes (Latham and Sapienza, 1998, Acton et al., 2004). Hence, the variability of embryonic developmental competence is reflected by oocyte quality. This variability has been recognised in ART clinics since not all the mature oocytes have the same developmental fate despite being produced at the same time under similar conditions. Only 10-20% of the transferred embryos into the human uterus implant to form a viable embryo while the average fertilisation rate of a mature oocyte is 60-70% (Wilding et al., 2009). This wide range of variability in embryonic developmental potential, has led to the identification of factors that might determine oocyte quality (Agarwal et al., 2006). Maternal factors that may lead to abnormal preimplantation development include proteins, cytoplasmic maternal messenger ribonucleic acid (mRNA), antioxidants and organelles including mitochondria, the most abundant organelles in the oocyte cytoplasm (Acton et al., 2004, Wilding et al., 2009).

1.2 Mitochondria

Mitochondria are double-membrane bound organelles which play an essential role in almost all the cells in the body. Their main function is to support aerobic respiration and produce, by OXPHOS, the majority of cellular energy in the form of ATP (Ramalho-Santos et al., 2009, Tuppen et al., 2010). In addition to ATP synthesis, mitochondria control cytosolic calcium homeostasis, control apoptosis and integrate cellular stress signals and have a role in steroid biogenesis (Reeve, 2012). They are known to produce a significant amount of endogenous reactive oxygen species (ROS) and host other important biochemical pathways, including the tricarboxylic acid (TCA) cycle and parts of the urea cycle. These organelles are central to iron–
sulphur cluster biogenesis, the only entirely conserved function of mitochondria in eukaryotes (Tuppen et al., 2010).

Mitochondria have distinct features based on the inner membrane invaginations called cristae and matrix structure (Figure 1–2). Based on the cell's function and type, mitochondria may present a wide range of morphologies, vary in number and function in a heterogeneous fashion. In fact, mitochondria have many similarities with a-proteobacteria (Gray, 2012). Most importantly, it has been characterised to have its own genome.

![Figure 1–2: The mitochondrial inner structure.](image)

*Mitochondria are comprised of two membranes, the outer and inner. Each membrane is made of a phospholipid bilayer. The outer membrane is permeable through which molecules (up to 10kDa) can pass. The inner membrane, contains all the enzymes involved in the respiratory chain, which serve to support the process of oxidative phosphorylation. The mitochondrial matrix, which is formed through the intricate folding of the inner membrane is made up of cristae. These serve to increase the surface area of the inner membrane, which contains both the enzymes of the tricarboxylic acid cycle (TCA cycle) as well other structures and molecules including mitochondrial DNA, ribosomes and matrix granules. Adapted from Frontiers in Genetics, 5 (448), Levin et al., Mito-nuclear co-evolution: the positive and negative sides of functional ancient mutations, pages 1-10, Copyright (2014), (https://creativecommons.org/licenses/by/4.0/). This diagram has been altered from its original version. Modifications:” mtDNA position”.*
1.2.1 Function of Mitochondria in the eukaryotic cell

1.2.1.1 ATP synthesis

ATP synthesis by OXPHOS is the most important function of mitochondria which has led to these organelles being recognised as the “powerhouse” of the cell. The OXPHOS is composed of more than 80 different polypeptides, organised into five trans-membrane complexes (C) (CI, CII, CIII, CIV and CV). The oxidation of fat and carbohydrates results in the production of reducing equivalents NADH which passes electrons into complex I and FADH$_2$ which passes electrons to complex II. Electrons, flow through the mitochondrial inner membrane (MIM) in the electron transport chain (ETC) from reduced to oxidised states, ultimately terminating with the reduction of oxygen to water. In the ETC this process starts with oxidation of NADH (NADH: coenzyme Q oxidoreductase or NADH dehydrogenase) by complex I or succinate (FADH$_2$: CoQ oxidoreductase or succinate dehydrogenase) by complex II. The electrons are then transferred to CoQ (COQ), complex III, cytochrome c, complex IV (cytochrome c oxidase or COX), and finally to oxygen. The energy released when the electrons traverse complexes I, III, and IV, is used to pump protons from the mitochondrial matrix across the MIM to the inter-membrane space (Wallace and Chalkia, 2013). This generates an “electrochemical ion gradient” which enables complex V, ATP synthase, to phosphorylate adenosine diphosphate (ADP) forming ATP (Tuppen et al., 2010) (Figure 1–3).

At present, one of the most important factors affecting the fertilisation outcome of an oocyte and the developmental competence of embryo is linked to the mitochondrial ability to balance ATP supply and demand (Van Blerkom, 2011).
The respiratory chain in mammals is arranged in five complexes embedded within the mitochondrial inner membrane. NADH–ubiquinone oxidoreductase representing complex I (CI), succinate–quinone oxidoreductase representing complex II (CII), cytochrome bc\(_1\) as complex III (CIII), cytochrome c-oxidase as complex VI(CVI) and ATP synthase as complex V(CV). Electron transport between complexes is mediated by membrane-embedded ubiquinone (Q) and soluble cytochrome c. CI and CII are the entry points for electrons from NADH and FADH\(_2\) respectively. Electrons from NADH, are used to reduce Q to ubiquinol (QH\(_2\)) which will be then used by CIII to reduce cytochrome c in the intermembrane space (IMS). CIV uses cytochrome c to reduce O\(_2\), which is the final electron acceptor. For each NADH oxidised molecule, protons are pumped across the membrane from the matrix to the IMS. The energy derived from the electron transport is used to force out protons (H\(^+\)) from the mitochondrial matrix across the MIM creating an electrochemical proton-motive force, associated with an inside-negative mitochondrial membrane potential and increased matrix pH. The backflow of H\(^+\) is used by CV to phosphorylate ADP forming ATP. Reproduced from Nature Reviews Molecular Cell Biology, 16, Sazanov, A giant molecular proton pump: structure and mechanism of respiratory complex I, pages 375-388, Copyright (2015), with permission from Springer Nature.

**1.2.1.2 Generation of ROS**

ROS are highly reactive oxidizing agents belonging to a class of free radicals. ROS are generated in the mitochondria as a result of cellular metabolism. This is important because it contributes to retrograde redox signalling from the organelle to the cytosol and nucleus and it underlies oxidative damage in many pathologies (Murphy, 2009). Between 0.2-2% of the oxygen taken up by the cells is transformed to ROS by mitochondria. Mitochondrial ROS production is very sensitive to the proton motive force (Ramalho-Santos et al, 2009). At several sites along the ETC (mainly CI and CIII), electrons can react directly with other electron acceptors or oxygen and produce free radicals. Damage to cellular macromolecules such as...
proteins, lipids and DNA may occur when cellular production of ROS exceeds its antioxidant capacity. Although mitochondria provide a continuous protection mechanism through various enzymatic defence systems and several antioxidants they contribute in the pathogenesis of some human diseases through the "oxidative stress" where the production of ROS exceeds the antioxidant defence (Venkatesh et al., 2009). Recent evidence suggests a specific role of ROS in cell signalling (Thannickal and Fanburg, 2000, Ramalho-Santos et al., 2009)(Figure 1-4).

Figure 1–4: ATP generation processes inside the mitochondria and the production of ROS.

Pyruvate and ATP are produced by the break-down of carbohydrates in the process of glycolysis. These can be transformed into acetyl-CoA which is the substrate for the Krebs cycle. In this cycle the acetyl-group is donated to oxaloacetate, which generates citric acid; this is dissimilated step-by-step resulting in the generation of ATP and the production of the reducing agent NADH. Mitochondrial beta-oxidation causes the biochemical dissimilation of Fatty acids (FAs), which also results in acetyl-CoA. NADH is also produced during this process along with another reducing agent; FADH₂. Both reducing agents, NADH and FADH₂, drive ATP production by OXPHOS. ROS can damage the mtDNA, they are produced as by-products of the OXPHOS pathway. Reproduced from Human Reproduction Update, 21, Otten and Smeets, Evolutionary defined role of the mitochondrial DNA in fertility, pages 671-698, Copyright (2015), with permission from Oxford University Press.
1.2.1.3 Apoptosis

Apoptosis (programmed cell death) is characterised morphologically by apoptotic body formation and membrane blebbing, nuclear condensation, DNA fragmentation, relative sparing of organelles and the absence of an inflammatory response. Mitochondria have a major role in apoptosis via the intrinsic apoptotic pathway which is believed to have a crucial role in the pathogenesis and progression of neurodegenerative disorders associated with a Respiratory Chain (RC) deficiency, several physiological functions and in embryological development (Leonard and Schapira, 2000).

Apoptosis is induced by the change in the permeability of the Mitochondrial Membrane Potential (MMP). MMP is generated by proton transfer though complexes I, III and IV from the mitochondrial matrix to the intramembranous space. A decrease in MMP opens the pore (permeability transition), allowing the release of small molecules from the mitochondrion including the apoptosis initiating factors. The mitochondrial apoptotic pathway is triggered by a wide range of stress stimuli such as high levels of ROS and other OXPHOS inhibitors. This leads to an increase in the permeability of mitochondrial proteins such as cytochrome c, which plays a prominent role in promoting the caspase cascade of cell execution (Khosravi-Far and Esposti, 2004). Once released to the cytoplasm, cytochrome c activates the caspase cascade pathway that can induce cellular degradation (Ramalho-Santos et al., 2009, Wang et al., 2009). Several factors can regulate mitochondrial mediated apoptosis. The Bcl-2 family members may either lead to cell death (Bid, Bik, Hrk, Bok, Bax, Bak, Bcl-Xs, Bad) or promote cell survival (Bcl-w, Mcl-1, Bcl-2, Bcl-xL, A1/Bfl-1) interacting to form homo- and heterodimers, their relative abundance being
the determinant of apoptotic threshold (Figure 1–5) (Zamzami et al., 1998, Leonard and Schapira, 2000, Ramalho-Santos et al., 2009).

![Diagram of mitochondria and apoptosis]

**Figure 1–5: Mitochondria and apoptosis.**

An intact RC which allows for effective proton pumping is necessary for high membrane potential and closure of the permeability transition pore (PTP). This closure can be caused by ligands such as benzodiazepine (BZP) agonists, cyclosporine and Bcl2. A decrease in membrane potential or response to ligands such as atractyloside, Bax or Bak can result in the opening of the pore. Apoptosis-initiating factors (AIF) and cytochrome c are released during permeability transition, which ends the activation of apoptosis and the caspase cascade. This release of cytochrome c as well as the PTP opening causes increased generation of free-radicals which impair OXPHOS and cause mitochondrial swelling. This pathway may result in necrosis (premature cell death). Reproduced from *The Lancet*, 355, Leonard and Schapira, Mitochondrial respiratory chain disorders II: neurodegenerative disorders and nuclear gene defects, pages 389-394, Copyright (2000), with permission from Elsevier.

In recent years an increasing number of reports have shown that mitochondrial dysfunction has been associated with aging and several pathologies, including infertility (Au et al., 2005, Ramalho-Santos et al., 2009).
1.2.2 The Mitochondrial Genetic System

1.2.2.1 The Mitochondrial DNA (mtDNA)

As mitochondria are maternally inherited, mitochondrial genetics differs from Mendelian genetics of nuclear encoded genes. Mitochondria is believed to have a bacterial evolutionary origin that originated between one and half to two billion years ago from the fusion of an ancient alpha proteobacterium with an ancient eukaryotic cell producing a “symbiont” (Bar-Yaacov et al., 2012).

Mitochondrial DNA (mtDNA) can be considered as the smallest chromosome in the human genome. It is composed of 16,569 base pairs. It is circular and double stranded DNA with a heavy strand (H-strand) and light strand (L-strand). The number of mitochondria varies from cell to cell. Unlike, nuclear DNA (nDNA) most somatic cells contain between 1000 and 10,000 mtDNA molecules depending on their energy demand (Trifunovic, 2006). Lymphocytes contain around 1000 while red cells and some terminally differentiated skin cells have no mtDNA. It is estimated that cells of tissue with high energy requirements (such as neurons and muscle cells) contain thousands of mtDNA molecules compared to leukocytes and monocytes which contain hundreds. Interestingly, the largest number of mitochondrial is found in the oocyte which possesses approximately 100 000 mitochondria having more than 150 000 mtDNA molecules while mature sperm contains 100 copies of mtDNA indicating the importance of mitochondria in early embryonic development (Wai et al., 2010). Multiple copies (~6–10) of mtDNA molecules are organised in nucleoprotein complexes known as “nucleoids” which lack protective histones (St. John, 2007, Scarpulla, 2008a, Scarpulla, 2008b, Chiaratti et al., 2011).
Mitochondrial DNA encodes 37 genes in vertebrates, including: two ribosomal RNA genes (rRNAs) (12S and 16S), 22 transfer RNA genes (tRNAs) and 13 protein-coding genes which encode components of the respiratory chain operating the OXPHOS system (Wallace, 2001). These include seven protein subunits (ND1-6, ND4L) of complex I, one subunit (cytochrome b) of complex III, 3 subunits (COI-III) of complex IV and two subunits (ATP 6, 8) of complex V. The rest of the mtDNA comprises two major noncoding elements: the displacement loop (the D-Loop) and the shorter noncoding region of the light strand. The D-Loop includes promoters of transcription of mtDNA light and heavy strands and at least some of the proposed mtDNA replication origins \((O_h)\) (Chang and Clayton, 1985). The mitochondrial DNA genes lack introns and intergenic noncoding nucleotides almost do not exist (Tuppen et al., 2010, Bar-Yaacov et al., 2012).

Although mtDNA repair system does exist (Alexeyev et al., 2013), it is not sufficient to overcome the oxidative damage sustained by the mitochondrial genome due to the proximity to the RC complexes in the MIM which generate ROS (Tuppen et al., 2010). Therefore, it is not surprising that the mtDNA mutation rate is considerably high (10-17 fold higher) compared to the nuclear DNA (Table 1-1) (Tuppen et al., 2010). Moreover, the mtDNA mutation rates is higher in animals compared to plants. This allows animals to adapt more rapidly to various environmental conditions in terms of energy production (Otten and Smeets, 2015).
Table 1-1: Comparison between the human mitochondrial and nuclear genomes.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Nuclear genome</th>
<th>Mitochondrial genome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>~3.3 x 10^9 bp</td>
<td>16,569 bp</td>
</tr>
<tr>
<td>Number of DNA molecules per cell</td>
<td>23 in haploid cells; 46 in diploid cells</td>
<td>Several thousand copies per cell (polyploidy)</td>
</tr>
<tr>
<td>Number of genes encoded</td>
<td>~20,000–30,000</td>
<td>37 (13 polypeptides, 22 tRNAs and 2 rRNAs)</td>
</tr>
<tr>
<td>Gene density</td>
<td>~1 per 40,000 bp</td>
<td>1 per 450 bp</td>
</tr>
<tr>
<td>Introns</td>
<td>Frequently found in most genes</td>
<td>Absent</td>
</tr>
<tr>
<td>Percentage of coding DNA</td>
<td>~3%</td>
<td>~93%</td>
</tr>
<tr>
<td>Codon usage</td>
<td>The universal genetic code</td>
<td>AUA codes for methionine; TGA codes for tryptophan; AGA and AGG specify stop codons</td>
</tr>
<tr>
<td>Associated proteins</td>
<td>Nucleosome-associated histone proteins and non-histone proteins</td>
<td>No histones; but associated with several proteins (for example, TFAM) that form nucleoids</td>
</tr>
<tr>
<td>Mode of inheritance</td>
<td>Mendelian inheritance for autosomes and the X chromosome; paternal inheritance for the Y Chromosome</td>
<td>Exclusively maternal</td>
</tr>
<tr>
<td>Replication</td>
<td>Strand-coupled mechanism that uses DNA polymerases α and δ</td>
<td>Strand-coupled and strand-displacement models; only uses DNA polymerase γ</td>
</tr>
<tr>
<td>Transcription</td>
<td>Most genes are transcribed individually</td>
<td>All genes on both strands are transcribed as large polycistrons</td>
</tr>
<tr>
<td>Recombination</td>
<td>Each pair of homologues recombines during the prophase of meiosis</td>
<td>There is evidence that recombination occurs at a cellular level but little evidence that it occurs at a population level</td>
</tr>
<tr>
<td>Histones</td>
<td>Present</td>
<td>absent</td>
</tr>
<tr>
<td>Repair system</td>
<td>Better</td>
<td>Poor</td>
</tr>
<tr>
<td>Mutation rate</td>
<td>Lower</td>
<td>10-17 folds higher</td>
</tr>
</tbody>
</table>

TFAM, mitochondrial transcription factor A; rRNA, ribosomal RNA. Adapted from Nature Reviews Genetics, 6 (5), Taylor and Turnbull, Mitochondrial DNA mutations in human disease, pages 1-29, Copyright (2005), Springer Nature.

The noncoding human mtDNA control region contains highly polymorphic sites known as the hypervariable sites. These sites evolve at a rate much faster than average and have been well documented via various analyses of human mtDNA variation (Hasegawa et al., 1993, Wakeley, 1993, Excoffier and Yang, 1999, Meyer et al., 1999). Studies of evolutionary rate examination have identified that germline and somatic mtDNA mutations occur preferentially at these sites suggesting that hypervariable sites in human mtDNA represent mutational hotspots (Stoneking, 2000). Although these mechanisms have been very successful at the level of the
species, they can have contrary effects for the individual. In humans it could result in severe or unpredictable phenotypes such as fertility problems and premature ageing (Otten and Smeets, 2015).

1.2.2.2 Mitochondrial haplogroups (Ancient adaptive polymorphism)

During evolution, the high mutation rate in mtDNA resulted in high levels of population-specific mtDNA base substitutions along radiating maternal lineages (Wallace and Chalkia, 2013). As humans migrated out of Africa to populate the globe, new branches of the mtDNA trees were generated with the rise of new mtDNA polymorphisms. If mitochondrial physiology was beneficially changed by a founder mutation in individuals within a region, that mtDNA lineage became enriched in that geographic environment. Each continent and geographical region is associated with a unique specific combination of single or short variations in mtDNA sequence, most often in the D-Loop, known as “mitochondrial haplogroups”. These population-specific polymorphisms have been used to establish historical human migrations, and the maternal ethnic origins (Wallace and Chalkia, 2013).

Mitochondrial haplogroups are classified according to letters (A-Z) in humans. All African mtDNAs are related and contained within one large continent-specific lineage specified “macro-haplogroup L” which developed between 130,000 and 200,000 years before present (YBP). This root was the source of four lineages specific for sub-Saharan Africa: L0, L1, L2 and L3. Haplogroup L3 gave rise to two mtDNAs designated M and N. Only the mtDNA descendants from M and N mtDNAs left Africa to colonise the rest of the world (Figure 1-6). As humans migrated, macro-haplogroup M carriers moved along tropical Southeast Asia, reaching Australia.

Later, M descendants moved north out of Southeast Asia to generate mtDNA
haplogroups: C, D, G, and M1 – M20 in central and eastern Asian. On the other hand, macro-haplogroup N carriers showed a bi-directional migration pattern. Haplogroups A and Z were formed through the first movement of N through Southeast Asia to Australia and from southern Asia into central Asia. In the other migration movement, the European haplogroups I, X, and W were formed when macro-haplogroup N moved northward out of Africa to Europe. Macro-haplogroup N also gave rise to sub-macro-haplogroup R which gave rise to the remaining European haplogroups H, J, Uk, T, U, and V which emerged 39,000–51,000 YBP in western Europe. The Asian mtDNA haplogroups B and F were also formed through the eastern movement of R (Wallace and Chalkia, 2013, Stewart and Chinnery, 2015).

1.2.2.2.1 Importance of mitochondrial haplogroups

Important mtDNA evolutionary variations that correspond with the major human geographic migrations could not have occurred by chance. It is most likely that these mtDNA changes allowed human ancestors to adapt to different regional environments, acting as an adaptive system (Ruiz-Pesini et al., 2004, Mishmar et al., 2006, Ruiz-Pesini and Wallace, 2006). For example; macro-haplogroup N, which moved directly from Africa into the Europe, harboured two polypeptide variants in ND3 and ATP6 which have been associated with alterations in the MMP and coupling efficiency of mitochondrial OXPHOS making them looser (Wallace et al., 1999) 2013). This resulted in an increase in the number of calories burned by the mitochondria to generate the ATP. Suggesting that adaptation took place since individuals had to burn more calories to resist the cold stress. By contrast, macro-haplogroup M mtDNA variations, did not acquire comparable functional mtDNA
mutations as it remained in the tropics. Thus, the tight coupling of OXPHOS and heat production is minimised, and ATP production is maximised (Ruiz-Pesini et al., 2004, Mishmar et al., 2006, Ruiz-Pesini and Wallace, 2006).

Another example of the importance of functional variants forming and defining the branches of the mtDNA haplogroup trees can be seen in haplogroups J and T which were founded by two polypeptide gene amino acid substitution variants in ND1 (in complex I) and cytb (in complex III). The lineages then split and each of the founding polypeptide substitutions changes an evolutionarily highly conserved amino acid (Wallace and Chalkia, 2013) raising the possibility that selection might have been important during human mtDNA evolution (Stewart and Chinnery, 2015).

Due to maternal inheritance, mtDNA single-nucleotide variants that have accumulated throughout human history remain in total linkage disequilibrium. Therefore, the significance of a mtDNA variant is strongly influenced by the pre-existing mtDNA variants on which it arose, and some variants may have a negative effect on mitochondrial function in some haplogroup backgrounds while showing an adaptive response when present in other haplogroup backgrounds. This was clearly demonstrated in the mtDNA variant in ND1. When this variant arises on macro-haplogroup N it causes an amino acid substitution, reduces mitochondrial complex I activity and markedly increases the penetrance of the milder Leber’s hereditary optic neuropathy (LHON) mutations (Brown et al., 1995, Liang et al., 2009). However, if the mutation arises on a macro-haplogroup M background, the variant is associated with maximum complex I activity and adaptation to high altitude (Ji et al., 2012).

Mitochondrial haplotypes can also be related to the tendency to the onset of disease phenotype (Brown et al., 1995). It has been proposed that different haplogroups present discrete changes in ATP production, mtDNA copy levels and
lactate levels, which may reflect adaptation to environmental conditions (St. John, 2014). Several studies have reported the association between common human diseases with polymorphic mtDNA variants. Different haplogroups have been associated with a wide range of pathologies by either promoting or conferring protection against others (Ruiz-Pesini et al., 2000, Baudouin et al., 2005, Saxena et al., 2006, Lorente et al., 2013, Nardelli et al., 2013). Alleles that increased disease risk are more common than protective alleles indicating that natural selection has not removed deleterious novel sub-haplogroup variants. The non-coding mtDNA D-loop contains the majority of high-risk alleles. This is evident in mtDNA gene expression and replication. Subtle differences, therefore in mtDNA replication or transcription may contribute to the pathogenesis of typical age-related disorders. Pathologies associated with mitochondrial haplogroups include; complex and age-related diseases, metabolic and degenerative diseases and various cancers (reviewed by Wallace and Chalkia, 2013).
Figure 1-6: Mitochondrial haplogroups and major human population migration.

The so-called ‘Mitochondrial Eve’, is suggested to have lived and originated in Africa. This first common ancestor was the root source of four distinct lineages specific to sub-Saharan Africa: L0, L1, L2 and L3. The haplogroups M and N arose from Africa and populated the rest of the world. MtDNA migration routes are shown in arrows. Haplogroup N was directed to Eurasia as humans migrated and haplogroup M lineages moved to Asia, which gave rise to haplogroups A, B, C, D, G and F. Haplogroup N in Europe led to haplogroup R; this is the root of the European haplogroups H, J, T, U and V. Australian haplogroups S, P, and Q were then formed. East Asia and the Americas were populated with haplogroups A, B, C and D (Stewart and Chinnery, 2015). Mitochondrial haplogroups are continent-specific, with little mixing of mtDNA haplogroups from different continents. Reproduced from Nature Reviews Genetics, 16, Stewart and Chinnery, The dynamics of mitochondrial DNA heteroplasmy: implications for human health and disease, pages 530-542, Copyright (2015), with permission from Springer Nature.

Despite the importance of mitochondrial haplogroups and their effect on mitochondrial function, a large number of genes (around 1500) required for the various activities involving mitochondria are encoded by the nuclear genome, necessitating the coordination between the two genomes. This includes 70 polypeptides involved in OXPHOS as well as factors responsible for regulating replication, transcription and parts of the mtDNA translation machinery (Tuppen et al., 2010). Hence, mitochondrial function depends on factors encoded by both nDNA and mtDNA suggesting the importance of mitochondrial-nuclear interactions to cellular health.
1.2.3 Mito-nuclear interaction

The high mutation rate in the mitochondrial genetic system must influence systems involving mitochondrial-nuclear (mito-nuclear) interactions. A particular example is in mitochondrial translation and the energy producing OXPHOS machineries, which are comprised of factors encoded by both genomes. Further, mtDNA replication systems and transcription of mitochondrial RNA are both operated by nDNA-encoded proteins which bind mtDNA regulatory elements (Levin et al., 2014). A tight selection towards mito-nuclear “co-evolution” or coadaptation is required to maintain both systems. The process must be strictly regulated in order to sustain the mitochondrial activity necessary for the viability of cells and organisms because no cellular function can proceed without sufficient energy.

There are several modes of mito-nuclear interactions including: subunit compatibilities in the OXPHOS system (protein-protein interactions), mitochondrial transcription and replication (protein-DNA interactions) and mitochondrial translation (protein-RNA interactions) (Bar-Yaacov et al., 2012) (Figure 1–7). It is unclear whether mito-nuclear co-evolution occurs only to maintain mitochondrial functions during evolution or serves as an adaptive tool to adjust for the evolving energetic demands as the complexity of species increases (Levin et al., 2014).
1.2.3.1 Subunit compatibilities (protein-protein interaction)

Oxidative phosphorylation depends on the functional compatibility between a large number of nuclear encoded subunits in the ETC while the rest are encoded by the mitochondria. Mitochondrial gene encoded proteins interact closely with nuclear gene encoded proteins (Figure 1-8). It has been suggested, that a mismatch caused by incompatibilities between subunits of OXPHOS complexes, results in small misalignments of respiratory chain subunits, which slow electron transfer, thus decreasing respiratory capacity and increasing free radical leak (Lane, 2011). This eventually leads to an increase in cellular oxidative stress, unbalanced ROS production & apoptosis (Burton and Barreto, 2012). The excessive production of ROS has been shown to result in redox damage and causes stress that can compromise preimplantation embryo development in culture conditions (Johnson,
Free-radical leak activates mtDNA replication to compensate for insufficient ATP generation, in a process known as reactive biogenesis. A threshold model has been proposed by Lane (2011) that determines which route a cell will take based on the free radical leak mechanism. New respiratory complexes are generated if the leak is below a threshold while apoptosis occurs if the leak is above the threshold.

Thus, cells with mismatched nuclear and mitochondrial encoded subunits face a decline in respiratory capacity and free radical leakage which has two important consequences for the evolution of eukaryotes. It creates a barrier between different species and it favours the evolution of two sexes.
A: Reactive biogenesis below a threshold, mitochondrial biogenesis is stimulated in response to free radical signals to optimize respiration. Mitochondria with low respiratory capacity are reduced (depicted in red) and leak free radicals. These activate redox-sensitive transcription factors, upregulating the expression of mtDNA and nuclear genes, thus improving respiration (blue mitochondria). Mitochondrial populations decline with falling demand, thereby balancing ATP production and mitochondrial density over time. Fluctuations in free-radical leak over time is represented by the red line; yellow arrow depicts ATP synthesis. Dotted line depicts the apoptotic threshold.

B: Reactive biogenesis above a threshold; mismatch between mtDNA and the nuclear background results in small misalignments of respiratory chain subunits, slowing the electron transfer, decreasing respiratory capacity and increasing free radical leak. In this case, the free-radical leak preferentially amplifies the most mismatched (i.e. most deficient) mitochondria. Free-radical leak rises above the threshold, leading to apoptosis.

C: Proactive biogenesis; stimulated by signal from outside the cell (e.g. calorie restriction; blue arrow) without respect to local demand or redox conditions; the entire population is amplified, without selection for deficient mitochondria, reducing demand per mitochondrion, so improving respiratory capacity (blue mitochondria). As long as the external signal persists This situation remains stable lowering free-radical leak and extending lifespan. Reproduced from BioEssays, 33, Lane, Mito-nuclear match: Optimizing fitness and fertility over generations drives ageing within generations, pages 860-869, Copyright (2011), with permission from John Wiley and Sons.

Therefore, investigating the reduction in mitochondrial function due to incompatibilities caused by certain polymorphisms affecting the expression of directly interacting mito-nuclear genes encoding subunits, at the preimplantation
stages of human embryonic development is essential to elucidate mitochondrial causes of infertility or the possible mechanisms resulting in embryonic developmental arrest.

1.2.3.2 Mitochondrial protein transcription and replication (protein-DNA interaction)

The human mtDNA replication mechanism is still unclear and is accomplished by nuclear-encoded factors. Cell signalling insures that energy requirement for replication is sufficient for the process to continue. Two different mechanisms have been proposed. The first mechanism suggests that the replication is initiated at the heavy strand origin (OH) leading to the L-strand displacement from the H-strand. The replication of the H-strand then continues until it reaches the L-strand origin (OL) while the L-strand remains single-stranded. The L-strand synthesis then starts in the opposite direction. The other suggested model is based on the coupled leading-lagging strand DNA replication that also starts at OH. In this mechanism, the lagging L-strand synthesis begins shortly after replication initiation as a result of the short Okazaki ribonucleotide fragments generation which will subsequently be converted to DNA. Although, the exact mechanism of mtDNA replication is still debated, it is likely that both models exist. Mitochondrial transcription machinery consists of a limited number of nuclear-encoded proteins: the transcription activator TFAM, the transcription factor TFB2M, the mitochondrial RNA polymerase (POLRMT), and the termination factor mTERF. It is initiated from a single promoter on the L-strand (LSP) and one of two promoters on the H-strand (HSP1 and HSP2) (Figure 1–10). The current understanding of mtDNA translation is incomplete and it is also known to include nuclear-encoded factors imported into the organelle.
**Figure 1–10**: Human mitochondrial genome.

The diagram represents the human mitochondrial genome with a greater view of the mammalian D-loop and transcription termination regions, shown in the linear form. The double stranded mtDNA is composed of the heavy (H) strand represented by the outer circle and the light (L) strand represented by the inner circle. Mitochondrial rRNAs are coloured in red while the 22 mt-tRNAs are coloured in black and identified by their single letter abbreviation. The 13 mitochondrial encoded subunits in complex I are colour coded according to the following: complex I (green), complex III (purple), complex IV (yellow) and complex V (blue). Major noncoding regions of the genome are represented in (grey) including the D-loop and (Ol). The origin of H-strand (Oh) replication is also indicated within the D-loop. H-strand transcription is initiated either from (HSP1), generating a short transcript that terminates at the RNR2/MTTL1 boundary (Term) under the guidance of the transcription termination factor MTERF, or from HSP2, generating polycistronic transcripts of the entire H-strand. Light strand promoter LSP denotes the L-strand initiation point that produces polycistronic transcripts for this strand and also generates RNA precursors for H-strand replication initiation. Transcription from all promoters requires the upstream binding of transcriptional activator TFAM, together with a single subunit RNA polymerase (POLRMT), which forms a heterodimeric complex with the transcription factor TFB2M (shown as TFB in the diagram). TFAM also binds to other regions of the D-loop; however, only binding to the CSB region is shown. Adapted from Biochimica et Biophysica Acta (BBA)-Bienergetics, 17979, Tuppen et al., Mitochondrial DNA mutations and human disease, pages 113-128, Copyright (2010), with permission from Elsevier.

The OXPHOS system depends on protein assembly and mitochondrial transport systems, as well as a number of nuclear-encoded factors, which are essential to maintain both the structure and function of mtDNA (Leonard and Schapira, 2000). A variety of factors are involved in coordinating mtDNA and nDNA expression. The most widely studied factor, encoded by the nDNA, is the peroxisome proliferator-activated receptor gamma coactivator 1 alpha (PPARGC1A). The nuclear respiratory
factors 1 and 2 (NRF1 and NRF2) are controlled by PPARGC1A, which in turn coordinates the expression of OXPHOS polypeptides (in the nucleus) and other factors controlling, transcription, translation and replication of the mtDNA at the nuclear level. At the mitochondrial level, however, both replication and transcription of mtDNA are regulated by factors imported from the cytoplasm by the mitochondria, which interact with the D-loop in mtDNA. Transcription factors B1 and B2 (TFB1M and TFB2M), a core RNA polymerase (POLRMT) and TFAM factor encoded in the nDNA also interact with the D-loop in mtDNA and are responsible for the initiation of its bidirectional transcription. A specific DNA polymerase, (POLG) as well as a helicase called the Twinkle helicase are essential for mitochondrial DNA replication (Spelbrink et al., 2001, Scarpulla, 2008a).

A variety of proteins are required to enable both the split and fusion of mitochondria. Mitochondrial dynamics is essential because defects in either splitting (Waterham et al., 2007) or joining can cause developmental anomalies (Chen et al., 2003) and disease (Alexander et al., 2000). Such processes enable the mitochondrial network to maintain mitochondrial quality by recycling damaged mitochondria and to regulate energy supply within the cell (Chiaratti et al., 2011). A new class of mitochondrial disease is now emerging OXPHOS dysfunction linked to defects of non-OXPHOS mitochondrial proteins. Those described so far are predominantly neurodegenerative diseases such as Parkinson’s and Huntington’s disease (Leonard and Schapira, 2000).

1.2.3.3 Mitochondrial translation (Protein-RNA interaction)
Since the mitochondrial genome is derived from once free-living bacteria, most of the bacterial genes have been removed to the cell nucleus. The remaining genes in mitochondria are believed to be the resources for any potential problems involved in
the interaction between mitochondrial and nuclear genomes (Alberts, 1994). Nuclear-encoded mitochondrial gene variants can become clinically relevant when joint with an incompatible mtDNA (Wallace and Chalkia, 2013). Interference with this mito-nuclear co evolution could disrupt mitochondrial function leading to devastating diseases, major evolutionary processes such as speciation and reproductive hurdles. Thus, the nuclear background derived from maternal and paternal genomes must be compatible with the maternally derived mitochondrial DNA to ensure embryonic survival. Therefore, couples may experience infertility due to mito-nuclear mismatch.

1.2.4 Mitochondrial disorders

Mitochondrial disorders are a highly heterogeneous group of diseases. All organ systems may be affected if there is a mitochondrial defect especially tissues with high energy demands such as the muscle, nerves and the brain. Mitochondrial disorders can be caused by mutations in mtDNA or nDNA (Figure 1-11). Therefore, such disorders show different modes of inheritance that may be autosomal dominant, autosomal recessive, x-linked and maternally inherited. Clinical diagnosis of mitochondrial diseases is based on clinical manifestations, enzymatic and physiologic analyses, tissue histochemical results, levels of biochemical analytes, and DNA analysis (Graham et al., 2012). Epidemiological studies confirm that pathogenic mtDNA mutations affect at least 1 in 5000 of the population with pathogenic alleles being present in more than one in 200 live births, and occurring de novo mutations in at least one in every 1000 births (Cree et al., 2009). Attempts to measure the true incidence of such diseases have been limited by disease
heterogeneity and the lack of definitive biomarkers (Manwaring et al., 2007, Chiaratti et al., 2011). Currently, more than 250 pathogenic mtDNA mutations have been identified (Tuppen et al., 2012) which could be encoding respiratory chain subunits or mitochondrial protein synthesis (e.g. tRNA or rRNA) and rearrangements (DiMauro and Schon, 2003). Mitochondrial disorders may occur at any age; mtDNA from late childhood to adulthood with the nDNA mutations present in childhood. This can be explained by the fact that the phenotypic expression of mtDNA mutations is determined by mutational load i.e.; proportion of mutant to normal mitochondrial DNA template in mitochondria which can change over time (section 1.2.4.1).

Figure 1–11: Graphical representation of the mitochondrial and nuclear encoded mitochondrial proteins involved in the OXPHOS system.

Nuclear encoded subunits are coloured in blue while the mitochondrial encoded subunits are coloured in red. Adapted from The New England Journal of Medicine, 348, DiMauro and Schon, Mitochondrial respiratory-chain diseases, pages 2656-2668. Copyright (2003), with permission from Massachusetts Medical Society. This diagram has been altered from its original version. Modifications: No. of nDNA encoded subunits from "~39"to"~45". According to the information derived from (Fassone and Rahman, 2012).
Mitochondrial disorders due to mutations in nDNA are numerous and have been reported in about 150 nDNA encoded factors (Landsverk et al., 2012). There is no effective way to treat and prevent mitochondrial disorders caused by mtDNA mutations, which indicates the need for methods to predict their transmission (Cree et al., 2009, Chiaratti et al., 2011).

1.2.4.1 Heteroplasmy & threshold effect

A large number of mtDNA copies are present in every nucleated cell. Normal individuals have nearly identical mtDNA copies. Individuals affected by defects in the mitochondrial genome can either have the mutation present in all copies of the genome (homoplasmic mutation) or have a mixture of both normal and mutant mtDNA molecules in most of their tissues. where the ratio of mutation to the wild type sequence (mutational load) exceeds a threshold before clinical symptoms may appear (heteroplasmic mutation) (Brown et al., 2006). Heteroplasmy can be present in different ratios between 0-99% which accounts for the phenotypic variability of mitochondrial disease (Tajima et al., 2007, Chiaratti et al., 2011). These thresholds vary with both mutation and tissue type. Based on the mutation load and its effect on the normal cellular activity, heteroplasmy can be developmentally toxic, ultimately lethal (Van Blerkom et al., 2008, Chiaratti et al., 2011) or harmless. The relationship of the percentage of a mtDNA mutation and clinical symptoms are often not straightforward and can change in time. Most pathogenic mutations leading to severe disease are heteroplasmic. Homoplasmic pathogenic mutations such as LHON mutations exist. Life-threatening severe homoplasmic mutations are rare, however, the heteroplasmic state of an individual, in principle, could have arisen either through the germline via maternal transmission or within the individual via a new somatic
mutation (Wallace and Chalika, 2013). Most of the studies that have reported heteroplasmy have not attempted to define which mechanism might be responsible.

1.2.5 Mitochondria in early mammalian development

During oogenesis, embryo development is sustained by maternally provided proteins, and transcripts that are stored in the oocyte. These factors are vital for early developmental events including embryonic genome activation and cell fate determination. Mitochondria not only provide energy to maintain cellular activity, but it also participates in a number of pathways which maintain amino acid metabolism, cellular homeostasis, apoptosis and signal transduction. These mitochondrial functions are crucial elements in early embryonic development, through the regulation of spindle organisation, chromosomal segregation, cell cycling, blastocyst hatching and other dynamic processes (Zhang et al., 2006, Johnson et al., 2007). In mammals, the loss of paternal mtDNA through ubiquitin tagging of sperm mitochondrial membrane and the proteolytic digestion upon fertilisation suggests that the paternal contribution to the zygote is predominantly through nuclear chromosomes only. After fertilisation, the oocyte cytoplasm is uniquely responsible for the remodelling of the incorporated paternal chromatin and plays a role in epigenetic modifications of the newly formed embryonic nucleus (Torres-Padilla et al., 2006), Yoshida et al. 2007) changing highly specialised germ cells into undifferentiated human embryonic stem cells which display a broad developmental potential (Yanicostas et al., 2011).

Recent data suggests that mitochondria can modulate embryonic differentiation into different cellular fates as well as stem cell pluripotency (Ramalho-Santos et al., 2009). Therefore, mitochondria are clearly important factors in early steps of
development and it is not surprising that a decline in oocyte quality has a profound impact on the developmental competence of the embryo and contributes to the high incidence of arrested embryos observed in ART clinics following IVF procedures (Jurisicova and Acton, 2004).

1.2.5.1 Mitochondrial inheritance and the bottleneck theory

In humans, homoplasmic mtDNA mutations are transmitted to all maternal offsprings. However, in some cases such as LHON, all offspring may inherit the mutation but only some may develop the disease phenotype. About, 50% of males, but only 10% of females, develop impaired vision which certainly indicates the influence of other factors in modulating the expression of the disease (Taylor and Turnbull, 2005). On the other hand, the inheritance of heteroplasmic mtDNA mutations is unpredictable. The rapid segregation of mammalian heteroplasmic mtDNA genotypes observed between generations, with a return to homoplasmy in some progeny has suggested the existence of mtDNA “bottleneck” at the early stages of development.

The mitochondrial bottleneck hypothesis postulates that only a small proportion of the total number of mitochondrial genomes are passed on from mother to offspring acting as a unique evolutionary mechanism in order to maintain mtDNA homoplasy and minimize heteroplasmy (Chiaratti et al., 2011). It may have also evolved to minimize problems arising from the need for nucleo-mitochondrial interactions. Thus, it has been suggested that the reduction in the mtDNA content and subsequent clonal proliferation of this small group of mtDNA molecules during oocyte growth could account for the rapid shifts in mtDNA genotype frequencies between generations (Hauswirth and Laipis, 1982, Chiaratti et al., 2011). However, the timing
at which this phenomenon occurs is still debatable and unclear. Most of the suggested theories suggest that mtDNA bottleneck most probably occurs during oogenesis since it has been proposed that mtDNA replication does not occur until after implantation. This was predicted when the disruption of mtDNA in mice embryos, by knocking out TFAM, did not affect the development of mutant embryos until after implantation (Larsson et al., 1998, Yanicostas et al., 2011). In humans, mtDNA synthesis measured by Bromodeoxyuridine (BrdU) incorporation was not observed until the late morula and blastocyst stages (Jacobs et al., 2006). Another proposed theory suggests the growth of the mitochondria population in the embryo is discontinuous with a growing burst during oogenesis followed by an arrest of replication during cleavage stages until mitochondrial replication resumes after gastrulation (Dumollard et al., 2006, St. John, 2007). This indicates that the embryonic mitochondria all originate from a restricted founder population present in the PGC, that is multiplied during oogenesis (Figure 1-12). Other proposed mechanisms suggest that the bottleneck occurs during postnatal folliculogenesis (Wilding et al., 2009, Tuppen et al., 2010). Another suggestion indicates that mtDNA replication barely occurs during the early stages of embryo development and that the pre-existing mtDNA molecules segregate among the cells of the blastocyst (Piko and Taylor, 1987, McConnell and Petrie, 2004, Thundathil et al., 2005, Cree et al., 2009, Chiaratti et al., 2011). The ICM at this stage contains a small number of cells, yet it will eventually give rise to the whole embryo. Thus, both the replication of mtDNA in the developing oocyte and mtDNA segregation into a few cells of the blastocyst could explain the selection of only a small number of mtDNAs transmitted to the next generation (Chiaratti et al., 2011).
Nevertheless, further investigations are needed to clarify exactly when and how this phenomenon occurs (Ramalho-Santos et al., 2009).

Figure 1–12: Proposed mechanisms for the mtDNA genetic bottleneck.

a. During cell division the unequal segregation of mutant and wildtype genotypes results in a variation in heteroplasmy levels, which leads to accelerated drift in heteroplasmy levels. This occurs mainly as a result of the dramatic reduction in mtDNA copy number immediately before expansion of the primordial germ cell (PGC) population (Cree et al., 2008). b. Variation in heteroplasmy in PGCs is generated through unequal segregation of the homoplasmic nucleoids (Cao et al., 2007). Multiple identical copies of mtDNA are contained within each nucleoid contains. c. In post-natal life, oocyte maturation results in variation in heteroplasmy which is generated through the replication of a subpopulation of mitochondrial genomes. Adapted from Mitochondrion, 11 (5), Carling et al., The implications of mitochondrial DNA copy number regulation during embryogenesis, pages 686-692, Copyright (2011), with permission from Elsevier.

1.2.6 Fertility and mitochondrial DNA

1.2.6.1 Mitochondrial number in oocytes and embryos

Quantification of mtDNA molecules in oocytes and embryos of mice as well humans have all indicated that the largest mtDNA content of the organism is found in the oocyte (St. John, 2007). This can be explained by the lack of mtDNA replication during the first cleavage stages of the embryo (Jacobs et al., 2006).
The mammalian oocyte contains $2 \times 10^5$ copies of the mitochondrial genomes in approximately $1 \times 10^5$ mitochondria (Cummins, 1998). Thus the number of mtDNA copies per mitochondrion in germ cells in thought to be as few as one or two molecules, compared to somatic cells (Poulton et al., 2010). Data from quantification of mtDNA in human cleavage stage embryos and mouse models suggests that the number of mitochondria remain stable during the first three days of preimplantation developmental stages (Steuerwald et al., 2000, Chan et al., 2005, May-Panloup et al., 2005). The overall amount of mtDNA must be distributed among all daughter cells during the cleavage divisions of embryos. Thus, by day six of development each embryonic cell should contain very few copies of mtDNA. Replication of mtDNA is not thought to be commenced until after the embryo has undergone the first cellular differentiation into ICM and trophectoderm (TE) and has become a blastocyst (St John et al., 2010, Eichenlaub-Ritter et al., 2011).

Based on the fact that mitochondrial replication does not occur before implantation, it has been suggested that mtDNA content in oocytes positively correlates with both fertilisation and embryo viability and may be used as a potential biomarker (Diez-Juan et al., 2015, Fragouli et al., 2015, Fragouli et al., 2017, Ravichandran et al., 2017, Wells et al., 2017). In humans (Santos et al., 2006) and other mammalian species (El Shourbagy et al., 2006, Hua et al., 2007) several reports suggest that there is a relationship between lower levels of oocyte mtDNA content and fertilisation outcome, connecting mtDNA content in unfertilised oocytes with fertilisation failure. However, these reports could not conclude if a lower number of mtDNA copies was the cause or was a result of fertilisation failure. Another group has reported a significant association of mitochondrial DNA haplogroup with the proportion of embryos that developed to the blastocyst stage in cattle (Tamassia et
This group did not find a significant difference in mtDNA number between the different haplogroups affecting embryo development.

Recently, there has been a huge debate on whether mtDNA number is associated with increased aneuploidy levels in embryos and if this has an effect on implantation potential. Some groups have reported the association between elevated mtDNA levels with aneuploidy and implantation failure in blastocysts (Fragouli et al., 2017, Ravichandran et al., 2017, Wells et al., 2017) while others have reported an association between elevated mtDNA levels with euploid cleavage stage embryos that failed to form a viable pregnancy (Diez-Juan et al., 2015). In contrast, other reports have shown that this relationship is not as direct as previously indicated (Tamassia et al., 2004, Victor et al., 2017), suggesting that neither the mtDNA copy number nor the ranges of mtDNA amounts are different between viable and nonviable embryos. Therefore, the usefulness of mtDNA template number as a biomarker for embryo viability is still debated and needs further confirmation by randomised controlled trials (Bayram et al., 2017).

1.2.6.2 Mitochondrial structure and activity in oocytes and embryos

The structure of mitochondria has been analysed using Electron microscopy showing that mitochondria in oocytes and early embryos have shorter and fewer cristae compared to differentiated and metabolically active cells (Trimarchi et al., 2000). Together with the low oxygen consumption, these observations have led to the hypothesis that mitochondria at the early stages of development have little capacity for respiratory activity and are believed to be immature (Trimarchi et al., 2000, Houghton and Leese, 2004, St. John, 2007).

A study using TFAM knockout mice revealed that decreasing the mitochondrial number in oocytes did not affect fertilisation or early development. On the other
hand, targeted deletion of genes essential to the replication, maintenance, and expression of mouse mtDNA resulted in embryonic lethality between E8.5 and E10.5 (Leonard and Schapira, 2000). Pharmacological approaches undertaken on mammalian and ascidian embryos, as well as the pattern of consuming energetic substrates seen in mammalian embryos suggest that mitochondrial activity is crucial for development activation and embryo survival. Even though a low level of metabolism has been indicated, the fact that oocyte has a very large number of mitochondria at the early stages of development indicates that the embryonic energy supply is definitely critical at this stage (St. John, 2007). Energy in the form of ATP is required to support important morphological processes, such as compaction, blastocoel formation and hatching as well as higher rates of macromolecular formation in the preimplantation development embryo development stages (Van Blerkom, 2009). Indeed, despite their undeveloped state, they are active in oxidative phosphorylation and are the primary source of ATP in the human oocyte and early embryo (Van Blerkom, 2011).

It has been suggested that failure of mitochondria to undergo structural changes to a state consistent with higher bioenergetics levels may lead to premature embryo arrest (Van Blerkom, 2011). However, embryos are contained in the hypoxic uterine lumen, which is not beneficial for ATP synthesis (Lonergan et al., 2007), indicating that ATP through the ETC is not the predominant pathway (Wang et al., 2009).

Structural analysis of mitochondria has shown that mitochondria undergo “stage specific” transformations during the preimplantation period in which they become elongated and develop a wide array of cristae that completely traverse a matrix of progressively lower electron density. In humans, it is not until the early blastocyst stage that such “structural transformation” typical of somatic cells occur; indicating its
high activity. This is logical since the blastocyst stage approaches the end of preimplantation embryogenesis and activities such as the maintenance of blastocyst cavity and hatching require a continuous and elevated energy supply in the trophectoderm (which will then form the placenta) compared to the ICM which is not involved in these processes (Van Blerkom, 2011). Research on mice blastocysts, has shown that the glucose metabolism of TE compared to the ICM was different. Isolated ICMs consumed almost three times more glucose per cell than TE suggesting that both the ICM and the TE can derive energy by aerobic glycolysis, but they differ in their capacity since 40% of the total lactate formed by a mouse blastocyst is due to the ICM and 60% to the TE (Hewitson and Leese, 1993). These findings are consistent with those of Barnett et al. (Barnett et al., 1996) who showed that active mitochondria were largely distributed in the TE rather than by the ICM in hamster blastocysts. Interestingly, it has also been reported that mouse blastocysts with low glycolysis and a high glucose consumption were the most viable after embryo transfer (Lane and Gardner, 1996).

Certainly, quantitative analysis of the amount of ATP produced in mouse blastocyst supports this interpretation. Houghton and colleagues showed that 80% of the net ATP content in the embryo was trophectoderm-derived, while the ICM was in a metabolically inactive state (Houghton, 2006). Differences in the number, structure and function of mitochondria between those two cell types were hypothesised to be the reason behind the differences in their metabolic capacities (Houghton, 2006). This may further explain why mitochondria in mouse and human trophectoderm are hyperpolarised (high IMMP), compared to the low potential of the mitochondria contained in the ICM. This work also suggests that the energy produced by
mitochondria in the trophectoderm is required to support (Na+, K+, ATPase) sodium pump (Houghton, 2006, Van Blerkom, 2011).

ATP demand and supply as well as cytoplasmic bioenergetic capacity is directly linked to embryonic competence (Van Blerkom, 2011). Developmental defects such as fertilisation failure (Reynier et al., 2001), chromosomal segregation disorders (Schon et al., 2000), abnormal fragmentation and cytokinesis, as well as arrested cell division may be associated with inadequate levels of mitochondrial ATP (Van Blerkom, 2011).

Several studies have been done in order to investigate the link of mitochondrial function with oocyte and embryo quality by measuring ATP at different stages of development. It has been shown, that the mean ATP content increases with development when the ATP in human unfertilised oocytes, zygotes and embryos was measured (Zhao and Li, 2012). ATP content within each embryo is suggested to be case specific (each couple may produce embryos with different ATP levels).

In humans, asymmetrical inheritance of mitochondria in blastomeres of two-cell embryos may result in severe adverse consequences for the embryo. These include cell lysis of blastomeres or slower cell division due to low ATP levels from the reduction in mitochondrial content. The developmental potential of mature human oocytes is also dependent on mitochondria and ATP content. Following uterine transfer, oocytes require a minimum of 2 pmol/oocyte in order for fertilised embryos to progress beyond the cleavage stage (Yanicostas et al., 2011) compared to viable somatic cells which will generally contain ~1 picogram (10^{-12} gram) or ~2 femtomoles (2 \times 10^{-15} moles) of ATP per cell (Lundin,1986, Kricka and Carter, 1982). Many reports have shown that ROS plays an important role in female reproduction (reviewed in (Agarwal et al., 2005). Both mammalian embryo and oocyte are,
however, sensitive to oxidative stress (Liu et al., 2000) and if physiological levels of ROS are beneficial, embryo development and oocyte maturation can be disrupted by increased levels of ROS (Harvey et al., 2002), which can promote fragmentation in the embryo (Johnson and Nasr-Esfahani, 1994, Yang et al., 1998). Indeed, apoptosis of the oocyte and early embryo can be induced by oxidative stress (Liu et al., 2000). Embryo development, implantation failure, miscarriages as well as the pre-and postnatal decline in germ cells, seem to be affected by mitochondria-dependent apoptosis.

Oocyte survival in mammals is determined by the balance between several anti- and pro-apoptotic members of the Bcl-2 family expressed in the oocyte (reviewed in (Liu et al., 2000, Jurisicova and Acton, 2004). Furthermore, it has been shown that the acute increase in mitochondrial ROS production in mouse is linked to the hyper stimulation of follicles (Chao et al., 2005, Van Blerkom, 2011).

Mitochondrial DNA copy number and the resulting ATP content in the oocyte are not directly related. Raised levels of mtDNA are most likely the result of a compensating mechanism to normalize ATP generation resulting from organelles with reduced function. Mitochondria in the oocytes of older mice and hamsters have been shown to produce less ATP and generate higher levels of ROS, which indicates less support towards dynamic processes, such as preimplantation development (Cummins, 2000). In humans, if this scenario occurs, an increase in the number of mitochondria may be necessary in the embryos or oocytes of older women, in order to maintain the required ATP levels for healthy development. Experimental data in human trophoblastic cells support the concept that energetic stress, induced pharmacologically, increases mitochondrial number and produces morphologic mitochondrial changes. Therefore, energetic stress in an embryo can
induce early mitochondrial maturation, with increased size and mtDNA copy numbers to compensate for a low energy reserve. This suggests that the embryo depends largely on energy accumulated during oocyte maturation; therefore, in some cases an increase in mtDNA copy number may be a consequence rather than the root cause of the oocyte's energy status. This is supported by the observation that changes in mitochondria are associated with mitochondrial hyper-proliferation and that the pathogenic consequence of “mitochondrial distress” is also an evident increase in mitochondrial proliferation (Monnot et al., 2013).

Consequently, an increase in mtDNA copy number in early embryos would be indicative of metabolic stress and that the embryo is trying to compensate for this stress by increasing its mtDNA content. This energetic stress could be related to intrinsic factors during oocyte maturation or could be in response to impaired respiratory capacities due to mtDNA mutations or mitochondrial mismatch. A very interesting study has shown that the m.3243A>G MELAS mutation results in a gradual increase in mtDNA amount from the germinal vesicle to the blastocyst stage (Monnot et al., 2013). Taken together, mature oocytes and early embryos maintain an overall low-level (i.e. ‘quiet’) metabolism, in order to minimize oxidative stress, while required ATP is generated to fulfil cellular functions (Leese, 2002, Leese et al., 2007, Ramalho-Santos et al., 2009). The effect of oxidative stress is illustrated in the Figure 1-13.

Thus, these observations suggest that the embryo largely depends on the energy accumulation during oocyte maturation.
Figure 1–13: Schematic diagram of the effect of oxidative stress on fertility.

The impact of oxidative stress on physiological balance of ROS over the mtDNA is represented in the image above. Consequently, the ATP production by OXPHOS is affected by this stress. On the right, the list of risks and concerns of the impact of oxidative stress on the embryos, gametes, pre-natal and post-natal stage of development is shown. Reproduced from The International Journal of Biochemistry & Cell Biology, 55, Benkhalifa, Mitochondria: Participation to infertility as source of energy and cause of senescence, pages 60-64, Copyright (2014), with permission from Elsevier.

1.2.6.3 Polymorphism and mutations in oocytes

Mitochondrial DNA mutations in oocytes have been described as de novo events which could potentially explain the incidence of new disease cases. They can also occur as a part of the transmission of pathogenic familial mutations which is important for carriers in families for the prediction of the recurrence risk of mtDNA disease. Deletion in mtDNA, although usually occurring in very low mutational loads, has been reported in 40-60% of unfertilised oocytes. They have also been reported in oocytes that failed to develop into mature metaphase II (Chen et al., 1995, Reynier et al., 1998, Barritt et al., 1999, Hsieh et al., 2002, Jacobs et al., 2006).

The entire mtDNA in oocytes was screened for mainly heteroplasmic point mutations and it was found that over 25% of the oocytes had point mutations. The mutation percentages varied from very low levels (<1%) to high levels (>50%) with most oocytes containing low level of mutation (<30%) (Jacobs et al., 2006). It has been assumed that at least 10% of the point mutations in the mtDNA will be pathogenic suggesting that more than 5% of the oocytes carry a possible pathogenic
mutation in the mtDNA. Mostly, these mutations are present in very low levels. De novo mutations may have a direct phenotypic effect if the mutational load rises above the threshold, (de Coo et al., 1996, Degoul et al., 1997, Maassen et al., 2002, Thorburn, 2004). Low mutation levels may get lost by cell division, but also fixed during life by random genetic drift, which has been observed in rapidly dividing cancer cells (Carew and Huang, 2002) in colonic crypt cells (Taylor et al., 2003). This also indicates that diseases such as Alzheimer and Parkinson’s, which are associated with mtDNA mutations, might be caused by the accumulation of very low levels of mtDNA mutations during life (Chinnery et al., 2002, Coskun et al., 2004, Jacobs et al., 2006).

A number of mitochondrial disorders have been linked to infertility. Some mtDNA defects that affect metabolic capacity have been found to be associated with oocyte maturation, fertilisation and early embryonic development. Most of these disorders are caused by variants that affect either mtDNA or the mitochondrial translation pathway. However, not all variants in this protein synthesis pathway cause infertility. Thus, many of the reported cases of infertility associated with mitochondrial syndromes have not been genetically confirmed (Demain et al., 2017). The occurrence of affected individuals indicates that the presence of mutant forms per se is not a barrier for fertilisation or gestation to term (Inoue et al., 2000, Moilanen and Majamaa, 2001), but whether certain threshold levels of heteroplasmy could contribute to infertility or recurrent pre-and post-implantation loss is unknown (Van Blerkom, 2011) and whether the mismatch between certain nuclear and mitochondrial haplotypes may have an effect on development remains to be determined. In light of this evidence, it is worth investigating the role of mitochondria at the very early stages of oogenesis and preimplantation embryogenesis for early
clinical diagnosis and research purposes with regards to infertility.

1.2.6.3.1 Mitochondrial haplotypes

Mitochondrial variants which have arisen from within a germ cell must have been heteroplasmic, having survived the mtDNA genetic bottleneck over several generations they ultimately become fixed (homoplasmic) within a maternal line. The lack of intermolecular recombination and the uniparental inheritance of mtDNA means that these variants have remained restricted to specific ethnic groups (Chinnery et al., 2015). Recent data have shown that the mitochondrial haplogroup may influence the amount of cellular ATP, whose variants mediate cell growth by the production of energy and cell signalling for the major molecular pathways, such as those regulating chromosome segregation (Kenney et al., 2013; Gianaroli et al., 2014). Haplogroups have physiological consequences and it has been reported that mitochondrial haplogroups U and T are associated with reduced sperm motility (Montiel-Sosa et al., 2006; Ruiz-Pesini et al., 2000). Haplogroups have also been shown to influence ovarian ageing (Torroni et al., 1996) although some haplogroups appear to be protective against ovarian reserve decline (May-Panloup, 2016). It was recently found that mtDNA haplotypes define gene expression patterns in mouse embryonic stem cells (Kelly et al., 2013), so certainly mtDNA-nDNA interaction does primarily depend on the mtDNA haplotype.

1.2.6.4 Mismatch effect on fertility and reproduction

Several studies investigating whether incompatibilities between nDNA and mtDNA encoded proteins, extensively interacting in oxidative phosphorylation, could contribute to interspecies reproductive barriers in drosophila (Camus et al., 2012,
Camus et al., 2015) and other species (Barreto et al., 2015). Mitalipov and colleagues show that reciprocal mtDNA replacement in zygotes between two mouse strains (B6 and PWD) results in post-implantation embryonic lethality, suggesting that mtDNA sequence divergence between mammalian species contributes to a reproductive barrier (Ma et al., 2016). On the other hand, (Latorre-Pellicer et al., 2016) has shown that the high sequence variability of mtDNA is of little consequence when mitochondrial replacement was performed between two different mice strains. They have shown that the replacement of mtDNA from a mouse strain called NZB to the nDNA background of another strain of mice C57BL/6 had a beneficial effect in mitochondrial function, insulin signalling and longevity.

These observations demonstrate that the substitution with a different wild-type mtDNA variant is sufficient to promote a cellular adaptive response in some species and mitochondrial function while showing a different effect in others. This highlights the influence of mtDNA-nDNA interplay, orchestrated by the complex network of mitochondrial stress response pathways in which proteostasis, ROS and mitochondrial Unfolded Protein Response (mtUPR) signals lead to adaptive responses that impact on the organism’s metabolic performance and ageing. Therefore, investigating this effect in humans is essential to better understand fertility and it will provide insights for the emerging new field of mitochondrial replacement therapy.

1.2.6.5 Nuclear transfer technologies

Nuclear transfer has been proposed as a novel approach to minimize the transmission of mutant mtDNA from a carrier mother to her child at the gamete (spindle transfer) or zygote level (pronuclear transfer) (Craven et al., 2010). However, the safety of this technique had to be tested as it involves the substitution
of two nuclear genomes that will interact with a different mitochondrial background. Experimental nuclear transfer in both animals and humans has been reported (Zhang et al., 1999; Tachibana et al., 2009; Liu et al., 1999, 2003; Zhang et al., 2016; (Tachibana et al., 2013); Zhang and Liu, 2015; Hyslop et al., 2016). Recent studies show that the spindle and pronuclear transfer techniques can be used to carry out nuclear transfer for mitochondrial replacement therapy: meta-phase II (MII) (Tachibana et al., 2009, Craven et al., 2010). In 2015, the UK parliament approved a change in the regulation of IVF techniques, allowing “Mitochondrial replacement therapy” to become a reproductive choice for women at risk of transmitting mitochondrial disease to their children. Therefore, understanding the interaction between mitochondria and the nucleus has never been more important.

The first report of a healthy born baby following the spindle transfer technology into the cytoplasm of enucleated donor oocytes was reported recently (Zhang et al., 2017). Interestingly, the mother was a carrier of Leigh syndrome (mtDNA mutation 8993T > G), suffering from multiple undiagnosed pregnancy losses and deaths of offspring as a result of this disease. Changing the mitochondrial background for the nuclear genome in the embryo not only treated the mitochondrial disorder but also treated the repeated miscarriage problem.

1.2.6.5.1 Cytoplasmic transfer
This form of mitochondrial transfer entails the transfer of some ooplasm, containing wild-type mitochondria, from a healthy donor oocyte into an oocyte from a female with reduced fertility in order to augment the number of mitochondria in the oocyte. Cytoplasmic transfer, using a donated part of the ooplasm, has been reported in patients experiencing repeated failure in embryonic development. Donation of ooplasm is thought to give mitochondria and other energy producing factors that may
support developmental competency (Van Blerkom et al., 1998; Huang et al., 1999; Jacobs et al., 2006). Despite the significant improvement in embryo development from ooplasmic transfer for certain patients experiencing infertility problems, some studies show that ooplasmic transfer resulted in an increased incidence of aneuploidy and birth defects (Barritt et al., 2001, Brown et al., 2006, Jacobs et al., 2006). Therefore, the clinical use of this technique for women at risk of transmitting mtDNA disease to their offspring was banned.

1.2.6.6 Mitochondrial aging

There is variation in oocyte mutational load over time, which could be associated with reproductive senescing. For example, the point mutation m.414T>G is highly prevalent in the oocytes of women aged >37 years (40%) compared to the oocytes of women aged <37 years (4%) (Barritt et al., 2000). A reduced number of mitochondria in oocytes was also noted by Reynier and colleagues from patients who had experienced fertilisation failure owing to unknown causes (Reynier et al., 2001). Fewer mitochondria are also found in ageing oocytes (de Bruin et al., 2004) meaning that it is not necessarily just energy production and ATP content, but specifically, the number of mitochondria which are important during embryo development. Fertility may also be affected through acquired mtDNA mutations which in turn affect mtDNA replication (Jacobs et al., 2006). A relationship between oocyte quality/fertility and mtDNA copy number has also been observed (Yesodi et al., 2002, Santos et al., 2006), as unfertilised oocytes present a lower mtDNA copy number than fertilised oocytes (Santos et al., 2006). It is perhaps, therefore not dysfunction in OXPHOS that primarily contributes to reduced fertility, but rather reduced copy number of mtDNA/mitochondria that leads to the dysfunction observed.
in OXPHOS and to the subsequent poor-quality oocytes or reduced fertility (Inoue et al., 2000, Jacobs et al., 2006, Ramalho-Santos et al., 2009).

The female ability to produce an offspring in humans reaches its peak at age 30 while it declines considerably as they reach menopausal age. This is due to many factors but in general subfertility in older females is related to the poor quality of aging oocyte, reflecting morphological, chromosomal and functional abnormalities (Kujjo and Perez, 2012). Oocyte/embryo quality decreases with the increasing maternal age affecting embryo development which is strictly correlated with the activity of mitochondria in oocytes, characterised by MMP. Compared to younger females, oocytes derived from aged women often present abnormal chromosomal alignment, aberrant spindle formation and a high occurrence of aneuploidy (reviewed in (Battaglia et al., 1996, Eichenlaub-Ritter et al., 2004) Studies have shown that maternal age has an effect on oocyte mitochondrial activity. A decline in MMP was observed in oocytes and preimplantation embryos from women of advanced maternal age (Wilding et al., 2001). Also, an accumulation of mtDNA point mutations (Barritt et al., 2000) and higher levels of mtDNA deletions (Keefe et al., 1995, Van Blerkom, 2011) and apoptosis (Wang et al., 2009) have also been reported. All these dysfunctions may cause a high level of developmental retardation and arrest of preimplantation embryos (Wang et al., 2009). Mitochondrial swelling and disruption of cristae appear to be a common feature of the oocytes of women of advanced reproductive age ≥ 40 years of age (Muller-Hocker et al., 1996, Van Blerkom, 2011) suggesting that this observation could be due to abnormalities resulting from the inadequate capacity to generate sufficient ATP levels to support these events (Gaulden, 1992, Ramalho-Santos et al., 2009). However, when ATP level was measured in normal meiosis II oocytes from older women, no significant differences
were found compared to their younger counterparts. This could be due to the very small number of women in this age group which undergo infertility treatment with their own oocytes (Ramalho-Santos et al., 2009, Van Blerkom, 2011).

A higher volume of mitochondrial fraction is seen in oocytes of older women, indicating the role of a potential compensatory mechanism (Muller-Hocker et al., 1996). Similarly, fragmentation in postovulatory aged oocytes seems to be associated with mitochondrial dysfunction. Indeed, a lower uptake of pyruvate is noted in aged oocytes (Hardy et al., 1989). Furthermore, [Ca2+] oscillations in these oocytes fail to trigger the production of ATP, inducing apoptosis instead (Gordo et al., 2002).

Damages in mitochondria caused by oxidative stress leading to apoptosis, appears to be the primary cause of fertility decline in both postovulatory and maternally aged oocytes (Perez et al., 1999, Fissore et al., 2002, Ramalho-Santos et al., 2009). The risk of infertility has dramatically increased in the last 30 years and although ART has been widely used to overcome this, it is limited to its ability to tackle general ageing. Female fertility is, according to Hassold and Hunt, influenced by a multitude of complex events which can occur at each or any of the three stages of oocyte development: when meiosis is initiated in the fetal gonad, when primordial follicles form during late gestation, during oocyte growth and maturation in the perinatal period and, finally, in the adult. It appears, based on this, that a complex and dynamic interplay between both the nuclear and cytoplasmic components are involved in the age-related changes affecting oocyte quality (Kujjo and Perez, 2012).
1.2.6.7 Other factors affecting mitochondrial function

Chronic early foetal demise has long been thought to be associated with cigarette smoking in females since mutagenic compounds commonly detected in tobacco smoke have been identified in the follicular fluid and oocytes. Levels of ROS increase significantly within fully developed antral follicles of smokers, indicating a possible mechanism for acute mtDNA damage (Van Blerkom, 2011). However, a direct association between mtDNA damage in the oocyte and tobacco smoking has not been confirmed yet (Van Blerkom, 2011). A large number of studies have shown an association between obesity and many health problems such as type 2 diabetes, liver disease, hypertension, cancer, psychological problems as well as infertility in females. However, a growing number of research data has indicated that paternal obesity may also be of concern. This significant association has been confirmed by Binder when paternal diet induced obesity in mice resulted in the reduction of the mitochondrial membrane potential, high pyruvate uptake, delayed embryo development (as early as syngamy) and differences in the ratio of cell allocation to TE and ICM lineages at the blastocyst stage (Binder et al., 2012). Since embryos derived from the sperm of obese males (but normal females) affected mitochondrial activity and kinetics in the developing embryo, it is essential to investigate the paternal effect of nuclear encoded mitochondrial genes on the mitochondrial functional capacity in embryos at the early stages of development to generate a high-quality embryo capable of survival. These effects could be due to epigenetic changes and / or DNA mutations.
1.3 Repeated miscarriage and repeated implantation failure

Pregnancy loss is one of the most common obstetric complications. Approximately 2-5% of women at reproductive age experience recurrent miscarriages (RM). RMs are defined when three or more repeated pregnancy losses (RPL) occur (Ford and Schust, 2009). These losses could result from many factors, such as infectious, hormonal, metabolic and immune changes. The exact cause of the majority of RM cases remains unknown since some cases tend to occur even if the foetus has a normal chromosome composition.

The failure of embryos to implant following the transfer of good quality embryos following IVF, or with ICSI or frozen embryo transfer cycles, is a complication that can be associated with RM. Recurrent implantation failures (RIF)s are described when 1-2 high-quality embryos transfers fail to implant following at least three consecutive IVF attempts (Simon and Laufer, 2012). RIF causing mechanisms include decreased endometrial receptivity which could be related to endometriosis, thrombophilias and/or endometriosis (Simon and Laufer, 2012).

The prevalence of pregnancy loss is higher amongst siblings with unexplained RM in comparison to the general population, suggesting a possible genetic cause (Nybo Andersen et al., 2000, Kolte et al., 2011). Understanding the causes of RM is important for the patients and scientific community in order to find targeted therapies (Grimstad and Krieg, 2016).

1.3.1 Genetic factors

Understanding the genetic aetiology of RM has been an area of intense investigation. Studies that examine the genetic causes of RPL aim to identify genetic markers (DNA/ RNA), which have a direct predictive value in order to determine the
pathways involved in implantation and pregnancy. In those studies, three different genetic profiles can be evaluated: the mother, the father, the foetus/placenta (along with the epigenomes of the foetus and placenta). Nearly half of the pregnancy losses are a result of chromosomal anomalies, which can be of parental origin, or arise de novo in the embryo (Aboulghar et al., 2001).

Well known parental abnormalities causing RPL are translocations, both balanced chromosomal rearrangements and Robertsonian. Studies on miscarried product of conception (POC) estimate that about 25%–39% have chromosomal imbalances (Carp et al., 2006, Stephenson and Sierra, 2006). Data from embryo biopsies report that nearly a quarter are chromosomally balanced, confirming the high level of chromosomal abnormalities in these embryos (De Rycke et al., 2015). Overall, despite the increased risk of pregnancy loss, the majority of couples with balanced translocations end up with healthy pregnancy outcomes (Stephenson and Sierra, 2006, El Hachem et al., 2017).

1.3.1.1 Aneuploidy
Over 90% of the chromosomal abnormalities seen in miscarriages are numerical (polyploidy, aneuploidy). There is a positive association between maternal age and aneuploidy. This could be due to the degradation of cell cycle controls that manage the formation and function of the spindle cell (Fritz and Speroff, 2011). The most well-known example is the association of maternal age (>35 years) and trisomy 16 (Marquard et al., 2010). However, these changes are de novo (Carvalho et al., 2010) therefore; the risk of aneuploidy in later pregnancies is low. Frequent miscarriages are less likely to be associated to chromosomal abnormalities. Therefore, women with RM have a lower incidence of embryonic chromosomal abnormalities in comparison to those with sporadic miscarriages. Other genetic causes of RPL may
involve varying allelic expression, with many studies evaluating different genes. Some of the genes differentially expressed in PRL patients have immunologic purposes (Kosova et al., 2015).

1.3.2 Major histocompatibility complex (MHC) and the human leucocyte antigen (HLA)

Several studies have highlighted the immune response to explain idiopathic RM and RIF suggesting that disturbance of the maternal immune tolerance to the allogeneic foetus could contribute to these observations. Despite the immune-privileged state of the foetus existence, at some point-appropriate inflammatory responses must occur for human implantation to happen. Trophoblastic implantation can either be suppressed or promoted by components of the immune system (Park and Yang, 2011, Haller-Kikkatalo et al., 2014). The entire state of pregnancy is a pro-inflammatory state with most markers being raised for the duration of pregnancy (Laird et al., 2003). These changes are thought to be mediated by estrogen (pro-inflammatory) and progesterone (immunosuppressive). The hormonal balance is a result of the pro-inflammatory and the immunosuppressive state typically resulting in a balance. It has been suggested that HLA region genes have a role in both pregnancy outcome and mate choice. Hence, selective action of these genes may occur pre-conceptionally as well as during pregnancy (Lessey and Young, 2014).

The human leukocyte antigen (HLA) alleles are on chromosome 6. While class I HLAs present peptides from within the cell, the Class II HLA alleles present antigens outside of the cell. Studies focusing on the associations between maternal HLA alleles (HLA class I [HLA-C2], HLA class II [DRB1, DQA1, DQB1 and DBP], HLA-E, HLA-G), as well as HLA sharing between couples with regards to unexplained RMs have reported inconsistent findings. Research into the HLA-Class II loci for maternal-
foetal histocompatibility has shown that compared with fertile control couples, significantly more couples with RPL shared two HLA DQA1 alleles. This led to the suggestion that HLA-DQA1 compatible foetuses might be miscarried early in pregnancy, before the time when fetal tissue can be recovered for genetic studies (Ober et al., 1993). Other studies have reported higher rates of RPL seen in couples with HLA compatibility (Beydoun and Saftlas, 2005). One study found increased frequencies of identical HLA-A and HLA-B alleles in families with higher rates of RPL (Christiansen et al., 1989).

On the other hand, there are other studies which show that HLA sharing has no effect on pregnancy outcome (Balasch et al., 1989, Aruna et al., 2010). Therefore, there is insufficient evidence to recommend the management of RM based on HLA phenotypes and whether the immune responses are in response or a cause of the pregnancy loss is still unknown (Grimstad and Krieg, 2016, El Hachem et al., 2017).

1.4 Association studies

For many years, linkage and candidate gene studies were widely used in attempting to decipher the genetics underlying many diseases. Linkage tests widely spaced markers across genomes of affected individuals, within families, to identify if they segregate with the disease phenotype (Botstein et al., 1980). Although most of the common human genetic variants are hypothesised to be neutral, it is believed that each complex phenotype is influenced by a mixture of genetic variants and environmental factors (Frazer et al., 2009, Manolio et al., 2009). This clearly shows the importance of clarifying the role of genetics variants when searching for an association between a phenotype and a genetic factor across the genome.
1.4.1 Genome Wide Association Studies (GWAS)

Genetic variants are present in different forms, including single-nucleotide polymorphisms (SNPs), block substitutions, insertion-deletion variants, inversion variants and copy number variants. Advances in large-scale genomic projects such as the HapMap project, the human genome sequence, SNP discovery efforts, as well as new genotyping technology has led to a better understanding of genomic variations (Hirschhorn and Daly, 2005). Thus, large-scale studies have been devised to detect variants that differ between individuals, to examine the effects of human genetic variation on phenotypic diversity (Frazer et al., 2009). The most recent and widely used approach to investigate the association between a large number of markers and a particular phenotype across the genome are Genome-Wide Association Studies (GWAS) (Hardy and Singleton, 2009).

Researchers have begun to use GWAS to scan human genetic samples for signals of individual differences (Baker, 2010) with the aim to identify “common disease-associated variants”. By mapping those susceptibility effects, GWAS highlights the most frequently occurring genetic variants between cases comparing them to controls. Using rigorous statistical methods, researchers can further investigate whether the association between these variants and the common complex trait are due to mere chance or result from a consequence change in biology (McCarthy et al., 2008, Baker, 2010, Witte, 2010). Different from disease-causing mutations, common functional variants could either be mildly deleterious or confer adaptive properties, or otherwise, they would have been removed due to negative selection. However, common mtDNA genetic variants causing functional changes could have survived a long evolutionary time because of functional compensation “co-evolution”.
GWAS mostly relies on indirect associations for the detection of causal variants on a large scale (Palmer and Cardon, 2005). Indirect associations measure the link between a marker (which is usually a ‘tagging’ SNP associated with the actual causal allele) with the phenotype, as a result of linkage disequilibrium (LD). This is illustrated in Figure 1-14.

![Indirect association](image)

*Figure 1-14: Graphical representation of an indirect association between a disease phenotype and marker locus.*

Since the causal locus is unobserved, the two direct associations cannot be easily identified. However, it may be possible to detect the indirect association between the marker locus and disease phenotype if the marker locus and causal locus are in linkage disequilibrium. Reproduced from Nature Reviews Genetics, 7, Balding, A tutorial on statistical methods for population association studies, pages 781-791, Copyright (2006), with permission from Springer Nature.

GWAS and other population studies have failed to identify any definitive relationships with variants, in genes encoding mitochondrial proteins. The unique inheritance pattern of mitochondria makes mtDNA variations susceptible to having geographical structure. Such a structure may have a dramatic impact on mitochondrial association studies, particularly in heterogeneous populations. Therefore, mtDNA sequencing in combination with whole-exome sequencing is currently undertaken (Vento and Pappa, 2013).
1.4.2 SNP-genotyping technology

Genotyping by GWAS is established through commercial ‘SNP chips’ or arrays which contain about 300,000 to 500,000 common SNPs to pinpoint differences in allele frequencies between cases and controls. These studies have presented convincing and significant statistical associations for over 300 different loci in the genome for more than 80 phenotypes. The main advantage of GWAS is that they enhanced the understanding of the molecular pathways that cause many human diseases, and also shed light on many genes and genomic loci that have not been previously known to be associated with the disease of interest (Frazer et al., 2009, Manolio et al., 2009). Most SNP chips include a set of Ancestry Informative Markers (AIM) that show high allele frequency divergence between different geographically distant or ancestral populations. These genetic markers are especially useful in inferring the likely ancestral origin of an individual or estimating the sharing of ancestry components in admixed populations or individuals. The study of AIMs is of great interest in clinical genetics research, particularly to detect and correct for population substructure effects in case-control association studies, but also in population and forensic genetics studies (Pereira et al., 2012) along with mitochondrial haplogroups.

1.4.2.1 Haplotype analysis

The combination of markers on a single chromosome is called a haplotype. Haplotype structures may provide critical information on human evolutionary history and aid in the identification of genetic variants underlying various human traits (Zhao et al., 2003). The disease-associated variants are localised in LD blocks overlapping known genes. To date, a portion of variants in LD blocks are known to involve protein-coding variation. Therefore, causal polymorphisms may exist in regulatory
regions affecting gene expression. Studies of gene expression variation within and between human populations can examine the effects of polymorphisms in these cis-regulatory sequences (Verlaan et al., 2009). Functional SNPs affecting gene expression may be present in the surrounding non-conserved sequences and are therefore not easily identified. Identification of cis-acting SNPs can provide insight on how noncoding variants can alter disease risk. Moreover, it can also provide a better understanding of how gene expression is regulated. Although GWAS studies have been widely used, they are limited by the observation that genetic variants detected by most of these studies may not provide a complete explanation of the disease heritability, may not be causal for disease, or may have limited impact on public health (Witte, 2010, Cortes and Brown, 2011). These issues have prompted researchers to design custom content genotyping assays including hundreds and thousands of variable points of interest across the human genome with a relatively low cost and small amounts of DNA.

1.4.2.2 Infinium CoreExome-24 v1.2 BeadChip.

The Infinium CoreExome-24 v1.2 BeadChip is high density, 24-sample BeadChip. It was developed in collaboration with several leading research institutions to deliver affordable, high quality, genome-wide information across diverse world populations. The chip includes all the tag SNPs found on the Infinium Core-24 BeadChip, plus over 240,000 markers from the Infinium HumanExome BeadChip. In addition to performing cost effective large-scale genotyping studies, this chip can be used to obtain baseline sample data sets for various downstream applications quickly and easily. These applications include common variants, mtDNA, AlMs, sex confirmation, loss of-variant, and insertion/deletion (indel) detection studies.
1.5 Hypothesis

Based on the previous reports indicating the importance of mitochondria in the early stages of embryo development, and since a considerable number of nuclear genes directly affect mitochondrial function, we aimed to sequence the mitochondrial and nuclear genomes of couples and their embryos to see if there are any groups or populations that struggle to conceive based on the incompatibilities between their mitochondrial and nuclear genome background affecting mitochondrial function. We hypothesise that this could be due to mismatches between nuclear-mitochondrial interacting proteins, within the mitochondrial OXPHOS system leading to apoptosis, which may affect embryo development. Such results could be used for the management of unexplained infertility since treatment would be focused on targeting mitochondrial function, but above all, would allow us to gain a deeper understanding of the biological underpinnings of embryo survival.

1.6 Aims

1.6.1 Aim 1: To assess the effect of mtDNA copy number and ATP with preimplantation embryo quality by:

- Examining the relationship between mtDNA copy number and embryo quality (morphology and aneuploidy) as a predictor for successful implantation. This was performed by testing several real-time PCR techniques comparing the developed blastocysts with the arrested embryos at various stages. In order to elucidate if this effect is only caused by the difference in mtDNA number, the aneuploidy status of some of the embryos was correlated with mtDNA number and embryo morphology.
The ATP status of frozen/thawed embryos, following PGS/PGD where the aneuploidy status was already known, was determined. The ATP analysis was performed by bioluminescent somatic cell assay, and a comparison based on the number of chromosomal abnormalities was performed.

1.6.2 Aim 2: To investigate the effect of parental ethnicity on parental fertility and embryo quality

Ethnicity was determined by the mitochondrial genome to find if there is an association between specific mtDNA haplogroups and fertilisation outcome. Sequencing of the whole mitochondrial genome was performed by long-range PCR using the MiSeq Illumina platform. Bioinformatics analysis was performed to link the mitochondrial haplotypes with fertilisation and blastocyst rates. Genotyping Ancestry informative markers (AIM) was performed to confirm the ethnicity determined from the nuclear genome. Since the embryos tested already had the aneuploidy data determined by a-CGH, a comparison of aneuploidy analysis between the Human Karyomap chip and a-CGH for the same embryo was performed.

1.6.3 Aim 3: To examine the effect of mtDNA/nDNA mismatch on human embryos quality and fertility

A selection of candidate nuclear genes encoding proteins essential for mitochondrial function was sequenced on Illumina platform using the Sure-Select QXT (Agilent, UK). Bioinformatics analysis was performed to find if a mismatch caused by SNPs within the parental mitochondrial haplogroups and nuclear genes in the embryos may affect fertility. Due to the difficulty in assessing the mitochondrial depth of read
in embryos, the relationship of mtDNA number to nuclear background was not evaluated.
Chapter 2 Materials & Methods

Materials and methods outline

The methods section has been organised into three main sections; sample collection and patient information, sample preparation and sample processing & analysis.

1. Sample collection and patient information: This is a detailed explanation of the samples used and collected for each study aim, including the patient cohort, the type of sample and the criteria used for grading the embryonic samples.

2. Sample preparation: This section describes the methods used for the preparation of samples, such as the single cell isolation procedure, tubing, lysis and WGA.

3. Sample processing & analysis
   3.1 General methods: describe all the general methods used for initial processing of samples including; primer design, PCR, gel electrophoresis, Qubit, NanoDrop, assessment of DNA library quality and quantity by the Bioanalyzer or TapeStation and aneuploidy screening.
   3.2 Specific methods: methods used for specific aims are listed in this section and described in chapters 3, 4 and chapter 5.

All reagents were purchased from VWR (International) and were of Analar Quality including; ethanol 99.7-100% v/v and salts used for the preparation of buffers, such as Tris (hydroxymethyl) methylamine (TRIS) and sodium chloride (NaCl), unless otherwise stated. Molecular biology grade reagents were supplied by Sigma (UK) including the following list of reagents: adenosine 5’-triphosphate disodium salt hydrate (ATP), ethylenediaminetetra-acetic acid anhydrous 99% (EDTA), bovine
serum albumin (BSA), phosphate buffered saline (PBS), polyvinyl alcohol (PVA), sodium acetate (3M solution), sodium dodecyl sulphate (SDS) and agarose type I. Of these reagents ATP was stored at -20°C, lyophilised BSA was stored at 4°C and the rest at RT. Enzymes such as; Shrimp alkaline phosphate (SAP), Exo I were obtained from New England Bio labs (UK) and and stored at -20°C while Proteinase K was supplied by Roche (UK) and stored at 4°C.

An overview of the work flow to achieve the project aims is represented in the following diagram (Figure 2-1) and table (Table 2-2).

Samples were collected from 3 groups of patients; a fertile group and repeated miscarriage (RM) / repeated implantation failure (RIF) group and a PGS group. The fertile group included couples undergoing PGD for single gene disorders. The PGS group included couples who underwent PGS for aneuploidy screening.

A total of 114 embryos which were diagnosed as affected following PGD or unsuitable for transfer were collected. These embryos were tubed then allocated for analysis for each research aim. In total, 76 embryonic samples were used for the mtDNA quantification and 56 embryos were whole genome amplified (WGA) then subjected to several other methods of analysis including: mtDNA sequencing, nDNA sequencing, SNP genotyping and aneuploidy screening. Results from the PGD fertile group of patients were compared to the RM/RIF group. Finally, another set of frozen/thawed embryonic samples (n=29) were collected from couples who underwent PGS for aneuploidy screening. The ATP levels were tested in these embryos and compared to another set of frozen thawed embryos (n=15) from the PGD group.
Samples were collected from; a fertile group and a RM/RIF group and a PGS group. The fertile group included couples undergoing PGD for single gene disorders. 114 spare embryos were collected from the PGD group. In total, 76 embryonic samples were lysed (including single cells and whole tubed embryos) for mtDNA quantification analysis while the other 56 embryos were WGA to be used for mtDNA sequencing, nDNA sequencing, SNP analysis & aneuploidy screening. Results from the PGD fertile group were compared to the RM/RIF group. Finally, another set of frozen/thawed embryonic samples were collected from patients who underwent PGS for aneuploidy screening. Their embryos were tested for ATP levels and compared to another set of frozen/thawed embryos from the PGD group of patients.
Table 2-1: Techniques used for sample collection, preparation, processing and analysis to fulfil all aims of this thesis.

<table>
<thead>
<tr>
<th>AIM</th>
<th>Samples collection</th>
<th>Sample preparation</th>
<th>Sample processing &amp; analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td>General technique</td>
</tr>
<tr>
<td>Quantification of mtDNA</td>
<td>Spare embryos (n=58)</td>
<td>Disaggregation, tubing &amp; cell lysis (76)</td>
<td>Real time PCR</td>
</tr>
<tr>
<td>Aneuploidy screening</td>
<td>Spare embryos (n=56)</td>
<td>Disaggregation, tubing</td>
<td>WGA</td>
</tr>
<tr>
<td>Investigate the effect of parental ethnicity on fertility and embryo quality</td>
<td>Parental samples (n=46)</td>
<td>DNA extraction</td>
<td>mtDNA sequencing by NGS Genotyping AIM</td>
</tr>
<tr>
<td>Examine the effect of mtDNA/nDNA mismatch on human embryos quality and fertility</td>
<td>Parental samples (n=46)</td>
<td></td>
<td>Selection of genes Targeted nuclear genes sequencing Bioinformatics analysis</td>
</tr>
<tr>
<td>Examine the relationship between cellular ATP and embryonic aneuploidy status</td>
<td>Frozen thawed spare embryos (n=46)</td>
<td>Embryo thawing and tubing</td>
<td>ATP bioluminescence somatic cell assay</td>
</tr>
</tbody>
</table>
2.1 General laboratory practice

Different laboratory activities were carried out in specific designated areas of the laboratory and are outlined in Table 2-1.

Table 2-2: Laboratory activities were carried out in specific designated areas

<table>
<thead>
<tr>
<th>Activity</th>
<th>Laboratory area</th>
<th>Techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRE-PCR</td>
<td>Single cell room</td>
<td>single cell isolations, blastomere and embryo tubing, primer reconstitution, PCR reaction set up and whole genome amplification (WGA)</td>
</tr>
<tr>
<td>PCR</td>
<td>Main laboratory</td>
<td>Addition of genomic DNA and amplified samples</td>
</tr>
<tr>
<td>POST PCR</td>
<td>Post PCR work station area</td>
<td>High-throughput sequencing reactions set up and POST-PCR products handling</td>
</tr>
</tbody>
</table>

The single cell room is maintained under positive pressure and the air in filtered so that the total volume of air was changed 20 times per hour. Genomic DNA or amplified samples were never taken into this room. Dedicated lab coats were available and disposable fresh hats, gloves and overshoes were worn before entering. All the consumables, apart from the primers and embryonic samples, brought in the single cell room were exposed to ultraviolet light (Template Tamer, Qbiogene, UK) for at least 10 minutes to ensure the removal of DNA contaminants. All micropipettes were also exposed to UV prior to use. The deionised water was autoclaved, and this was used to make up the solutions. Solutions were reconstituted and prepared using deionised water. All the reactions were set up using DNase- and RNase-free water (Promega, UK) in DNA-, DNase-, RNase- and pyrogen-free tubes (Molecular BioProducts, Inc, UK) and prepared on an ice tray under a micro flow advanced Bio-safety cabinet class II. The pH of all the buffers was measured at room temperature (RT). Gloves were frequently changed to reduce
contamination. All the centrifugation was carried out in a bench top microfuge (MSE Microcentaur, Sanyo, UK and Labofuge 400, Heraeus Instruments, UK).

All the equipment and the designated laboratory area were cleaned with 10% hypochlorite solution, rinsed with distilled water and dried with clean tissue before each experiment. All the PCRs and WGA reactions were carried out using any of the following thermocyclers; the Applied Biosystems® GeneAmp® PCR System 9700, Mastercycler Gradient® thermal cycler (Eppendorf, UK) and the LightCycler® Nano (Roche, UK). Next generation sequencing was performed using the MiSeq and the NextSeq sequencers (Illumina Inc, San Diego, CA).

POST-PCR products (PCR analysis including mini-sequencing following PCR) were handled in a separate room using pipettes and pipette tips dedicated for post-PCR products. A PCR workstation (Template Tamer, Qbiogene, UK) was used to carry out all the post-PCR analysis and all the consumables, including pipette tips and pipettes, were exposed to UV for at least 10 minutes after each use for decontamination before the next use.

2.2 Sample collection and patient information

2.2.1 Human embryos from PGD and PGS cycles

Samples were collected from patients undergoing in vitro fertilisation intracytoplasmic sperm injection (ICSI)/ Preimplantation Genetic Diagnosis (PGD)/ Preimplantation Genetic Screening (PGS) at the Centre for Reproductive and Genetics Health (CRGH) after providing an informed consent to research. This study was approved by Human Fertilisation and Embryology Authority (HFEA) by the (project number R0113) and the Health Research Authority Committee (reference number 10/H0709/26). The samples included spare embryos that were diagnosed as
affected with a single gene disorder following PGD or were unaffected but unsuitable for transfer due to developmental arrest. Another set of embryos at blastocyst stage were collected from couples who underwent PGS (aneuploidy) / PGD families (karyomapping) that were thawed following vitrification. Those embryos were thawed using clinical procedure described in (Material and methods section 2.2.1.2). Aneuploidy and karyomapping data was provided by the IVF unit.

2.2.1.1 Embryo collection from the Centre for Genetic and Reproductive Health

Embryos in culture dishes were transported from the CRGH in an insulated box. Embryo morphology was scored by the embryologists at the CRGH on the day of sample collection (usually day 6 or 7) post fertilisation or thawing. Embryos at the blastocyst stages, were graded according to the Cornell grading system; individual scores were given to ICM compaction, blastocyst expansion and TE morphology (Veeck et al., 2004). This grading system was also used to score the embryos at the time of tubing if the embryologist failed to provide this information. This was based on counting the number of cells and recording any fragments for cleavage stage embryos.

2.2.1.2 Vitrification & thawing of embryos

Embryos were vitrified and thawed by Kaliopi Loutradi or Rabi Odia at the CRGH based on the method described in (Kuwayama et al., 2005, Amoushahi et al., 2013) with slight modifications. The embryos were transferred briefly into equilibration solution at room temperature for 7 min, containing 7.5% ethylene glycol, 7.5% DMSO, and 20% human serum albumin (HSA) in Ham’s F10. Embryos were then placed in a vitrification solution for 1 min. The vitrification solution contained 15%
ethylene glycol, 15% DMSO, 20% HAS and 0.5 mol sucrose in Ham’s F10. The embryos were loaded to cryotop with a small volume of vitrification solution (<1.0 ul) and rapidly plunged into liquid nitrogen (LN<sub>2</sub>). To warm the embryos, the cryotop was immersed at 37°C for 1 min into a solution of 1 mol sucrose containing 20% HSA in Ham’s F10. Following thawing, the embryos were sequentially transferred into a solution of 0.5 mol sucrose for 3 min followed by a solution of 0.25 mol sucrose for 3 min then washed with a medium supplemented with 20% HSA. The survival rate of embryos was assessed after 1 h of equilibration.

### 2.2.2 Genomic samples from fertile and infertile patients

Blood samples from patients undergoing PGD with no previous history of infertility (fertile group) were collected in EDTA and lithium heparin tubes. DNA extraction was performed for the fertile group of patients by Roy Naja and Seema Dhanjal at the UCL centre for PGD. DNA samples from couples who underwent three or more repeated miscarriages and/or repeated implantation failures (RM/RIF) for unknown reasons (infertile group) were from patients attending the UCH RM clinic. Couples were selected when the male partner had normal sperm parameters and the female partner was 35 years old or under. Lymphocyte separation and single cell isolations were carried for some cells (n=50) to set up the mitochondrial DNA quantification experiments.
2.3 Sample preparation

2.3.1 Single cell isolation and embryo tubing

Embryo cell tubing was performed directly after transport of the sample to the IFWH. Depending on the research objective and embryo quality, embryos were either disaggregated into single cells, clumps of cells or tubed as a whole. Usually embryos arrested at cleavage stages were disaggregated and cells were tubed separately. Most blastocysts were tubed as a whole since it was difficult to separate the cells. The disaggregation and tubing of embryos was completed using phosphate buffered saline with 0.1% polyvinyl alcohol solution (PBS/PVA, Sigma, USA and Invitrogen, UK respectively) using an inverted microscope with a 0.2mm polycarbonate microcapillary (Biohit, Finland) and then stored at -80°C.

In general, the tubing process involved removing the zona pellucida (ZP) prior to tubing any blastomere. The ZP of the biopsied embryos was removed easily by pipetting in and out. The cells freed from the ZP were washed three times in fresh drops of PBS/PVA whereas; non-biopsied embryos were exposed to acidified Tyrode’s solution (MediCult Ltd, UK) to remove the ZP. The washing and aspiration of cells was repeated until the desired number of cells from the embryo was obtained. The capillary was always washed with the fresh adjacent droplet. The washed cells from the embryo were transferred into a thin walled 0.2mL or 0.5mL microfuge tubes as described in Figure 2-2. One μL from the last wash droplet was added to another tube to serve as a negative control for that cell.
Figure 2–2: Diagram illustrating the embryonic disaggregation scheme on a petri-dish. As illustrated, the washing process of embryos and single cells was performed on a petri dish. The embryo was first transferred to drop 1 for washing from oils and remainder of the culture medium. The capillary was washed in drop 2 and the embryo was then transferred to well 3 which contain the acidified Tyrode’s to remove the ZP. As soon as the ZP was detached, the embryonic cells were aspirated to drops, 4, 5, 6 etc. to establish picking up the desired number of cells for each experiment.

2.3.2 Lymphocyte separation using the Ficoll-Paque plus method

Lymphocyte separation from blood samples was performed in the laminar flow cabinet located in the designated area in main laboratory. Upon receiving the blood sample (placed in Lithium Heparin tube), the sample was diluted 1:1 with normal saline (0.9% NaCl) in 15ml Falcon tube and mixed well. Ficoll®-Paque Premium (Sigma Aldrich) was then added in an empty 15 ml Falcon tube and layered by the diluted blood on top so that the ratio of diluted blood to Ficoll was 4:3 (Figure 2–3). The maximum volume of the blood to Ficoll, was 8:6ml. The samples were centrifuged at 1300 rpm for 30 min at room temperature. To avoid mixing the two layers, the centrifuge brake was removed. The lymphocyte layer (buffy coat) was removed and placed in a new 15 mL Falcon tube topped up to 15ml with 0.9% NaCl and centrifuged at 1300 rpm for 15 minutes in a centrifuge (with the brake on). Supernatant was discarded in a beaker containing water with diluted PreSept tablets and the buffy coat pellet was washed two more times. The buffy coat pellet was re-
suspended in a fresh 2ml 0.9% NaCl. The cell suspension was stored at 4°C until single cell isolation for no longer than two days.

Cell clumps consisting of 50, 100 and 150 cells were isolated and tubed similar to embryo tubing as described in (materials and methods section 2.3.1) excluding the exposure of the cells to acidified Tyrode’s solution step. Once the cells were isolated, they were either stored at -80°C until the day of usage or used immediately after adding the appropriate lysis buffer. These isolated single lymphocytes were used to optimise the mtDNA quantification techniques (Materials and methods section 2.4.2.1) prior to embryo processing.

Figure 2–3: Lymphocyte separation using the Ficoll-Paque plus method.

Adapted from Springer eBook, Bharadwaj et al., Detection and Characterisation of Alloreactive T Cells, pages 309-337, Copyright (2012), with permission from Springer Nature.

2.3.3 Single cell lysis

Single cells were lysed by either the alkaline lysis buffer (ALB) or proteinase K (PK) method following cell isolation as described in section 2.3.1. The efficacy of lysis for nuclear DNA has been previously determined (Thornhill et al., 2001). In this study, the efficacy of each method with mtDNA was tested. Details of each method are described below.
2.3.3.1 Alkaline Lysis Buffer (ALB)

ALB lysis depends on two processes; firstly, mechanical lysis resulting from freezing and thawing followed by chemical lysis performed by denaturing the DNA using Dithiothreitol (DTT) under alkaline conditions (50mM DTT & 200mM NaOH) pH 8.3. The buffer was prepared in the laminar flow in the single cell room. Following that, 2.5ml of the freshly prepared ALB was added to the frozen cells/ blastomeres which were frozen at -80°C for a minimum of 2 hours after tubing. Tubes were then placed on the thermocycler for incubation at 65°C for 10 min, followed by cooling at 4°C. Sample were neutralised before the PCR amplification by adding 2.5ml of 200mM Tricine (a neutralizing agent for the salts found in the alkaline lysis buffer (NaOH)) (Cui and Kim, 2007).

2.3.3.2 Proteinase K (PK) Lysis

Prior to amplification 3μL of proteinase K lysis buffer (1.25μg/μL proteinase K Roche, UK) 19.2 mg/mL, (14-22 mg/ml in 10 mM Tris-HCL, pH 7.5), 17.5μM sodium dodecyl sulphate (SDS) (Sigma Aldrich) was added to the cells. This was prepared in the single cell room according the following: 1μL Sodium dodecyl sulphate (SDS) was placed in 1000μL water in a 1.5mL microfuge tube and vortexed. 10μL from the diluted SDS solution was taken and added to another 1.5mL microfuge tube and 204μL of water was added and vortexed. In a third new 1.5mL microfuge tube 76.75μL of the diluted SDS was aliquoted and mixed with 151μL of nuclease-free water and 2.5μL of PK solution. This was all performed on an ice tray since PK might be denatured by heat. Cells were stored at -80°C until required and 3μL PK lysis buffer solution was then applied to the frozen sample (on an ice tray). This was
followed by a 1-hour incubation period at 37°C. Following this, PK was inactivated by 15 minutes incubation at 99°C then cooling to 4°C.

2.3.4 Whole genome amplification

Human embryonic samples (single cells, clump of cells or whole embryos) were whole genome amplified by Multiple displacement amplification (MDA) using the Repli-g Midi Kit from (Qiagen, UK). MDA is a non-PCR based DNA amplification technique. This technique can rapidly amplify minute amounts of DNA samples to reasonable quantity for genomic analysis. This kit was chosen because a small fraction of mtDNA is amplified and performs almost as efficiently as the single cell kit (Manufacturer’s claims).

- Primers (arrows) anneal to the template

- Primers are extended at 30°C as the polymerase moves along the gDNA or cDNA strand displacing the complementary strand while becoming a template itself for replication.

In contrast to PCR amplification, MDA:

- Does not require different temperatures
- Ends in very long fragments with low mutation rates

Figure 2–4: Multiple displacement amplification.
(Adapted from Qiagen, UK).

Following the manufacturer’s instructions, buffers and reagents were prepared and aliquoted into separate tubes sufficient for 5 reactions per tube. After the denaturing lysis buffer (DLB) was reconstituted by adding 500µL of water, the denaturation buffer (D2) was prepared by adding 5µL of Dithiothreitol (DTT) to 55µL DLB buffer. Following that, the prepared (D2) (17.5µL), the stop solutions (SS)
(17.5μL), the midi reaction buffer (MRB) (145μL), and the Phi29(5μL), enzyme was aliquoted into separate tubes sufficient for 5 reactions per tube and left at -20°C until used. Samples were placed on an ice rack, 3.5μL of D2 buffer was added gently to the wall of each sample tube ensuring the tip was changed between samples. The samples were then mixed by brief shaking on the vortex and then centrifuged. The denaturation step was then completed by incubating the samples on ice for 10 minutes. Fifty μL of deionised water was added to the MRB mix tube for 5 reactions. After vortexing and centrifuging, the total volume of MRB was added to the enzyme tube while it is placed on ice. Denaturation of samples was stopped by adding 3.5μL Stop Solution to each sample. After vortexing and centrifuging, 40μL of the MRB/enzyme mix was added to the denatured DNA. A final vortex and centrifugation took place. Reactions were incubated at 30°C for 16 hours and then the REPLI-g Midi DNA polymerase was inactivated by heating to 65°C for 3 minutes. Some embryos (Family-10, Family-20) were amplified using the Agilent GenetiSure Amplification kit. This kit is also an MDA based whole genome amplification technique with a slight modification in the denaturation and extension steps as per the protocol.

2.4 Sample processing and analysis

2.4.1 General methods

2.4.1.1 Selection of DNA region and Primer design

The complete updated sequence (Revised Cambridge Reference Sequence (rCRS) of Human Mitochondrial DNA (GenBank sequence NC_012920) was obtained from Ensembl genome browser, http://www.ensembl.org/index.html and MITOMAP http://www.mitomap.org/MITOMAP where the mtDNA sequence was based on the L
strand. Primers were designed using the primer 3 tool in a product size range of 200bp for mtDNA and 250-300bp for the nDNA (hg 19 reference genome (build 37) (http://frodo.wi.mit.edu/). The best primer pair with a length of 18-23bp was chosen based on the guanine-cytosine (GC) base content (maximum of 55%), ensuring the absence of SNPs and similar melting temperature (Tm) values between the forward and reverse primers. The alignment of the chosen primers within the whole genome was analysed on Basic Local Alignment Search Tool (BLAST) program. Primers were ordered in Tris-EDTA buffer (Eurogentec, UK) and to prevent multiple freeze-thawing, upon arrival, each primer was aliquoted in 20µL volumes at 50µM. These aliquots were kept at -20ºC whereas the original stocks were stored at -80ºC.

2.4.1.2 Polymerase chain reaction (PCR)

Cells and embryos lysed by PK required only a single round of PCR and therefore (after adding the real time PCR master mix) they were transferred from the 0.5ml microfuge tubes to the Light Cycler® Nano reaction tubes and processed immediately. Cells lysed by ALB required neutralisation of the alkaline lysis buffer using Tricine prior to real time PCR analysis. The reaction consisted of the following: 12.5µL Taq PCR Master Mix (Qiagen, UK), 0.1µL of each primer (to obtain a final concentration of 0.2µM) and 1µL of DNA was added, if a genomic DNA sample (100-120ng/µL) was used while if a single cell was tested, 3µL of the sample in lysis buffer was added to the mixture. The reaction master mix was topped up with the required amount of RNAase free water to obtain a final reaction volume of 25µL. As the genomic DNA was added in the main laboratory, an additional blank control was included in order to detect if there was any contamination in the single cell room, main laboratory and POST-PCR work station areas.
The first step in PCR included denaturation at 95°C for 2 minutes followed by 2 thermal cycling series. The first series included a three-step amplification process which involved denaturation at 96°C for 15 seconds, annealing at 60°C for 45 seconds and extension at 72°C for 1 minute under 10 cycles. The other amplification step was performed by denaturation at 94°C for 15 seconds, annealing at 60°C for 45 seconds and extension at 72°C for 1 minute. This was repeated 29 times. Finally, the extension step of the reaction was at 72°C for 8 minutes. PCR products were then stored at 4°C.

2.4.1.3 Agarose gel electrophoresis

PCR products and WGA products were analysed using agarose gel electrophoresis. DNA molecules were separated by size using 0.8% Agarose gel (agarose gel powder type 1, Sigma, USA) in 50ml of 1xTris/Borate/EDTA (TBE) buffer (ThermoFisher, UK). Agarose and TBE mixture was heated in a microwave for 2 minutes. Following that, 5µL of SYBR Safe DNA gel stain was aliquoted to the melted agarose (Sigma, UK) to visualize the DNA bands on the gel. The melted agarose was then poured into the gel mould with well combs placed in position and allowed to cool down. When set, the comb was removed, and the buffer was added to the tank gel. Five µL from each amplified sample was mixed with 6x loading dye (10 Mm Tris-HCL at pH 7.6,0.03% bromophenol blue, 0.03% xylene cyanol FF,60% glycerol and 60mM EDTA) and aliquoted into the gel wells. One and a half µL of GeneRulerTM 100bp-1kb plus DNA ladder (Fermentas, UK) was loaded on to the gel for DNA band size comparison. The samples were run on the gel in 1XTBE buffer at 75V for 15-20 minutes. Analysis of DNA bands on the gel was then carried
out under UV light using the BioDoc-It™ Imaging Systems (UVP, UK). An image of the gel was taken for comparison with gels of other amplified products.

2.4.1.4 Quantification of DNA

2.4.1.4.1 NanoDrop
The concentration of DNA was determined using the NanoDrop technology (NanoDrop ND-1000, ThermoFisher Scientific, UK) according to the manufacturer’s protocol. The machine was calibrated using DNAase and RNase-free water and the concentrations of nucleic acids were measured in sample volumes of 1µL at 260 nm. The measurements of DNA concentration were displayed in ng/µL.

2.4.1.4.2 Qubit
DNA samples used for next generation sequencing were quantified by the Qubit® dsDNA HS (High Sensitivity) Assay Kit. This is a highly sensitive assay designed to accurately measure the double-stranded DNA (dsDNA) over RNA for initial sample concentrations from 10pg/µL to 100 ng/µL. In brief, the Qubit working solution was prepared by diluting the Qubit dsDNA HS Reagent 1:200 in Qubit dsDNA HS Buffer according the manufacturer’s instruction. A clean falcon plastic tube was used each time the Qubit working solution was prepared. A maximum of 25 reactions were prepared each time which included 2 standards and one extra reaction to overcome minor pipetting errors. Standard 1 was composed of TE buffer while Standard 2 was made of λ DNA (0.35-0.6µg/µL). To reach a final volume of 200µL in each tube, each standard tube required 190µL of Qubit working solution plus 10µL of standard, and each sample tube required 199µL plus 1µL of sample. Sufficient Qubit working solution was prepared to accommodate all standards and samples. All tubes were mixed by vortexing 2-3 seconds while bubbles were avoided. Following that all tubes
were incubated at room temperature for 2 minutes. Finally, the concentrations were read on the Qubit Fluorometer.

2.4.1.5 Assessment of DNA library quality and quantity

The quantity and quality of the DNA libraries was assessed using the Agilent 2100 Bioanalyzer and/or Agilent TapeStation (Agilent, CA, USA).

2.4.1.5.1 Bioanalyzer

The Bioanalyzer was used to assess the quality and size of the amplified DNA libraries. DNA 100 chip (Agilent, UK) was used following manufacturer’s protocol for quantification. This analysis revealed the size and concentration of DNA libraries through the peak shown by the electropherogram. The size of the DNA amplicons was integrated through the position of the peak shown by the electropherogram (for example; 245 to 325 bp). The concentration of each library was measured by integrating the area under the entire observed peak. The concentration of each library was measured again after the indexes were added. The peak of DNA fragment size and concentration of each library was measured again by integrating the area under the entire peak. It was confirmed that the concentration was within the linear range of the assay for accurate quantification (Figure 2-5).
2.4.1.5.2 TapeStation

The Agilent D1000 ScreenTape and the associated reagent kit were used to analyse the quantity and quality of amplified next generation sequencing libraries using the 2200 TapeStation (Agilent, UK). This was performed by diluting 1μL of each amplified library DNA sample with 3μL of D1000 sample buffer on a plate. The plate, loading tips and D1000 screen tape were then loaded into the TapeStation. The concentration of each library was measured by integrating the area under the observed peak by the electropherogram. Similarly, the High Sensitivity D1000 ScreenTape and reagent kit were used to analyse the amplified indexed DNA libraries using the 2200 TapeStation. Two μL of each indexed DNA sample was diluted with 2μL of High Sensitivity D1000 sample buffer for the analysis. The High Sensitivity D1000 ScreenTape, the sample plate and loading tips were loaded into the TapeStation. The DNA fragment size and the concentration of each library was measured by integrating the area under the entire peak shown by the electropherogram.

2.4.1.6 Purification of adapter tagged libraries using Magnetic beads

AMPureXP magnetic beads were used to purify the adaptor-tagged libraries. Before starting AMPureXP beads were incubated at room temperature for at least 30 minutes. Fresh 70% ethanol was prepared each time the beads were used for purification. When starting the AMPureXP bead suspension has been vigorously re-suspended by vortexing until it appeared homogeneous and consistent in colour. 52μL of the homogeneous bead suspension was added to each well containing the DNA samples. Wells were sealed with a fresh aluminium cover, then vortexed for 5 seconds. Plates were centrifuged so that the samples get mixed, without pelleting
the beads. Samples were then incubated for 5 minutes at room temperature. The plate was placed in a magnetic stand and left at room temperature until the solution was clear for approximately 3 to 5 minutes. While keeping the plate or tubes on the magnetic stand, the cleared solution was carefully removed and discarded from each well using a micropipette. 200μL of fresh 70% ethanol was added to each sample well while the plate was still on the stand. After one minute, the ethanol was removed by a micropipette and the remaining few drops were dried. This step was repeated for a total of two washes. Samples were then dried on the thermal cycler (with lids open) at 37°C for 1 to 3 minutes making sure that the precipitate did not dry. 11μL of nuclease-free water was then added to each sample well. Samples were sealed, mixed and centrifuged briefly to collect the liquid. After incubation for 2 minutes at room temperate the plate was placed on the magnetic stand again and left for two minutes until the solution was cleared. Each cleared supernatant (approximately 10μL) was removed to the wells of a fresh plate and kept on ice until use.

2.4.1.7 Aneuploidy screening by array-CGH

Array CGH was performed in order to examine the aneuploidy status of all the chromosomes in the embryos. Agilent GenetiSure Pre-Screening Microarrays (Agilent, UK) were used to establish the aneuploidy status of the embryonic samples. The first step was WGA of the embryonic samples. Most of these embryos were previously whole genome amplified by MDA using the Repli-g Midi Kit from (Qiagen, UK). The remaining embryos (Family-10 and Family-20) were whole genome amplified by the same technique (MDA) but using the Agilent GenetiSure Amplification kit for 85 minutes, as per the protocol, (Agilent GenetiSure Pre-Screen
Kit for Single Cell Analysis v. A0, UK). The difference in denaturation technique between both experiments was tested and there were no significant differences. Labelling, hybridisation and microarray kits were implemented according to manufacturer’s protocol (Agilent, UK) described briefly in Figure 2-6.

![Diagram of embryo analysis for chromosomal imbalances using a-CGH](image)

**Figure 2–6: Schematic diagram of embryo analysis for chromosomal imbalances using a-CGH.** 8x60K slides were used for a total of 14 experimental samples on one slide. Two experimental samples were hybridised on the same array and compared to a Male and Female Reference sample co-hybridised to another array on the slide. If there was an odd number or insufficient number of samples to fill the slide, one or more additional male or female reference samples were added to pair with the odd experimental sample or to fill the empty microarrays was labelled.

8x60K array slides were used to test 14 independent samples per slide. The male and female reference DNA provided by Agilent were used alongside the samples and amplified for the same duration and under the same conditions to account for any artefacts incorporated into the samples during the amplification process. MDA was applied so that more than one test could be performed on the same embryo including next generation sequencing (NGS). Labelled test and reference samples...
were hybridised on oligonucleotide microarrays specifically designed for single cell screening. Single cell analysis was carried out with algorithms implemented in the Agilent CytoGenomics v3.0 software. All processes included in the preparation and hybridisation, were completed in the dark in order to reduce the negative impact of light on the labelled samples. Samples were tested over a few days in batches; to avoid variation between samples which may be caused by the atmospheric ozone. Following amplification, the concentration of samples was measured by Qubit and diluted to 50ng/μL. The amplified embryonic DNA samples and references (genomic DNA sample provided by Agilent that were diluted and whole genome amplified) were fluorescently labelled with either Cyanine 3 or Cyanine 5 (Cy3 or Cy5) for 55 minutes in a thermal cycler. Briefly, 13μL of each amplified sample was transferred to a new 200μL tube, 2.5μL random primers solution for a total volume of 15.5μL and mixed by pipetting up and down. Samples were then transferred to the thermal cycler for denaturation at 98ºC for 3 minutes and left at 4ºC to hold. During that time, a labelling master mix was prepared according to the associated protocol; 5μL of reaction, 2.5μL of dNTPs mix, 1.5μL of either Cy3-dUTP or Cy5-dUTP was added to the sample, 0.5μL of Exo Klenow enzyme was added. Following that 9.5μL of this master mix was added to each sample to make a total of 25μL.

The male reference was labelled by Cy3 and the female reference was labelled by Cy5. Half of the test samples were labelled by Cy3 labelling master mix and the other half was labelled by Cy5 labelling master mix. If the number of samples used was odd, the male reference sample was used in duplicate. The samples were labelled by incubation at 37ºC in a pre- warmed thermal cycler for 45 minutes followed by 10 minutes at 65ºC. Once this step was completed, the differentially labelled male and female reference samples were combined to create a Cy3-Cy5-
pair with a final volume of 50µL. Similarly, the differentially labelled test samples were combined to make Cy3-Cy5-pairs.

Following that, the paired Cy3-Cy5 samples were purified using the post labelling purification columns and collection tubes provided by the kit. Briefly, the 50µL of each labelled sample was transferred to a 1.5mL tube and 430µL of 1X TE (pH 8.0) was added. Samples were loaded into purification columns and centrifuged for 15 minutes at 9,000xg at room temperature. The flow through was then discarded and columns were inverted into fresh 1.5mL collection tubes and centrifuged for 2 minutes at 2,500xg to collect the purified sample. 15µL of each purified sample was transferred to a fresh 200µL reaction tube. When the sample volume was less than 15µL, 1xTE (pH 8.0) was added to bring the total volume to 15µL. After purification, the excess samples were frozen at -80°C to allow for repeats to be carried out.

2.4.1.7.1 Hybridisation and washing
The hybridisation master mix was prepared according to the following (Table 2-3).

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume (µL)</th>
<th>x4 reactions including excess</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human Cot-1 DNA</td>
<td>5</td>
<td>42.5</td>
</tr>
<tr>
<td>10x array CGH Blocking agent *</td>
<td>5</td>
<td>42.5</td>
</tr>
<tr>
<td>2x HiRPM Hybridization buffer</td>
<td>25</td>
<td>212.5</td>
</tr>
<tr>
<td>Final volume</td>
<td>35</td>
<td>297.5</td>
</tr>
</tbody>
</table>

*Blocking agent was prepared by adding 1.35 DNase/RNase free distilled water to the vial containing the lyophilised 10x a-CGH Blocking agent. The mixture was left at room temperature for 60 minutes then mixed on a vortex mixer to reconstitute the sample before usage. This was then stored at -20°C.

A total of 35µL of hybridisation master mix was added to each sample tube to make a volume of 50µL. Samples were mixed by pipetting up and down and spanned briefly. Sample tubes were then incubated at 98°C for 2 minutes followed by 18 minutes at 37°C on a thermal cycler for preparation of hybridisation.
A gasket slide was loaded into the Agilent SureHyb chamber base (according to the manufacturer’s instruction). Each sample (45μL) was placed onto the slide in the centre of the assigned gasket well, avoiding contact between the samples leaving a bubble to allow the mobility of the hybridisation buffer. A microarray slide was then loaded on top of the gasket slide creating a sandwich-pair with proper alignment. The SureHyb chamber cover was put onto the sandwiched slides and sealed properly. The assembly chamber was vertically rotated to wet the slides and assess the mobility of the bubbles before placing them in the hybridisation oven. Each assembled chamber was then loaded into the oven rotator rack for hybridisation. Slides were hybridised at 67°C overnight for 18 hours.

Following that, the microarray wash procedure was performed for all the samples in a batch at the same frame time to avoid the effect of ozone level. Samples pretested by FISH and showed to be normal were confirmed by a-CGH. Wash buffer 1 and wash buffer 2 were used for the washing steps. Three coplin jars were prepared for washing the slides (Table 2-4).

**Table 2-4 Washing conditions following a-CGH hybridisation**

<table>
<thead>
<tr>
<th>Slide staining dish</th>
<th>Wash</th>
<th>Dish</th>
<th>Wash buffer</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Disassembly</td>
<td>500mL</td>
<td>Wash buffer 1</td>
<td>RT</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>1st wash</td>
<td>500mL</td>
<td>Wash buffer 1</td>
<td>RT</td>
<td>5 minutes</td>
</tr>
<tr>
<td>3</td>
<td>2nd wash</td>
<td>500mL</td>
<td>Wash buffer 2</td>
<td>37°C</td>
<td>2 minutes</td>
</tr>
</tbody>
</table>

Cover slips were removed manually by agitating in a slide staining dish 1 containing wash buffer 1. The slides were then washed to remove any un-hybridised DNA in slide staining dish 2 by placing the slide with a stirrer for 5 minutes on a magnetic stir plate on, to adjust the settings to get thorough mixing without disturbing
the microarray slides. The final wash was in the pre-warmed wash buffer 2 for two minutes with the magnetic stirrer switched on. The slides were left to dry at room temperate for 3 minutes then scanned.

2.4.1.7.2 Scanning and analysis of a-CGH slides

Array CGH slides were scanned, immediately after washing, within an Agilent scanner in The Doctor’s Laboratory. TIF images were checked for evenness by (Alessandra Callegari) at the Doctor’s Laboratory. The second part of the evaluation was carried out using the CytoGenomics software (v.3, Agilent) to analyse the chromosomal complement of the embryonic samples. All samples were analysed using the “single cell recommended analysis method”.

The cut-off for a gain of a segment/whole chromosome was considered when the log ratio was more than +0.35 following the CytoGenomics Software. When the log ratio was less than -0.45, this was considered to be a loss of that particular segment/whole chromosome. Validation of the arrays was performed as part of a master’s project done by Arwa Al Mutlaq (Appendix A section 7.1.1). Segmental losses and gains were only considered if they were greater than 50 Mb.

Due to the data distribution, embryos were then classified as euploid, aneuploid, aneuploid ≤5 or having multiple chromosomal aneuploidies > 5. The number of aneuploidies within each embryo was categorised as aneuploid when a single monosomy or single trisomy was detected, aneuploid ≤5 when chromosomal aneuploidies were less or equal to five and multiple aneuploidy >5 when more than five chromosomal aneuploidies were present. Repeated analysis was carried out if there were significant differences in the software calls between the male and female
references, or if the analysis failed more than one of the 14 of the internal quality control checks (Table 2-5).

**Table 2-5: QC metric thresholds used by Agilent for experimental samples analysed by GenetiSure.**

<table>
<thead>
<tr>
<th>Metric</th>
<th>Preferred range</th>
<th>Metric description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is Good Grid</td>
<td>&gt;=1.0</td>
<td>Tracks how well the software was able to align the grid to the array spots.</td>
</tr>
<tr>
<td>AnyColorPrcntFeatNonUnifOL</td>
<td>&lt;5.0</td>
<td>The percentage of features that are of poor quality in either channel, and therefore excluded from the analysis.</td>
</tr>
<tr>
<td>DerivativeLR_spread</td>
<td>&lt;0.7</td>
<td>Measurement of probe-to-probe noise along the chromosomes in log ratio space. Critical metric for aberration calling.</td>
</tr>
<tr>
<td>Dye Flip</td>
<td>NA</td>
<td>Refers to the polarity input by the customer. &quot;Normal&quot; is reference in green, &quot;Sample&quot; is red. Flipped is the reverse.</td>
</tr>
<tr>
<td>gRepro</td>
<td>&lt;0.2</td>
<td>Measurement of consistency of Cy3 signal across replicate probes on the array surface.</td>
</tr>
<tr>
<td>g_BGNoise</td>
<td>&lt;15.0</td>
<td>Standard deviation of the background-subtracted signals of the negative control probes in the Cy3 channel. These numbers are affected by hybridization and wash stringency.</td>
</tr>
<tr>
<td>g_Signal2Noise</td>
<td>&gt;10.0</td>
<td>Ratio between the probe signal and the background noise, for the Cy3 channel.</td>
</tr>
<tr>
<td>g_Signal Intensity</td>
<td>&gt;30.0</td>
<td>Measurement of overall intensity of the Cy3 channel after software processing.</td>
</tr>
<tr>
<td>rRepro</td>
<td>&lt;0.2</td>
<td>Measurement of consistency of Cy5 signal across replicate probes on the array surface.</td>
</tr>
<tr>
<td>r_BGNoise</td>
<td>&lt;15.0</td>
<td>Standard deviation of the background-subtracted signals of the negative control probes in the Cy5 channel. These numbers are affected by hybridization and wash stringency.</td>
</tr>
<tr>
<td>r_Signal 2 Noise</td>
<td>&gt;8.0</td>
<td>Ratio between the probe signal and the background noise, for the Cy5 channel.</td>
</tr>
<tr>
<td>r_Signal Intensity</td>
<td>&gt;25.0</td>
<td>Measurement of overall intensity of the Cy5 channel after software processing.</td>
</tr>
<tr>
<td>Restriction Control</td>
<td>NA</td>
<td>Measurement of restriction enzyme digest quality.</td>
</tr>
<tr>
<td>Log Ratio Imbalance</td>
<td>NA</td>
<td>The amount of amplifications versus deletions per chromosome. Not applicable to single cell analysis.</td>
</tr>
</tbody>
</table>

Adapted from QC Metrics Guide for Agilent CytoGenomics Software (Agilent, UK).
Chapter 3 **Assessment of mtDNA copy number and ATP with preimplantation embryo quality**

3.1 Aims Summary

Mitochondria from the oocyte are essential for the embryo since they support early development until the resumption of mitochondrial replication which occurs after implantation. They are the principal site of energy production in the form of ATP and have various other critical cellular functions, such as regulation of spindle organisation, chromosomal segregation, cell cycle and other dynamic processes such as: compaction, cavitation and blastocyst hatching.

A certain level of ATP is required for successful maturation of the nucleus and cytoplasm in preparation for fertilisation and completion of meiosis II (Chappel, 2013). Good quality oocytes containing optimal mitochondrial numbers (Reynier et al., 2001, Santos et al., 2006) and sufficient levels of ATP (at least 2pM per oocyte) (Van Blerkom et al., 1995) produce higher quality blastocysts after fertilisation (Takeuchi et al., 2005). On the other hand, low ATP and decreased mtDNA copy number are associated with poor oocyte quality, poor embryo development and suboptimal implantation and placentation rates (Chappel, 2013). Little is known about the extent of variation in mtDNA copy number between individual human embryos prior to implantation. Recently it has been shown that the quantity of mtDNA was significantly higher in aneuploid embryos and embryos from older women (Fragouli et al., 2015). Others have shown a link between mtDNA haplotype and aneuploidy in the oocyte (Gianaroli et al., 2015). These reports point to two different sources of aneuploidies (prezygotic, post zygotic errors) and mtDNA number as a reflection of mitochondrial function. The aim of this study was:
A. To investigate the effect of mtDNA copy number on embryo development using two real time-PCR approaches:

- Mitochondrial DNA absolute quantification by Resolight & TaqMan SNP genotyping assays; samples compared to a synthetic oligonucleotide.
- Relative quantification by TaqMan copy number assay; samples normalised to an internal nuclear gene ($RNAase \, P$).

B. To examine the link between ATP levels and aneuploidy in blastocysts. Embryos following PGD and PGS, which had been frozen after biopsy for the assessment of aneuploidy status, were thawed and used in this analysis.

3.2 Materials and methods

- A synthetic oligonucleotide with sequence from position 8926 to 9075 in the $MT$-$ATP$-$6$ gene of the mitochondrial genome (Heavy strand) was synthesised (Eurogentec, UK) in order to validate the qPCR method for the quantification of mtDNA number. The oligonucleotide was diluted to specific concentrations and used to generate a standard curve from qPCR of the oligonucleotide using primers ATP6-F and R. Embryos and single cells were then tested by qPCR using the same primers and the standard curve was used for inferring the mtDNA template number in the sample. In total, 76 embryonic samples were collected and analysed by this method.

- Relative quantification to the nuclear $RNAase \, P$ gene by TaqMan copy number assay targeting a specific mtDNA fragment on the $MTRNR2$ gene was used to test another set of embryonic samples ($n=54$) which had already been WGA as described in 2.3.4. The aneuploidy status of these embryos was determined. An aliquot of the WGA product for analysis with GenetiSure
(Agilent, UK) according to the procedure described in 2.4.1.7. The data was used to investigate the relationship of mtDNA copy number with aneuploidy, embryo development and maternal age.

- The ATP content in embryos was determined using the Bioluminescence Somatic cell assay kit, Sigma. Samples included 46 frozen / thawed blastocysts post PGD/PGS. The aneuploidy data of the selected embryos was previously assessed by BlueGnome arrays and/or karyomapping. Samples were classified into 3 groups based on the number of chromosomal aneuploidies/imbalances.

The workflow for assessing the mtDNA copy number and ATP levels in preimplantation embryos and single cells is summarised in Figure 3-1.

**Figure 3–1: The work flow for assessing mtDNA template number and ATP levels in embryonic samples.**

Embryos at different developmental stages were collected, tubed and lysed. A synthetic oligonucleotide (mtDNA from position 8926 to 9075 in the MT-ATP-6 gene) was used as standard. The oligonucleotide was diluted in serial dilutions. Primers were designed to amplify the region covered by the oligonucleotide and to amplify the same region in the test sample for absolute quantification by the Resolight and TaqMan genotyping assay. The mtDNA copy number of another set of embryos, which were WGA by MDA, was quantified by the TaqMan copy number assay. An aliquot from the WGA product was also used to test for aneuploidy. The third set of embryos consisted of frozen/thawed blastocysts following PGS or PGD. The ATP analysis of those embryos was assessed and linked the aneuploidy data which was already available.
3.2.1 Assessment of mitochondrial DNA copy number by qPCR

Three types of qPCR using the Light Cycler® Nano (Roche, UK) were tested in order to develop a method that would accurately quantify and genotype mtDNA in single cells and blastomeres. These methods included: Absolute quantification and genotyping by high resolution melting (HRM) and TaqMan genotyping assays, as well as relative quantification by TaqMan copy number assays. Although all methods are capable of quantifying and detecting genotypic differences between samples, the principle behind each method is slightly different requiring different primers to be designed.

3.2.1.1 Absolute quantification of mtDNA by Resolight and TaqMan genotyping assay

To determine the level of mtDNA present in embryos at different stages of preimplantation development a synthetic oligonucleotide sequence corresponding to positions 8926-9075 in \( MT-ATP6 \) of mtDNA (heavy strand), was used as a standard. A pair of forward and reverse primers was designed to amplify the region in the mitochondrial genome position 8926-9075 shown below (Figure 3-2). Those primers were used to amplify the mtDNA of the oligonucleotide and the test samples.

![Figure 3-2: The synthetic oligonucleotide used for quantification of mtDNA.](image)

A single stranded oligonucleotide (149 bases) between positions 8926-9075 in \( MT-ATP6 \) of the mitochondrial genome was designed. Forward and reverse primers were designed to amplify this region of the oligonucleotide as well as the test sample.

To establish mtDNA number in single embryonic cells, the oligonucleotide was diluted to a concentration similar to what has been estimated in the literature according to the calculation mentioned in the (Appendix B section 7.2.1).
Quantitative PCR assay was optimised by running a 1/10 serial dilution (Figure 3–3) so that each dilution (containing known amounts of mtDNA template) represented a different target quantity in the reaction. The analysis of amplification provided a specific cycle of quantification (Cq) value that was then automatically plotted against the known mtDNA number in each dilution forming a standard curve. Comparison of the Cq values from test samples (with unknown amounts) to this standard curve provided an estimate of the starting amount of template in each sample. An additional Tm Calling analysis was performed to check if any nonspecific products were present.

![Diagram of serial dilution](image)

**Figure 3–3: Serial dilution of synthetic oligonucleotide.**
1/10 serial dilution of the synthetic oligonucleotide was performed 8 times to reach the level of mtDNA molecules in a single blastomere on day 3.

3.2.1.1.1 Resolight PCR master mix and conditions

The setting up of the High-Resolution Melting (HRM) PCR master mix and addition of embryonic samples was performed in the single cell room while the oligonucleotide was added to each reaction tube in the main laboratory to avoid contamination. The master mix was prepared in a 0.5ml microfuge tube to prepare a final reaction volume of 10µL. Briefly, the reaction mix contained: 5µL of the Light
Cycler® 480 HRM master mix, 0.05μL of 50μM (to obtain a final concentration of 0.25μM per reaction) of each primer, 1μL (2.5μM final concentration) of MgCl₂ and 2.9μL of deionised H₂O for analysis of genomic samples or 0.9μL for single cells lysed by PK (Chapter 2 section 2.3.3.2). When genomic DNA was analysed, 1μL of genomic DNA was added in the main laboratory. However, if amplified product was used from a first round PCR (such as in the case of single cells and embryos lysed by ALB (Chapter 2 section 2.3.3.1) 1.5μL of amplified product was added in the post PCR area.

The PCR conditions were as follows: Denaturation at 95°C for 10 minutes with a rate of amplification (Ramp) of 4°C/s. Following that a 3-step amplification of 45 cycles (denaturation at 95°C (Ramp 5°C/s) for 10 seconds, annealing at 60°C for 25 seconds and extension at 72°C for 15 seconds (Ramp speed 4°C/s) was performed. In order to quantify the amount of DNA, another denaturation step at 95°C, for 1 minute with a Ramp speed of 4°C/s followed by cooling at 40°C (Ramp speed 4°C/s) for 1 minute was applied. The HRM initial stage was set at 65°C with a Ramp speed of 4°C/s for 1 minute and the final stage was set at 95°C with a Ramp speed of 0.05°C/s for 1 minute respectively.

3.2.1.1.2 TaqMan genotyping assay PCR conditions

The TaqMan SNP genotyping assays were custom designed and ordered from Applied Biosystems. The PCR reaction set up was prepared in a 0.5mL microfuge tube to obtain a final reaction volume of 25μL. This process was performed in the single cell room on an ice tray. And the master mix was prepared by adding 12.5μL of TaqMan Universal PCR, 0.625μL of the custom designed 40x Assay Mix, 1μL of DNA from a genomic sample and 3μL when a single cell lysed by PK was used. The
rest of the reaction was topped up with RNAase free water. When amplified product from a first round PCR was used as in the case of (ALB lysis of single cells), the 1.5µL of amplified product was added to the master mix in a post-PCR area. PCR conditions were as follows: denaturation at 95°C for 10 minutes at Ramp of 5°C/s followed by a two-step amplification under 60 cycles. The amplification step was initiated by denaturation at 95°C (Ramp 5°C/s) for 15 minutes followed by an annealing step at 60°C (Ramp 4°C/s) for 1 minute. The acquisition setting was set at elongation.

3.2.1.2 Relative quantification of mtDNA by TaqMan Copy Number Assay

A custom-designed TaqMan copy number Assay (ThermoFisher Scientific, UK) AATTTAACTGTAGTCCAAAGAG was used to target and amplify a specific mtDNA fragment of the mitochondrial 16s ribosomal RNA sequence on the MTRNR2 gene, adapted from (Fragouli et al., 2015). Whole genome amplified embryonic samples were all diluted to 5ng/µL. Normalisation of input DNA was performed by the internal RNase P gene. Normalisation relative to a nuclear DNA sequence was performed to ensure that any variation in mtDNA levels related to technical issues such as differences in the number of cells within the biopsy specimen or due to inefficiency of WGA could be controlled. The PCR reaction set up was prepared in a 0.5mL microfuge tube to obtain a final reaction volume of 20µL. This process was performed in the single cell room on an ice tray and the master mix was prepared by mixing 10µL of 2X TaqMan Genotyping Master Mix, and 1µL of TaqMan mtDNA designed Copy Number Assay, 20X, 1µL TaqMan Copy Number Reference Assay, 20X and 4µL of WGA amplified product (5ng/µL). The rest of the reaction was topped up with RNAase free water. The Light Cycler® Nano was used (Roche, UK).
PCR conditions were as follows: incubation at 50°C for 2 minutes, incubation at 95°C for 10 min and then 40 cycles of 95°C for 15s and 60°C for 1 minute. Relative quantification was performed using the 2^{-ΔΔCT} (User Bulletin #2; Applied Biosystems). The two groups were compared based on the non-parametric Mann-Whitney U test with 95% confidence interval (i.e. p<0.05 was considered statistically significant).

### 3.2.2 Assessment of ATP content in human preimplantation embryos

The ATP content in the frozen/thawed blastocysts was determined by the measurement of the luminescence generated in an ATP-dependent luciferin—luciferase bioluminescence assay by the (Bioluminescence Somatic cell assay kit, Sigma) using the (FLUOstar Omega Luminometer) adapted from the methods described previously (Blerkom et al., 1995, Zahao et al., 2012) and the manufacturer's instructions. The principle of the reaction is described in Figure 3–4.

All reagents provided in the kit were aliquoted in to several tubes and placed at -20°C for not more than 5 days to avoid freezing-thawing effect and contamination. An ATP standard, provided in the kit, was also aliquoted into several tubes and a fresh dilution series was made up with every reaction set up. The standard was diluted in a 1/2.5 serial dilution as explained in the Appendix B section 7.2.2. However, only the last 5 dilutions were used in the analysis.

Before using the Luminometer, a series of two priming steps were performed in the following order: dH₂O, 100% ethanol, dH₂O respectively. This was repeated after using the Luminometer to avoid contamination. When measuring ATP in embryos, after the following steps were performed, an extra step of priming was performed using the ATP assay mix. The Luminometer protocol was set on Luminescence
mode with the gain=2500, Kinetic windows =1, number of intervals = 20 seconds. A maximum 16 samples per run were analysed.

![Diagram](image)

**Figure 3–4: Diagram illustrating the principle for ATP measurement in this procedure.**

The luciferase enzyme in fireflies catalyses the modification of luciferin, consuming both luciferin and ATP, and producing light. Assuming ATP is in excess, the reaction follows the Briggs-Haldane mechanism, with luciferin as the limiting substrate. Reactions 1 and 3 are essentially irreversible and Reaction 2 is reversible. The light emitted is proportional to the ATP present, which is in turn proportional to the number of cells in the sample. Adapted from Methods of Enzymatic Analysis, 4 (1974), Stehler., Adenosine-5′-triphosphate and Creatine Phosphate Determination with Luciferase, pages 2112-2126, Copyright (1971), with permission from Elsevier.

Briefly, 98µL of ultrapure water (17MW.cm) was added to each tubed sample (embryo) prior to the analysis. After lyophilizing and preparing all the reagents according to the manufacturer’s instructions, 100µL of somatic cell ATP releasing agent was then added to each sample and mixed by pipetting 10 times; then 100µL was taken from the sample mixture and placed in an empty well. Following that, the plate containing the sample was placed in the Luminometer with a shaking time set for 8 seconds before plate reading, orbital 600 rpm and an injection of 100µL of the ATP Assay Mix containing (luciferase, luciferin, MgSO₄, DTT, EDTA, bovine serum albumin (BSA), and Tricine buffer salts) for 10 seconds (average 5 readings from 1 second after assay mix injection i.e. 11 seconds. was performed by the Luminometer to each sample in the well. The light was then immediately measured by the Luminometer.
A standard curve containing 5 ATP concentrations was generated. ATP concentration in the embryos was inferred by plotting the relative light units (RLU) of the samples against the standards of known concentrations and was calculated using the formula derived from the linear regression of the standard curve. This was repeated for each run which was analysed separately.

### 3.3 Results

In total 76 samples were collected from 16 couples (Table 3-1) and used for mtDNA copy number quantification by Resolight and TaqMan assays. All samples were graded prior to tubing (day 5 grading). Samples analysed by relative quantification to nDNA and ATP were graded on the day of collection and processing and analysed accordingly.

**Table 3-1: Samples used for mtDNA quantification by Resolight and TaqMan assays & ATP assessment:**

<table>
<thead>
<tr>
<th>Test</th>
<th>Quantity</th>
<th>Sample type</th>
<th>Grade</th>
</tr>
</thead>
</table>
| mtDNA quantification by Resolight and TaqMan compared to morphology on day 5 | 76       | 50 Single cells           | 16 Arrested  
4 blastocysts  
16 Morula  
4 Pre-Morula |
|                                                            |          | 13 Two tubed cells        | 5 Arrested                 |
|                                                            |          | 13 Whole embryos          | 3 Arrested  
3 Blastocysts  
3 Morula  
3 UF  
1 Pre-Morula |
| Relative quantification of mtDNA by TaqMan copy number *  | 56       | WGA                       | Details mentioned in Table  
presented in Appendix B section 7.2.3 |
| ATP measurement*                                          | 46       | Frozen thawed blastocysts | Blastocysts                |

*The test mtDNA copy number / ATP measurement was compared to morphology on the day of collection and/or aneuploidy.
3.3.1 Quantification of mtDNA by Resolight and TaqMan assays

3.3.1.1 Results from the synthetic oligonucleotide

3.3.1.1.1 Quantification of mtDNA copy number using a single round of PCR

Absolute quantification and genotyping of the synthetic oligonucleotide (1/10) dilution series using HRM, resulted in an approximate Cq difference of three with each dilution. This was clearly seen with samples of dilutions 4 through 8 (Figure 3–5(a)). The specificity of the PCR reaction and the absence of non-specific products was also confirmed by the presence of a single melting peak corresponding to that of the expected amplicon (Figure 3–5(b)).

![Figure 3–5: Absolute quantification of the normal oligonucleotide dilutions by Resolight and HRM.](image)

Genomic DNA had a Cq value falling between dilutions 5 & 6 of the synthetic oligonucleotide. This corresponded to 3.01x 10^7 copies of mtDNA. b) Shows the consistent single melting peak (Tm=84°C) indicating the specificity of the PCR reaction and the absence of contamination.

A standard curve plotting the concentrations of the oligonucleotide dilutions was generated by the software (Figure 3–6). This was used in order to infer the concentrations on the test samples (Table 3-2).
As shown in the figure, the calculated concentrations of each dilution were plotted on the x axis. While the measured Cq values were plotted on the y axis.

Further validation was carried out using the combined data of the standard curves generated by diluted standards per experimental run (Appendix B 7.2.4 & 7.2.5).

The measured Cq values and melting temperatures of the standards as well as the test genomic control sample are presented in the table below (Table 3-2).

**Table 3-2: Quantification of mtDNA template of the synthetic oligonucleotide by a single round of PCR.**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Type,</th>
<th>Cq</th>
<th>Quantity (mtDNA molecules)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oligo-Dilution 1</td>
<td>Standard</td>
<td>N/A*</td>
<td>3.01X10(^{12})</td>
</tr>
<tr>
<td>Oligo-Dilution 2</td>
<td>Standard</td>
<td>N/A*</td>
<td>3.01X10(^{11})</td>
</tr>
<tr>
<td>Oligo-Dilution 3</td>
<td>Standard</td>
<td>N/A*</td>
<td>3.01X10(^{10})</td>
</tr>
<tr>
<td>Oligo-Dilution 4</td>
<td>Standard</td>
<td>7.888</td>
<td>3.01X10(^{9})</td>
</tr>
<tr>
<td>Oligo-Dilution 5</td>
<td>Standard</td>
<td>11.433</td>
<td>3.01X10(^{8})</td>
</tr>
<tr>
<td>Oligo-Dilution 6</td>
<td>Standard</td>
<td>15.345</td>
<td>3.01X10(^{7})</td>
</tr>
<tr>
<td>Oligo-Dilution 7</td>
<td>Standard</td>
<td>19.564</td>
<td>3.01X10(^{6})</td>
</tr>
<tr>
<td>Oligo-Dilution 8</td>
<td>Standard</td>
<td>23.161</td>
<td>3.01X10(^{5})</td>
</tr>
<tr>
<td>Negative Control</td>
<td>Negative control</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Genomic-sample</td>
<td>Test sample</td>
<td>12.706</td>
<td>136,497,000</td>
</tr>
</tbody>
</table>

*Oligo; oligonucleotide, Cq; Cycle of quantification, Tm; melting temperature. Using the standard curve, the amount of mtDNA was calculated based on the Cq value. * The first 3 standard dilutions were highly concentrated and therefore the Cq peaked so early and mtDNA could not be calculated.

As shown in the figures and the table above, dilutions 1, 2 and 3 had very high concentrations of the oligonucleotide and reached the plateau even before 3 cycles
were completed. Based on the Cq values and the mtDNA quantities in each oligonucleotide used as a standard, the corresponding mtDNA quantity in the test sample was quantified by the software indicating that there were 136,497,000 mtDNA molecules in a genomic sample.

3.3.1.1.2 Quantification of mtDNA template in split rounds of PCR

The optimisation of a mtDNA quantification method that would be suitable for lysed single cells and blastomeres required testing the efficacy of two lysis methods; PK and ALB lysis methods. Therefore, several rounds of PCR using the HRM master mix were used to amplify mtDNA on isolated single lymphocytes using both methods. The amplification of single cells using PK worked efficiently. Therefore, mtDNA quantities obtained from the normal oligonucleotide dilution were used as standards when PK lysis method was applied on single cells and embryos, since only a single round of PCR was performed.

On the other hand, optimisation of mtDNA quantification method that would be suitable for single cells and blastomeres lysed by ALB required two rounds of PCR reactions. This was because the addition of Tricine (a neutralizing agent required for the salts in ALB) to the HRM master mix, affected the reaction conditions and resulted in no amplification. When ALB lysis buffer was used, samples were amplified using 7 cycles of PCR to avoid the effect of Tricine on the HRM master mix. An aliquot from each amplified product was then added to the HRM master mix. Absolute quantification of mtDNA copies in the synthetic oligonucleotide dilution series was also applied using two rounds of PCR. However, this approach resulted in the amplification of non-specific product (Figure 3–7 & Figure 3-8).
The non-specific product amplification is clearly observed.

The obtained Cq values were slightly different (Table 3-3) compared to the Cq values obtained when a single round of PCR was performed (Table 3-2). This also resulted in obtaining a lower Cq value from the test genomic sample. Hence, samples lysed by ALB were compared to the Cq values obtained from the standards amplified using two rounds of PCR.

**Figure 3–7**: Absolute quantification of normal oligonucleotide after performing a 7 cycle PCR. The non-specific product amplification is clearly observed.

**Figure 3–8**: HRM of the absolute quantification of normal oligonucleotides after 7 cycle PCR. The HRM shows two different melting profiles meaning that there is non-specific formed product.
Table 3-3 : Absolute quantification of normal oligonucleotide after performing a 7 cycle PCR.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Type</th>
<th>Cq</th>
<th>Quantity (mtDNA molecules)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oligo-Dilution 1</td>
<td>Standard</td>
<td>N/A*</td>
<td>Standard, 3.01X10^{12}</td>
</tr>
<tr>
<td>Oligo-Dilution 2</td>
<td>Standard</td>
<td>N/A*</td>
<td>Standard, 3.01X10^{11}</td>
</tr>
<tr>
<td>Oligo-Dilution 3</td>
<td>Standard</td>
<td>N/A*</td>
<td>Standard, 3.01X10^{10}</td>
</tr>
<tr>
<td>Oligo-Dilution 4</td>
<td>Standard</td>
<td>N/A*</td>
<td>Standard, 3.01X10^{9}</td>
</tr>
<tr>
<td>Oligo-Dilution 5</td>
<td>Standard</td>
<td>7.121</td>
<td>Standard, 3.01X10^{8}</td>
</tr>
<tr>
<td>Oligo-Dilution 6</td>
<td>Standard</td>
<td>10.938</td>
<td>Standard, 3.01X10^{7}</td>
</tr>
<tr>
<td>Oligo-Dilution 7</td>
<td>Standard</td>
<td>16.139</td>
<td>Standard, 3.01X10^{6}</td>
</tr>
<tr>
<td>Oligo-Dilution 8</td>
<td>Standard</td>
<td>19.583</td>
<td>Standard, 3.01X10^{5}</td>
</tr>
<tr>
<td>Normal genomic control</td>
<td>Test sample</td>
<td>10.851</td>
<td>38,696,749.12</td>
</tr>
<tr>
<td>Negative control SC1</td>
<td>Negative</td>
<td>24.033</td>
<td>31,084.19</td>
</tr>
<tr>
<td>Negative control-Main-Lab</td>
<td>Negative</td>
<td>23.773</td>
<td>35,778.23</td>
</tr>
<tr>
<td>Negative control SC2</td>
<td>Negative</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Negative control-post</td>
<td>Negative</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Oligo; oligonucleotide, SC1; negative control used when the first round of PCR was performed in the single cell room, SC2; negative control used when the second round of PCR was performed in the single cell laboratory, post; negative control used in the post PCR laboratory when the second round of PCR was performed. Using the standard curve, the amount of mtDNA was calculated based on the Cq value. * The first 3 standard dilutions were highly concentrated and therefore the Cq peaked so early and mtDNA could not be calculated. ♦ The presence of primer dimer was confirmed by gel electrophoresis indicating the absence of contamination.

Agarose gel separation confirmed that the smaller peaks were produced by primer-dimerization (Figure 3–9).

Figure 3–9: Gel electrophoresis confirming the primer-dimer formation when 2 sets of PCR reactions were performed. As shown in the figure above, all the dilutions showed a PCR product of (150bp) which is the size expected from the amplification. The amplification marked with green arrow had a different product size confirming that the different melting peak was due to the primer-dimer formation and is not a specific product due to contamination.
3.3.1.2 Results from embryos

3.3.1.3 Quantification of mtDNA in embryos lysed by PK

Whole embryos lysed by PK peaked earlier than single and two isolated blastomeres (Figure 3-10 A&B) indicating the presence of more mtDNA copies and confirming the efficiency of the PCR. There were slight variations in the melting peaks even between single blastomeres from different embryos which could be due to polymorphisms in the mitochondrial sequence.

A. Differences in amplification and melting peaks were observed between whole embryos & single blastomeres.

B. Differences in amplification and melting peaks were observed between 2 blastomeres and single blastomeres.

Figure 3–10: Quantification of embryonic samples lysed by PK showing differences in amplification based on cell numbers.
3.3.1.3.1 Absolute quantification of mtDNA in embryonic samples lysed by ALB

Embryos lysed by ALB also resulted in the formation of primer dimer which was seen in the analysed negative controls from the first round PCR reaction as shown in (Figure 3-11). However, this did not affect the quantification of the samples.

![Image](image.jpg)

**Figure 3–11: Absolute quantification of mtDNA in embryonic samples lysed by ALB.**

Similar results were obtained from TaqMan genotyping assays and samples were plotted on the generated standard curve from dilutions measured by TaqMan (Appendix B 7.2.5, 7.2.4). Single cells derived from different embryos gave different Cq value. This could be due to the difference in embryo qualities and developmental stages. In general, single cells from arrested embryos at the morula stage had fewer mtDNA copy numbers compared to single cells from embryos arrested at cleavage stage. Single cells isolated from a blastocyst (E1 from Family-F (mtDNA copy = 2.6x10^3) had fewer mtDNA copies compared to a single cell isolated from an embryo that arrested at cleavage stage divisions (E4 from Family-C (mtDNA copy= 1.2x10^5)).

Cells lysed by ALB (which required two rounds of PCR appeared to contain more mtDNA copies compared to embryos lysed by PK (which required a single round to be performed). So, it is not that there are more copies, but the methodology used affected the actual mtDNA copy number. The difference was one order of
magnitude. This was clearly identified when the Cq values of single cells from the same embryo E3 from Family-C were different. Mitochondrial DNA copy number was $9.7 \times 10^4$ copies for the single cell isolated by PK while showing $8.7 \times 10^5$ copies when another cell from the same embryo was lysed by ALB. Therefore, two standard curves were used for the identification of mtDNA copy number in embryos. One, for comparison of embryos lysed by PK, while the other for comparison of embryos lysed by ALB.

Although previous studies have shown that ALB lysis may have lower ADO rates for amplification of nDNA (Thornhill et al., 2001), as split PCR was required in order to use ALB for lysis, and mtDNA is present in multiple copies rather than a single template, all subsequent quantification analysis was carried out using PK lysis buffer. Further validation of mtDNA quantification technique on the small number of cells using both qPCR methods and the standard curve used for mtDNA quantification are listed in Appendix B 7.2.3 and 7.2.4.

3.3.1.3.2 Analysis of mtDNA quantification based on morphology on day 5

The ideal situation of comparing embryonic samples is according to the morphology on the day of collection and processing. However, since the grading information of some of the embryos was incomplete; those were compared according to the morphology of the previous day (Day 5 grading) to provide some preliminary results for statistical analysis to be performed. Thus, the analysis was performed according to day 5 morphology. Any significant finding was accounted for if $p < 0.01$.

Random effect regression analysis was performed. This is a test that recognizes the clustered nature of data taking into consideration that each couple has a number of embryos. Testing the random effect regression of log mtDNA copy number on the morphology, graded on day 5 based on the number of cells, showed no significant
differences. In general, single cells from embryos arrested at cleavage stages had more mtDNA template compared with single cells from embryos arrested at morula stages \( p=0.018 \). Moreover, embryos arrested at cleavage and morula had more mtDNA template compared with the embryos which developed to the blastocyst stage. Analysis of 21 single blastomeres showed that the average mtDNA copies was \( 4.9\times10^5 \) in embryos arrested at cleavage (\( n=21 \)). The average number of mtDNA copies in blastomeres (\( n=16 \)) collected from morula stage embryos was \( 3.1\times10^4 \). Mitochondrial DNA quantification in blastomeres (\( n=6 \)) collected from blastocysts showed an average of \( 5.3\times10^3 \) in human embryos (Table 3-4, Table 3-5).

Analysis of 2 cells from embryos arrested at cleavage based on day 5 grading showed more mtDNA copies compared with 2 cells tested from embryos arrested at the morula stages (\( n=26 \)), \( p=0.02 \). However, the number of observations was considered small and more samples need to be tested for validation.

Whole embryo analysis has shown that there was an average of \( 2.1\times10^5 \) in arrested at cleavage stage embryos (\( n=3 \)). Embryos arrested at the morula stage (\( n=3 \)) contained \( 5.1\times10^5 \) mtDNA copies while the blastocysts (\( n=3 \)) showed a higher number \( 9.7\times10^6 \). This was established using two different lysis methods PK and ALB.

3.3.1.3.3 Analysis of mtDNA template number based on morphology on the day of collection (day 6 / 7)

Analysis of whole embryos (including 2 abnormally fertilised (0 Pronuclei), 2 hatched blastocysts and 2 degenerated embryos) based on the morphology on the day of collection in this study showed, that there appears to be a difference in mtDNA copy number between whole embryos that developed to the blastocyst stages compared with the embryos which arrested at the early divisions suggesting that blastocysts
have more mtDNA copies compared with arrested, DG and abnormally fertilised embryos.

Quantification of mtDNA in single blastomeres, showed few variations among blastomeres from a single embryo and among embryos from a single patient (Table 3-3). Single cells from abnormally fertilised embryos had more mtDNA template compared with single cells derived from a normally fertilised degenerated embryo. This could be because those cells were from abnormally fertilised samples from the same couple while the normally fertilised embryo was from a different couple. Analysis of mtDNA copies in single cells from the same embryo that arrested at cleavage divisions were compared; not much difference was found. For example; embryo 5 from Family-4 mtDNA template had $4.1 \times 10^4$ in cell 1, and mtDNA template in the other cell was $4.8 \times 10^4$. However, single cells from the same embryo that developed to a hatched blastocyst as in the case in E1 from Family-F had a large difference in mtDNA copy number. The difference was one order of magnitude. It could be because at the hatched blastocyst stage cells are split into ICM and Trophectoderm and mtDNA in segregating differently. No significant difference was noticed when maternal age was taken into account.
Table 3-4: Absolute quantification of mtDNA in single cells, 2 cells and clump of embryonic cells.

<table>
<thead>
<tr>
<th>Couple</th>
<th>Maternal age</th>
<th>Embryo ID</th>
<th>Number of pronuclei</th>
<th>Number of cells</th>
<th>Cq</th>
<th>mtDNA template</th>
<th>Day of collection</th>
<th>Day 5 Morphology</th>
<th>Day 6/7 Morphology</th>
<th>RT-PCR &amp; Lysis buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family-1</td>
<td>42</td>
<td>1</td>
<td>2PN</td>
<td>1</td>
<td>24.6</td>
<td>1.4 x 10^5</td>
<td>6</td>
<td>Arrested at cleavage</td>
<td>Pre-morula</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-1</td>
<td>42</td>
<td>7</td>
<td>2PN</td>
<td>2</td>
<td>24.5</td>
<td>1.43 x 10^5</td>
<td>6</td>
<td>Arrested at cleavage</td>
<td>Arrested at cleavage</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-2</td>
<td>35</td>
<td>4</td>
<td>2PN</td>
<td>1</td>
<td>23.6</td>
<td>2.4 x 10^5</td>
<td>7</td>
<td>Arrested at cleavage</td>
<td>Arrested at cleavage</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-2</td>
<td>35</td>
<td>6</td>
<td>2PN</td>
<td>1</td>
<td>20.7</td>
<td>1.4 x 10^6</td>
<td>7</td>
<td>Arrested at cleavage</td>
<td>Pre-morula</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-2</td>
<td>35</td>
<td>7</td>
<td>2PN</td>
<td>1</td>
<td>19.8</td>
<td>2.4 x 10^6</td>
<td>7</td>
<td>Arrested at cleavage</td>
<td>Hatched blastocyst</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-2</td>
<td>35</td>
<td>3</td>
<td>2PN</td>
<td>1</td>
<td>24.4</td>
<td>1.5 x 10^5</td>
<td>7</td>
<td>Morula</td>
<td>Hatched blastocyst</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-2</td>
<td>35</td>
<td>2</td>
<td>2PN Clump</td>
<td>1</td>
<td>22.4</td>
<td>4.9 x 10^5</td>
<td>7</td>
<td>Morula</td>
<td>Pre-morula</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-4</td>
<td>36</td>
<td>5</td>
<td>2PN</td>
<td>1</td>
<td>26.6</td>
<td>4.1 x 10^4</td>
<td>7</td>
<td>Arrested at cleavage</td>
<td>Arrested at cleavage</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-4</td>
<td>36</td>
<td>5</td>
<td>2PN</td>
<td>1</td>
<td>26.332</td>
<td>4.8 x 10^4</td>
<td>7</td>
<td>Arrested at cleavage</td>
<td>Arrested at cleavage</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-5</td>
<td>29</td>
<td>7</td>
<td>2PN</td>
<td>1</td>
<td>26.124</td>
<td>5.33 x 10^4</td>
<td>7</td>
<td>NA</td>
<td>Blastocyst</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-6</td>
<td>28</td>
<td>1</td>
<td>2PN</td>
<td>1</td>
<td>28.03</td>
<td>1.7 x 10^4</td>
<td>7</td>
<td>Arrested at cleavage</td>
<td>Pre-morula</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-6</td>
<td>28</td>
<td>3</td>
<td>2PN</td>
<td>1</td>
<td>25.7</td>
<td>7.1 x 10^4</td>
<td>7</td>
<td>Arrested at cleavage</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-B</td>
<td>37</td>
<td>9</td>
<td>2PN</td>
<td>1</td>
<td>21.81</td>
<td>1.4 x 10^5</td>
<td>6</td>
<td>Morula</td>
<td>N/A</td>
<td>Resolight ALB</td>
</tr>
<tr>
<td>Family-B</td>
<td>37</td>
<td>8</td>
<td>2PN</td>
<td>1</td>
<td>28.214</td>
<td>3.2 x 10^3</td>
<td>6</td>
<td>Morula</td>
<td>N/A</td>
<td>Resolight ALB</td>
</tr>
<tr>
<td>Family-B</td>
<td>37</td>
<td>4</td>
<td>2PN</td>
<td>1</td>
<td>29.4</td>
<td>7.6 x 10^3</td>
<td>6</td>
<td>DG</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-B</td>
<td>37</td>
<td>7</td>
<td>2PN</td>
<td>1</td>
<td>33.32</td>
<td>740</td>
<td>6</td>
<td>Morula</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-B</td>
<td>37</td>
<td>7</td>
<td>2PN</td>
<td>1</td>
<td>36.743</td>
<td>96</td>
<td>6</td>
<td>Morula</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-B</td>
<td>37</td>
<td>9</td>
<td>2PN</td>
<td>1</td>
<td>36.6</td>
<td>110</td>
<td>6</td>
<td>Morula</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-B</td>
<td>37</td>
<td>8</td>
<td>2PN</td>
<td>1</td>
<td>36.96</td>
<td>85</td>
<td>6</td>
<td>Morula</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-B</td>
<td>37</td>
<td>8</td>
<td>2PN</td>
<td>1</td>
<td>33.1</td>
<td>850</td>
<td>6</td>
<td>Morula</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-B</td>
<td>37</td>
<td>9</td>
<td>2PN</td>
<td>1</td>
<td>29.6</td>
<td>6.9 x 10^3</td>
<td>6</td>
<td>Morula</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-B</td>
<td>37</td>
<td>9</td>
<td>2PN</td>
<td>2</td>
<td>20.81</td>
<td>1.8 x 10^5</td>
<td>6</td>
<td>Morula</td>
<td>N/A</td>
<td>Resolight ALB</td>
</tr>
<tr>
<td>Family-B</td>
<td>37</td>
<td>4</td>
<td>2PN</td>
<td>2</td>
<td>38.7</td>
<td>30.4</td>
<td>6</td>
<td>DG</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-B</td>
<td>37</td>
<td>9</td>
<td>2PN</td>
<td>2</td>
<td>29.01</td>
<td>96</td>
<td>6</td>
<td>Morula</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-B</td>
<td>39</td>
<td>14</td>
<td>2PN</td>
<td>1</td>
<td>23.6</td>
<td>3.95 x 10^4</td>
<td>6</td>
<td>Morula</td>
<td>N/A</td>
<td>Resolight ALB</td>
</tr>
<tr>
<td>Family-C</td>
<td>39</td>
<td>2</td>
<td>2PN</td>
<td>1</td>
<td>28.64</td>
<td>1.2 x 10^4</td>
<td>6</td>
<td>Arrested at cleavage</td>
<td>N/A</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-C</td>
<td>39</td>
<td>11</td>
<td>2PN</td>
<td>1</td>
<td>27.62</td>
<td>2.2 x 10^4</td>
<td>6</td>
<td>Morula</td>
<td>N/A</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-C</td>
<td>39</td>
<td>3</td>
<td>2PN</td>
<td>1</td>
<td>24.7</td>
<td>1.3 x 10^5</td>
<td>6</td>
<td>Arrested at cleavage</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-C</td>
<td>39</td>
<td>14</td>
<td>2PN</td>
<td>1</td>
<td>27.8</td>
<td>1.98 x 10^4</td>
<td>6</td>
<td>Morula</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-C</td>
<td>39</td>
<td>6</td>
<td>2PN</td>
<td>1</td>
<td>31.132</td>
<td>2.71 x 10^3</td>
<td>6</td>
<td>Blastocyst</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-C</td>
<td>39</td>
<td>3</td>
<td>2PN</td>
<td>2</td>
<td>21.45</td>
<td>1.3 x 10^5</td>
<td>6</td>
<td>Arrested at cleavage</td>
<td>N/A</td>
<td>Resolight ALB</td>
</tr>
</tbody>
</table>

Cq: cycle of quantification; which corresponds to the number of mtDNA molecules. ABF: abnormally fertilised embryo, PN: pronuclei; which corresponds to the number of pronuclei in the embryo after fertilisation. N/A: not available.
<table>
<thead>
<tr>
<th>Couple</th>
<th>Maternal age</th>
<th>Embryo ID</th>
<th>Number of pronuclei</th>
<th>Number of cells</th>
<th>Cq</th>
<th>mtDNA template</th>
<th>Day of collection</th>
<th>Day 5 Morphology</th>
<th>Day 6/7 Morphology</th>
<th>RT-PCR &amp; Lysis buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family-C</td>
<td>39</td>
<td>3</td>
<td>2PN</td>
<td>2</td>
<td>21.3</td>
<td>9.7 x 10^5</td>
<td>6</td>
<td>Arrested at cleavage</td>
<td>N/A</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-C</td>
<td>39</td>
<td>11</td>
<td>2PN</td>
<td>2</td>
<td>29.71</td>
<td>6.3 x 10^3</td>
<td>6</td>
<td>Morula</td>
<td>N/A</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-D</td>
<td>37</td>
<td>8</td>
<td>2PN</td>
<td>1</td>
<td>20.5</td>
<td>2.14 x 10^5</td>
<td>6</td>
<td>Pre-morula</td>
<td>N/A</td>
<td>Resolight ALB</td>
</tr>
<tr>
<td>Family-D</td>
<td>37</td>
<td>7</td>
<td>2PN</td>
<td>1</td>
<td>23.643</td>
<td>3.84 x 10^4</td>
<td>6</td>
<td>Morula</td>
<td>N/A</td>
<td>Resolight ALB</td>
</tr>
<tr>
<td>Family-D</td>
<td>37</td>
<td>8</td>
<td>2PN</td>
<td>1</td>
<td>25.634</td>
<td>7.134 x 10^4</td>
<td>6</td>
<td>Pre-morula</td>
<td>N/A</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-D</td>
<td>37</td>
<td>7</td>
<td>2PN</td>
<td>1</td>
<td>25.71</td>
<td>6.8 x 10^4</td>
<td>6</td>
<td>Pre-morula</td>
<td>N/A</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-D</td>
<td>37</td>
<td>8</td>
<td>2PN</td>
<td>1</td>
<td>25.5</td>
<td>7.8 x 10^4</td>
<td>6</td>
<td>Morula</td>
<td>N/A</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-D</td>
<td>37</td>
<td>8</td>
<td>2PN</td>
<td>2</td>
<td>25.3</td>
<td>8.7 x 10^4</td>
<td>6</td>
<td>Pre-morula</td>
<td>N/A</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-E</td>
<td>23</td>
<td>9</td>
<td>2PN</td>
<td>1</td>
<td>30.9</td>
<td>3.2 x 10^3</td>
<td>7</td>
<td>Morula</td>
<td>Blastocyst</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-E</td>
<td>23</td>
<td>7</td>
<td>2PN</td>
<td>1</td>
<td>29.531</td>
<td>7.02 x 10^3</td>
<td>7</td>
<td>Morula</td>
<td>Morula</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-E</td>
<td>23</td>
<td>7</td>
<td>2PN</td>
<td>1</td>
<td>26.8</td>
<td>3.6 x 10^4</td>
<td>7</td>
<td>Morula</td>
<td>Morula</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-F</td>
<td>40</td>
<td>5</td>
<td>2PN</td>
<td>1</td>
<td>37.83</td>
<td>50</td>
<td>6</td>
<td>Arrested at cleavage</td>
<td>Degenerated</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-F</td>
<td>40</td>
<td>1</td>
<td>2PN</td>
<td>1</td>
<td>31.14</td>
<td>2.7 x 10^3</td>
<td>6</td>
<td>Blastocyst</td>
<td>Hatched</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-F</td>
<td>40</td>
<td>1</td>
<td>2PN</td>
<td>1</td>
<td>25.3</td>
<td>8.703 x 10^3</td>
<td>6</td>
<td>Blastocyst</td>
<td>Hatched</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-F</td>
<td>40</td>
<td>10</td>
<td>2PN</td>
<td>1</td>
<td>25.2</td>
<td>9.504 x 10^4</td>
<td>6</td>
<td>Blastocyst</td>
<td>Hatched</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-F</td>
<td>40</td>
<td>1</td>
<td>2PN</td>
<td>2</td>
<td>20.6</td>
<td>1.44 x 10^6</td>
<td>6</td>
<td>Blastocyst</td>
<td>Hatched</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-F</td>
<td>40</td>
<td>ABF1</td>
<td>0PN</td>
<td>2</td>
<td>24.98</td>
<td>1.1 x 10^5</td>
<td>6</td>
<td>Arrested at cleavage</td>
<td>Arrested at cleavage</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-F</td>
<td>40</td>
<td>1</td>
<td>2PN</td>
<td>3</td>
<td>25.8</td>
<td>6.6 x 10^4</td>
<td>6</td>
<td>Blastocyst</td>
<td>Hatched</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-G</td>
<td>31</td>
<td>10</td>
<td>2PN</td>
<td>1</td>
<td>28.2</td>
<td>1.6 x 10^6</td>
<td>7</td>
<td>NA</td>
<td>NA</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-G</td>
<td>31</td>
<td>10</td>
<td>2PN</td>
<td>1</td>
<td>32.4</td>
<td>1.3 x 10^3</td>
<td>7</td>
<td>NA</td>
<td>NA</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-G</td>
<td>31</td>
<td>12</td>
<td>2PN</td>
<td>2</td>
<td>27.142</td>
<td>2.9 x 10^4</td>
<td>7</td>
<td>Arrested at cleavage</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-G</td>
<td>31</td>
<td>10</td>
<td>2PN</td>
<td>2</td>
<td>27.9</td>
<td>1.9 x 10^4</td>
<td>7</td>
<td>NA</td>
<td>NA</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-I</td>
<td>41</td>
<td>1</td>
<td>2PN</td>
<td>1</td>
<td>22.623</td>
<td>4.2 x 10^6</td>
<td>6</td>
<td>Arrested at cleavage</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-I</td>
<td>41</td>
<td>5</td>
<td>2PN</td>
<td>1</td>
<td>25.74</td>
<td>6.7 x 10^4</td>
<td>6</td>
<td>Arrested at cleavage</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-I</td>
<td>41</td>
<td>6</td>
<td>2PN</td>
<td>1</td>
<td>24.243</td>
<td>1.6 x 10^5</td>
<td>6</td>
<td>Arrested at cleavage</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-I</td>
<td>41</td>
<td>12</td>
<td>2PN</td>
<td>1</td>
<td>32.5</td>
<td>1.2 x 10^3</td>
<td>6</td>
<td>Arrested at cleavage</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-I</td>
<td>41</td>
<td>2</td>
<td>2PN</td>
<td>1</td>
<td>24.51</td>
<td>1.3 x 10^5</td>
<td>6</td>
<td>Arrested at cleavage</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-I</td>
<td>41</td>
<td>7</td>
<td>2PN</td>
<td>1</td>
<td>25.5</td>
<td>7.7 x 10^4</td>
<td>6</td>
<td>Pre-morula</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-J</td>
<td>36</td>
<td>ABF4</td>
<td>0PN</td>
<td>1</td>
<td>23.31</td>
<td>2.9 x 10^6</td>
<td>6</td>
<td>Not available.</td>
<td>Degenerated</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-j</td>
<td>36</td>
<td>ABF5</td>
<td>0PN</td>
<td>1</td>
<td>19.333</td>
<td>3.03 x 10^8</td>
<td>6</td>
<td>Not available.</td>
<td>Degenerated</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-K</td>
<td>34</td>
<td>ABF1</td>
<td>3PN</td>
<td>1</td>
<td>24.8</td>
<td>1.1 x 10^5</td>
<td>6</td>
<td>NA</td>
<td>NA</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-K</td>
<td>34</td>
<td>ABF3</td>
<td>3PN</td>
<td>1</td>
<td>29.93</td>
<td>5.5 x 10^3</td>
<td>6</td>
<td>Not available.</td>
<td>Arrested at cleavage</td>
<td>TaqMan PK</td>
</tr>
</tbody>
</table>

Cq: cycle of quantification; which corresponds to the number of mtDNA molecules. ABF: abnormally fertilised embryo, PN: pronuclei; which corresponds to the number of pronuclei in the embryo after fertilisation. N/A: not available.
Table 3-5: Absolute quantification of mtDNA in whole embryos.

<table>
<thead>
<tr>
<th>Couple</th>
<th>Maternal age</th>
<th>Embryo ID</th>
<th>Number of pronuclei</th>
<th>Number of cells</th>
<th>Cq</th>
<th>mtDNA template</th>
<th>Day of collection</th>
<th>Day 5 Morphology</th>
<th>Day 6/7 Morphology</th>
<th>RT-PCR &amp; Lysis buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family-2</td>
<td>35</td>
<td>ABF1</td>
<td>0PN</td>
<td>Whole</td>
<td>24.5</td>
<td>1.44 x 10^5</td>
<td>7</td>
<td>N/A</td>
<td>Arrested at cleavage</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-2</td>
<td>35</td>
<td>ABF2</td>
<td>0PN</td>
<td>Whole</td>
<td>20.8</td>
<td>1.3 x 10^6</td>
<td>7</td>
<td>N/A</td>
<td>Undivided</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-A</td>
<td>42</td>
<td>2</td>
<td>2PN</td>
<td>Whole</td>
<td>28.5</td>
<td>1.3 x 10^4</td>
<td>6</td>
<td>Arrested at cleavage</td>
<td>N/A</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-A</td>
<td>42</td>
<td>4</td>
<td>2PN</td>
<td>Whole</td>
<td>22.23</td>
<td>5.4 x 10^5</td>
<td>6</td>
<td>Arrested at cleavage</td>
<td>N/A</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-A</td>
<td>42</td>
<td>3</td>
<td>2PN</td>
<td>Whole</td>
<td>23.53</td>
<td>2.5 x 10^5</td>
<td>6</td>
<td>Morula</td>
<td>N/A</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-A</td>
<td>42</td>
<td>5</td>
<td>2PN</td>
<td>Whole</td>
<td>17.53</td>
<td>8.9 x 10^6</td>
<td>6</td>
<td>Blastocyst</td>
<td>N/A</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-B</td>
<td>37</td>
<td>5</td>
<td>2PN</td>
<td>Whole</td>
<td>23.922</td>
<td>1.9 x 10^5</td>
<td>6</td>
<td>Morula</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-C</td>
<td>39</td>
<td>1</td>
<td>2PN</td>
<td>Whole</td>
<td>18.83</td>
<td>5.61 x 10^6</td>
<td>6</td>
<td>Morula</td>
<td>N/A</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-C</td>
<td>39</td>
<td>4</td>
<td>2PN</td>
<td>Whole</td>
<td>18.7</td>
<td>4.42 x 10^6</td>
<td>6</td>
<td>Blastocyst</td>
<td>Hatched</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-F</td>
<td>40</td>
<td>6</td>
<td>2PN</td>
<td>Whole</td>
<td>25.2</td>
<td>9.3 x 10^4</td>
<td>6</td>
<td>Arrested at cleavage</td>
<td>DG</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-F</td>
<td>40</td>
<td>4</td>
<td>2PN</td>
<td>Whole</td>
<td>16.5</td>
<td>1.6 x 10^7</td>
<td>6</td>
<td>Blastocyst</td>
<td>Hatched</td>
<td>Resolight PK</td>
</tr>
<tr>
<td>Family-G</td>
<td>31</td>
<td>11</td>
<td>2PN</td>
<td>Whole</td>
<td>26.61</td>
<td>3.99 x 10^4</td>
<td>7</td>
<td>Pre-morula</td>
<td>DG</td>
<td>TaqMan PK</td>
</tr>
<tr>
<td>Family-H</td>
<td>39</td>
<td>ABF3</td>
<td>0PN</td>
<td>Whole</td>
<td>21.144</td>
<td>1.03 x 10^5</td>
<td>7</td>
<td>N/A</td>
<td>N/A</td>
<td>TaqMan PK</td>
</tr>
</tbody>
</table>

Whole: whole embryo. Cq: cycle of quantification; which corresponds to the number of mtDNA molecules. ABF: abnormally fertilised embryo, PN: pronuclei; which corresponds to the number of pronuclei in the embryo after fertilisation. N/A: not available.
### 3.3.2 Relative quantification by TaqMan copy number assays

Embryos which had been whole genome amplified for further NGS analysis (discussed in subsequent chapters 4 and 5) were used in this assay after normalisation and dilution to 5ng/µl. A comparison was made between all embryos in general. Embryos that developed to the blastocyst stage were compared with embryos that degenerated or arrested at cleavage. Relative quantification analysis was performed using the $2^{-\Delta\Delta CT}$ (User Bulletin #2; Applied Biosystems). The two groups were compared based on the non-parametric Mann Whitney U test with 95% confidence interval (i.e. $p<0.05$ was considered statistically significant). No significant difference in mtDNA copy number was found when embryos were compared based on their morphology (Figure 3–12). Even though blastocysts seemed to have more mtDNA copy number, it was not statistically significant $p=0.229$. However, clinically it may still be a useful measure considering which embryos may be suitable for transfer.

![mtDNA copy number comparison between embryos based on morphology](image)

**Figure 3–12: mtDNA copy number comparison between embryos based on morphology.**

No significant differences in mtDNA copy number were found between embryos that developed to the blastocysts stage (n= 18) and embryos that arrested and/or degenerated at cleavage (n= 36) on the day of collection (6–7 post fertilisation). The median of the blastocysts=0.991, minimum=0.51, maximum=1.38., Median of the DG/arrested embryos =0.9239, minimum=0.56, maximum=1.78.
A similar comparison was performed based on maternal age and no significant associations were found between maternal age and embryo morphology (p value ranged between p=0.396 and p=0.447).

![Figure 3-13: Relative quantification comparison between embryos based on maternal age](image)

No significant differences in mtDNA copy number were found between embryos that developed to the blastocysts stage (n=18) and embryos that arrested and/or degenerated at cleavage (n=36) on the day of collection when maternal age was considered. Blastocysts (n=5) from women older than 35 showed a median=0.98, minimum=0.73, maximum=1.71. The arrested embryos (n=15) were =0.97, minimum=0.57, maximum=1.27. Blastocysts (n=9) from women younger than 35 showed a median=0.943, minimum=0.76, maximum=2.05. On the other hand, median of the DG/arrested embryos (n=22) was =0.985, minimum=0.59, maximum=1.53. *o: outliers.

3.3.3 Aneuploidy detection by GenetiSure a-CGH

An aliquot was taken from embryos which were WGA and tested in relative quantification and was used to assess the aneuploidy status according to the method described in section 2.4.1.7. Of the 54 embryos for which the aneuploidy status was determined, 12 (21.43%) were euploid.
### Table 3-6: The number of euploid and aneuploid embryos.

<table>
<thead>
<tr>
<th>Aneuploidy status</th>
<th>Number of embryos</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Blastocysts</td>
<td>Arrested/DG</td>
</tr>
<tr>
<td><strong>Euploid</strong></td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td><strong>Aneuploid</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple aneuploidies &gt; 5</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>Aneuploid ≤ 5</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>Single aneuploid</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>

The number and percentage of PGD embryos within the donated sample with each form of aneuploidy.

CytoGenomics software reports for embryos that were euploid or had a single aneuploidy are shown in Figure 3-14.

**Figure 3–14:** Array-CGH traces from cytoreports, generated by the CytoGenomics Software. An example of a euploid and aneuploid embryonic samples.
An example of the scoring of aneuploidy ≤5 and multiple aneuploidies is presented in Figure 3-15. The aneuploidy data for all the embryos is displayed in Table 7-3 presented in Appendix B section 7.2.3.

The non-parametric Mann Whitney U test was performed for the comparison of mtDNA template number between euploid and aneuploid embryos. A significant difference was found between the euploid and aneuploid embryos according to their mtDNA copy number (p=0.004) suggesting that the euploid embryos had more mtDNA copies compared with aneuploid embryos collected on the same day post fertilisation regardless of their morphology (Figure 3-16).
Since the distributions are not normal between the two groups, the median, minimum and maximum differences were compared. The median=0.9931, minimum= 0.81 and maximum= 1.22 for euploid (n=12). For aneuploid embryos (n= 42) median=0.8050, minimum= 0.45 and maximum= 1.57. (p=0.004). o: outlier.

Samples were normalised to RNAse P (nuclear reference gene) to exclude the effect of differences in cell number and the WGA process.

To confirm the association of euploidy with mtDNA number, aneuploid embryos with losses and gains on chromosome 14 including the RNase P gene were excluded from the analysis (n=20). Differences between both groups was not affected by losses and gains on RNAse P gene located on chromosome 14. The results were still significant p=0.008 (Figure 3-17). More detailed analysis confirmed that only euploid arrested at cleavage embryos had greater mtDNA copy number p=0.002 while the blastocysts euploid/aneuploid comparisons were not significant (Figure 3–18).
Figure 3–17: Comparison between euploid and aneuploid embryos based on mtDNA copy number excluding embryos with losses and gains on RNAse P.

Since the distributions are not normal between the two groups, the median, minimum and maximum differences were compared. The median=0.993, minimum= 0.81 and maximum= 1.22 for euploid embryos (n=12). For aneuploid embryos (n=24) the median=0.8156, minimum= 0.45 and maximum= 1.16. (p=0.008).

Figure 3–18: Comparison of mtDNA (normalised to RNAse P) template number between euploid and aneuploid arrested and blastocysts embryonic samples.

On the left, a comparison between euploid (n=6) and aneuploid (n=14) arrested at cleavage and/or degenerated embryos is illustrated by the box plots. Since the distributions are not normal between the two groups, the median, minimum and maximum differences were compared. The median=1.0096, minimum= 0.91 and maximum= 1.11. The aneuploid DG/arrested embryos median=0.704, minimum= 0.53 and maximum= 0.99. A significant difference was found (p=0.002). On the right, a comparison between euploid blastocysts(n=6) and aneuploid blastocysts (n=10) is illustrated by the box plots. Since the distributions are not normal between the two groups, the median, minimum and maximum differences were compared. The median=0.952, minimum= 0.89 and maximum= 1.34 for euploid blastocysts. The aneuploidy blastocyst median=0.99, minimum= 0.50 and maximum= 1.27. No significant difference was found (p= 0.953). *: outlier.
Further analysis, has shown that arrested aneuploid embryos had greater copy template numbers compared with the aneuploid embryos which reached the blastocyst stages $p=0.001$. Euploid embryos comparisons showed a similar pattern but did not reach statistical significance $p=0.126$.

![Box plots comparing mtDNA template numbers between euploid and aneuploid arrested and blastocysts embryonic samples.](image)

**Figure 3–19:** Comparison of mtDNA (normalised to RNAse P) template number between euploid and aneuploid arrested and blastocysts embryonic samples.

On the left, a comparison between aneuploid blastocysts ($n=10$) and arrested at cleavage and/or degenerated embryos ($n=14$) is illustrated by the box plots. Since the distributions are not normal between the two groups, the median, minimum and maximum differences were compared. The median=1.1, minimum= 0.55 and maximum= 1.40 for aneuploid blastocysts. The aneuploid DG/arrested embryos median=1.6, minimum= 1.11 and maximum= 2.1. A significant difference was found ($p=0.001$). On the right, a comparison between euploid blastocysts ($n=6$) and arrested at cleavage and/or degenerated embryos ($n=6$) is illustrated by the box plots. Since the distributions are not normal between the two groups, the median, minimum and maximum differences were compared. The median=0.953, minimum= 0.89 and maximum= 1.34 for euploid blastocysts. The euploid DG/arrested embryos median=1.2, minimum= 1.1 and maximum= 1.31. No significant difference was found ($p=0.126$). *: outlier.

### 3.3.4 Analysis of ATP levels with aneuploidy in embryos

Forty-six frozen thawed blastocysts were analysed. ATP levels in embryos were calculated based on the concentration of the standard used as described in the Appendix B section 7.2.2. As shown in Figure 3–20 euploid blastocysts ($n=5$) showed a stable level of ATP which ranged between (1-3.4pM). Blastocysts with more than two aneuploid chromosomes had ATP levels with greater variance. ATP levels started to fall in blastocysts which had more than 10 aneuploid chromosomes.
A total of 46 frozen / thawed blastocysts were tested in this analysis: 29 from the PGS group and 15 from the PGD group. This included euploid embryos (n=5), embryos with multiple aneuploid chromosomes <5 (n=28) and multiple aneuploid chromosomes ≥5 (n=13). As shown in the figure, the ATP level in euploid embryos ranged between (1-3.4 pM) compared to embryos showing > 2 chromosome aneuploidies which had a greater variance in ATP levels. ATP levels started to fall in embryos with > 10 chromosome aneuploidies.

3.4 Discussion

Quantification of mitochondrial DNA copy number in human embryos is very challenging. Recently, several reports were published, some suggesting the possibility of using mtDNA as a predictor for successful implantation (Diez-Juan et al., 2015, Fragouli et al., 2015, Bayram et al., 2017, Fragouli et al., 2017, Ravichandran et al., 2017, Wells et al., 2017). These reports suggested that embryos can be assessed for mtDNA number at later stages of development to determine embryo quality or implantation potential respectively. Overall, embryos with lower mtDNA copy number showed better implantation outcomes (Diez-Juan et al., 2015, Fragouli et al., 2015, Fragouli et al., 2017, Ravichandran et al., 2017, Wells et al., 2017). Other groups showed no association of mtDNA number with
embryo quality (Gianaroli et al., 2015, Victor et al., 2017). Clinical trials are ongoing to determine the clinical benefit of this kind of assessment.

Many possible reasons were proposed for those differences in results between the groups. Some criticised the techniques used as they lacked normalisation to an internal reference, others pointed to the WGA process which may affect aneuploidy scoring and more recently, it has been reported that different clinics showed different results. The aim of this study was to calculate mtDNA copy number in human embryos post PGD. Embryos that developed to blastocysts or arrested at earlier stages were tested in order to correlate mtDNA copy number with morphology, aneuploidy status and indirectly with the ATP level using two approaches. The first method was by measuring mtDNA template in embryonic samples by Resolight and TaqMan genotyping assay compared with a synthetic oligonucleotide. The other test was by the relative quantification to an internal nuclear gene using the TaqMan copy number assays. Since the absolute quantification method required the full sample to be processed, there was no direct method to compare both methods using the same sample.

3.4.1 Absolute quantification of mtDNA copy number

Due to difficulties in obtaining human embryos, the PCR conditions were first optimised using a single stranded custom designed synthetic oligonucleotide that was intended to be used as a template corresponding to a region in the \textit{MT-ATP6} gene in the heavy strand of mitochondrial genome. According to the literature, the number of mtDNA molecules in a mammalian oocyte is approximately between $10^5$ and $10^6$ molecules (Cummins, 1998, Reynier et al., 2001). Therefore, using a specific calculation, we were able to establish that diluting the synthetic
oligonucleotide, eight times, would provide a sample that is most likely to represent a single blastomere.

Using HRM, the mtDNA was accurately quantified based on the normal oligonucleotide dilutions. Embryonic samples were compared to normal oligonucleotide template dilutions. To standardize the protocol, we then tried to quantify mtDNA using the HRM reaction on single isolated lymphocytes. Since there are two methods that can be used to lyse the cells, PK and ALB lysis methods were tested. When PK was applied on single cells and different number of embryonic cells, the real-time PCR reaction worked efficiently. However, lysing the cells using ALB which required adding Tricine (a neutralizing agent necessary for the salts found in the alkaline lysis buffer (NaOH) (Cui X et al., 1989), affected the HRM PCR reaction resulting in no amplification. Therefore, two rounds of PCR were needed to optimize the reaction conditions for real time single cell amplification using ALB. A nonspecific product always appeared whilst the first round of PCR negative controls was tested. The handling process, the cleanliness of the lab, and the primer design were suspected. Excluding all these possibilities and ordering another set of primers, we concluded that the nonspecific amplified product was a result of primer-dimer formation in the reaction due to the increased number of cycles used. Since the Cq value of this nonspecific amplified product was consistent in all reactions showing a different melting peak, the nonspecific primer-dimer amplification was disregarded.

Various attempts were made to optimize the reaction for the TaqMan genotyping assay, resulting in a deviation from the advised minimal amplification from 5 to 1 and increasing the number of cycles to 60. The absolute quantification of mtDNA molecules was determined from single cells as well as diluted templates of the oligonucleotide (section 3.2.11) (dilution 7 and dilution 8) suggesting that the
TaqMan assay was suitable to quantify mtDNA in single cells. In order to estimate the average value of mtDNA copies required for embryos to develop to the blastocyst stage, the analysis was performed by comparing the mtDNA copy number in whole embryos and tubed single cells separately.

Based on the analysis of whole embryos, blastocyst stage embryos had more mtDNA template compared with embryos arrested at cleavage divisions (P=0.05). Although many reports suggest that mtDNA replication does not occur until after implantation, the developmental significance of increased mtDNA number was reported by Spiking et al., 2007 when mtDNA replication was inhibited at various times during the 3 days of porcine oocyte maturation in vitro (Spikings et al., 2006, Van Blerkom, 2011). This, supports the possibility that mtDNA replication could be happening at the early stages of embryo development. However, no conclusion regarding replication can be drawn from our findings as these embryos were tested when arrested at day 5, not when they were at cleavage division stages.

Although measuring mtDNA copy number in single cells is not an accurate way to estimate the total number of mtDNA in an embryo, it was used to evaluate the possibility of comparing mtDNA copies by testing single cells from each sample since it would show if a single cell from a whole embryo is representative for future clinical application. Single cells from arrested embryos had greater mtDNA copy number compared with single cells from morula and blastocyst. This is not surprising since cells from cleavage embryos are bigger and have more mtDNA P=0.018.

Testing two cells (tubed separately) from the same embryo; showed that there is not much difference in mtDNA template number between cells from the same embryo. Although there is not enough data to conclude that, others have found that
cells from the same embryo have more or less the same number of mtDNA copies (Smeets, 2013).

3.4.2 Relative Quantification of mtDNA by TaqMan copy number assays

Relative quantification to the nuclear RNAase P gene showed no significant difference in mtDNA copy number between blastocyst and embryos that degenerated or arrested at the cleavage stage. This supports what has been mentioned in literature, that mtDNA replication does not occur until after implantation (Piko and Taylor, 1987, Ebert et al., 1988, el Meziane et al., 1989, El Shourbagy et al., 2006, Kameyama et al., 2007) and contradicts our findings from the previous quantification approach. The differences in results between both techniques could be because at WGA it is not enough to reach significance. Also, different embryos were tested, and different tests have different sensitivities.

Despite the differences in both results, we still cannot confirm if replication is happening at cleavage divisions or not, since our embryos are arrested, and embryos collected and tested on day 3 should be considered to validate those results.

3.4.2.1 Aneuploidy and maternal age effect

Maternal age differences showed no association with embryonic mtDNA copy number. However, a significant difference was found with aneuploidy. Euploid embryos had more mtDNA compared with aneuploid embryos. More detailed analysis has shown that this is mainly a reflection of the group of samples arrested at cleavage rather than the blastocysts. Euploid embryos arrested at cleavage have significantly more mtDNA compared with the aneuploid embryos arrested at cleavage (p=0.002).
3.4.2.1.1 Mitochondrial DNA number compared to embryo aneuploidy status and morphology

Morphology is an indication of embryo quality, but aneuploidy and other factors may cause implantation failure. Recently, the MitoScore test, introduced by Igenomix suggests that fewer mtDNA is present in euploid cleavage stage embryos (tested on day 3) results in better implantation outcomes (Diez-Juan et al., 2015, Bayram et al., 2017). Others, have introduced the MitoGrade test which used mtDNA copy number for the same purpose by measuring the mtDNA copy number in trophectoderm samples from blastocysts and found that more mtDNA is present in aneuploid blastocysts (Fragouli et al., 2015, Fragouli et al., 2017, Ravichandran et al., 2017). Based on our analysis, no significant association was found between mtDNA copy number and aneuploidy in blastocysts. The differences in our results and others may be in the WGA amplification kit used which may have affected the results by introducing bias. Also, even though we used the same nuclear gene for normalisation, which is more likely to be compared to single cell lysis method than whole embryo method, normalisation to more than one nuclear template is better.

In general, our data complement the findings of other studies in that elevated levels of mitochondria suggests poor prognostic outcome by testing samples arrested at cleavage. The conclusion from our results is that euploid embryos arrested at cleavage have more mtDNA copies than aneuploid embryos arrested at this stage. Mitochondrial quantification might be a good predictor for embryo quality.

3.4.3 ATP

Assessment of ATP levels in embryos has shown that euploid blastocysts (n=5) had a stable ATP level which ranged between (1-3.4pM). This agrees with previous reports (Van Blerkom et al., 1995) which stated that the ATP content of normal-
appearing MII oocytes differed significantly between cohorts, but a higher potential for development and implantation was associated with an ATP concentration > 2pmol/oocyte. Blastocysts with more than two aneuploid chromosomes had ATP levels with greater variance than euploid blastocysts. ATP levels fell when blastocysts showing more than 10 aneuploid chromosomes were tested.

Mitochondrial functions are important for the formation of meiotic spindles and for maintenance of the MII spindle before fertilisation (Zhang et al., 2006). Insufficient ATP generation may result in aneuploidy. It has been reported that the mean ATP increased with developmental stage; zygotes had a significantly higher ATP content than oocytes and embryos had a significantly higher ATP content than both zygotes and oocytes. Within the embryo group, the ATP content was significantly lower in lower grade embryos with substantial fragmentation compared to poly-pronuclear embryos (Zhao and Li, 2012). Thus, reduced ATP content may be related to arrested division, fertilisation failure and abnormal embryonic development. The differences observed in oocyte and embryo ATP content between women who achieved a pregnancy compared with those who did not indicates that mitochondrial function is correlated with individual fertility (Zhao and Li, 2012). This could also be the case in our samples and the differences in ATP levels are because they were from different families.

Even though all our samples were blastocysts, the detected aneuploidy in embryos following PGS/PGD could be a result of low levels of ATP. The increase in glucose uptake at the blastocyst stage is accompanied by a substantial increase in ATP generation and O$_2$ consumption (Houghton and Leese, 2004), suggesting that OXPHOS also takes place (Dumollard et al., 2007, Ramalho-Santos 2009). Since both glycolysis and OXPHOS occur at the early stages of embryo development with
a substantial increase at the blastocyst stage, a change in ATP may occur as a response to stress. It has been suggested that abnormally high levels of mtDNA in aneuploid embryos at the blastocyst stage may be symptomatic of some form of stress that results in compensation for poor mitochondrial function (Fragouli et al., 2015). In our analysis, a greater variance in ATP levels were observed in aneuploid embryos, suggesting dysregulation of ATP synthesis or consumption compared with the euploid embryos. This possibility would be consistent with the 'quiet embryo hypothesis', proposed by Leese, which proposes that viable embryos have relatively lower or 'quiet' metabolism, while those under stress, and of reduced developmental potential, tend to be more metabolically active (Leese, 2002, Fragouli et al., 2015). If mtDNA copy number is a reflection of ATP levels in embryos, this would explain why aneuploid embryos have high variance in ATP levels.

### 3.5 Results summary

1. Absolute quantification showed that the average mtDNA copy number was $4.9 \times 10^5$ in single blastomeres from embryos arrested at cleavage ($n=21$). Single blastomeres ($n=16$) collected from morula stage embryos had an average of $3.1 \times 10^4$ mtDNA copies and an average of $5.3 \times 10^3$ mtDNA copies in single cells ($n=6$) collected from blastocysts.

2. Whole embryo analysis showed that blastocysts have more mitochondrial templates compared with degenerate and arrested embryos; however, this was not statistically significant.

3. Analysis of single blastomeres from the same embryo showed similar mtDNA numbers. However, single blastomeres from blastocysts showed a difference in mtDNA copy number (difference one order of magnitude).
4. No association was found between mtDNA copy number and the developmental stage of the embryo with maternal age by absolute quantification and/or relative quantification to RNAse P.

5. Euploid embryos arrested at cleavage had significantly greater mtDNA copy numbers than aneuploid arrested embryos by relative quantification to RNAse P.

6. Assessment of ATP levels in embryos, indicated that euploid blastocysts (n=5) had stable ATP level which ranged between (1-3.4pM) compared with aneuploid blastocysts where ATP levels were variable.

7. ATP levels in blastocysts with multiple aneuploidy >10 showed a massive decline.

3.6 Limitations and future work

The main limitation of this study was the small sample size due to the difficulty in obtaining a sufficient number of human embryos. For this reason, embryos were grouped according to the overall stage on the day of collection and were not subdivided further. If more samples were available, a more specific analysis by testing embryos collected and tested on day three could have been performed and compared to embryos at later stages of development in order to confirm if mitochondrial replication is happening at cleavage divisions.

With regards to the methodology, the absolute quantification method required the full sample to be processed, there was no direct method to compare qPCR methods using the same sample. For future mtDNA quantification tests in human embryos, a non-human template (NHT) could be used for normalisation in quantitative PCR analysis. This NHT should be normalised to several nuclear genes as well as a
mitochondrial gene to compare the ratios of NHT: mtDNA and NHT: nuclear genes measured in embryos. This may give a better indication of the overall amount of mtDNA and avoid the effect of artefacts introduced by WGA.

The ATP assay required processing of the whole embryo. Therefore, mtDNA number could not be directly correlated to the ATP level in the embryo and only blastocysts which had been frozen/thawed with previously identified aneuploidy data were available. Moreover, ATP could be a reflection of ATP produced by glycolysis or the ETC. Therefore, the ADP/ATP ratios would give a better and clearer understanding. In the future, embryos with previously identified mitochondrial numbers and aneuploidy data should be tested. The interpretation of Genetisure arrays was very difficult because of the sensitivity of the arrays which picked up very small segmental losses and gains and gave a generally noisy result. Therefore, only losses and gains above 50 Mb were considered.
Chapter 4 The effect of parental ethnicity on fertility and embryo quality

4.1 Aims summary

Recurrent miscarriage (RM), defined as the occurrence of three or more consecutive pregnancy failures is estimated to occur in 1-2% of women (Ford and Schust, 2009). Similarly, Repeated implantation failure (RIF) occurs when 1-2 embryos of good quality fail to implant following at least three consecutive IVF treatment cycles (Simon and Laufer, 2012). Carriers of a balanced translocation can produce unbalanced gametes, which can cause fertilization failure, implantation failure or embryo loss (Mau-Holzmann, 2005). An increased prevalence of chromosomal structural abnormalities has been documented in RIF patients (Stern et al., 1999, Raziel et al., 2002, De Sutter et al., 2012) suggesting the involvement of karyotype abnormalities in the pathogenesis of implantation failure.

On the other hand, approximately 50% of RM cases remain unexplained, and a significant proportion of these is suggested to be caused by as yet unknown genetic mechanisms (Carrington et al., 2005, Rai and Regan, 2006, Kaare et al., 2009).

Mitochondrial dysfunctions affecting ATP production during early developmental stages may be lethal for the embryo, even if only a portion of the mitochondria are dysfunctional. Consequently, it is possible that some mtDNA mutations cause developmental arrest before the pregnancy is clinically recognised (Van Blerkom et al., 1998, Van Blerkom, 2004). Recent data have shown that the amount of cellular ATP may be influenced by the mitochondrial haplogroup, whose variants mediate not only cell growth through the production of energy, but also cell signalling for the major molecular pathways, including those regulating chromosome segregation (Kenney et al., 2013).
Alleles present in population groups reflect adaptation to environmental conditions and could have an influence on disease susceptibility and embryo development at the preimplantation stages. Self-reported ancestry or ethnicity are considered to be biased and inaccurate compared to genotypically-determined ethnicity (Sucheston et al., 2012, Mersha and Abebe, 2015, Ramos et al., 2016), and the effect of ethnicity on fertility is debated between those who report a significant effect (Sharara and McClamrock, 2000, Palep-Singh et al., 2007, Purcell et al., 2007) and those who report no meaningful difference (Randolph et al., 2003, Matalliotakis et al., 2008, Dayal et al., 2009, Bhide et al., 2015). I hypothesised that differences in ethnicities between partners defined by the mitochondrial haplogroups or nuclear genes could constitute one of the unknown genetic factors affecting fertility. This study was therefore performed to explore the effect of ethnicity differences determined by mitochondrial haplogroups from a fertile (PGD) group and repeated miscarriage (RM) / repeated implantation failure (RIF) group of patients. Within the fertile group, the analysis was performed by assessing preimplantation embryo development along with aneuploidies and mitochondrial haplogroups. To obtain information about population groups defined by nuclear genes, genomic samples were genotyped using the Infinium® CoreExome-24 v1.2 Bead Chip and the HumanKaryomap-12 Bead Chip (Illumina, San Diego, CA) due to several hundred thousand single-nucleotide polymorphisms (SNPs) spread throughout the genome present on both chips. The aneuploidy status and mitochondrial DNA number of the embryonic samples, already analysed from the previous section, were used in this chapter for additional analysis. The analysis of SNPs was performed to:
Determine if there is a relationship between mtDNA haplogroups and fertility by comparing the fertile group with the RM/RIF group following the mitochondrial genome sequencing.

Determine if different mitochondrial haplogroups have an effect on embryo development based on the proportion of embryos that reached the blastocyst stage in each family and to investigate if this is related to aneuploidy in the embryos from couples in the fertile group.

Investigate the directly interacting nuclear genes with SNPs in identified mitochondrial haplogroups.

Investigate whether the genetic difference in ethnicity between partners based on the nuclear genome (expressed in terms of the proportion of different alleles to the total number of alleles compared) is associated with RM / RIF. The same analysis was performed separately for the Major Histocompatibility Complex (MHC) SNPs.

4.2 Materials and methods

Samples: Included 12 couples suffering from RM/RIF for unknown reasons as well as 11 couples attending the IVF clinic for Preimplantation genetic diagnosis (PGD) treatment with no history of infertility. Some embryos (n=56) which were WGA according to the method described previously (Chapter 2 section 2.3.4) and used in other sections of this thesis (Chapter 3 section 3.2.1.2, Chapter 5 section 5.2.3), were sequenced in order to confirm the they have the same mitochondrial haplotype as the mother.

The workflow of the experiments is illustrated in Figure 4-1.
Figure 4–1: An illustration of the workflow applied for sequencing the mitochondrial genome of parental genomic samples as well as some of their embryos.

The mitochondrial genome of the fertile group, RM/RIF couples as well as some WGA embryos was sequenced. Two methods were tested for sequencing. The first method involved the amplification by Long Range PCR using 5 overlapping primers while the other method was following the Illumina protocol using two primers.
4.2.1 Ethnicity determined by mtDNA haplogroup

To determine mitochondrial haplogroups, the mitochondrial genome of the fertile group, RM/RIF couples as well as embryos that had been WGA, were prepared for sequencing by two different methods. The first method involved the amplification by Long Range PCR using 5 overlapping primers while the other method was following the Illumina protocol using one set of primers. There is a high possibility of allele drop out (ADO) when sequencing WGA samples by long range PCR (LR-PCR) primers. Therefore, embryonic samples were sequenced using the first method as they were WGA and needed smaller fragments for sequencing. The mitochondrial genome sequencing was performed in three separate runs. Firstly, 5 samples were loaded to validate the protocol. Then two runs of 24 then 96 samples respectively were sequenced to confirm results were reproducible. The aneuploidy status of all the embryos from the couples undergoing PGD cycles for different single gene disorders was assessed using a-CGH as described previously in Chapter 2 section 2.4.1.7.

4.2.1.1 The whole mitochondrial genome sequencing

Two methods were tested for sequencing the mitochondrial genome. In the first method, the mitochondrial genome was sequenced by LR-PCR according to the following primers cited in Dames et al., 2013. The primers sequences are shown in Table 4-1.
### Table 4-1: Long Range PCR primers for mtDNA amplification

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Length (bp)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mito1-F</td>
<td>5'-ACATAGCACATTACAGTCAAATCCCTTCTCGTCCC-3'</td>
<td>3968</td>
</tr>
<tr>
<td>Mito1-R</td>
<td>5'-TGAGATTGTGTTTGGCTACTGCTCGAGTGC-3'</td>
<td></td>
</tr>
<tr>
<td>Mito2-F</td>
<td>5'-TACTCAATCCTCTGTGAGCTGACGATCAAACCTC-3'</td>
<td>5513</td>
</tr>
<tr>
<td>Mito2-R</td>
<td>5'-GCTTGGATTAAGGCGACAGCGATTTCTAGGATAGT-3'</td>
<td></td>
</tr>
<tr>
<td>Mito3-F</td>
<td>5'-TCATTTTTATTGCCAAACTAACCTCGCTGGACTC-3'</td>
<td>7814</td>
</tr>
<tr>
<td>Mito3-R</td>
<td>5'-CGTGATGTCATTATTTAAGGGGAAGGTGTGGGTAT-3'</td>
<td></td>
</tr>
<tr>
<td>hmt1-F1</td>
<td>5'-AACCAAAACCCAAAGACACC-3'</td>
<td>9289</td>
</tr>
<tr>
<td>hmt1-R1</td>
<td>5'-GCCAATAATGACGAGTGC-3'</td>
<td></td>
</tr>
<tr>
<td>hmt2-F1</td>
<td>5'-TCCCACTCCTAAACACATCC-3'</td>
<td>7626</td>
</tr>
<tr>
<td>hmt2-R2</td>
<td>5'-TTTATGGGATGTGAGCC-3'</td>
<td></td>
</tr>
</tbody>
</table>

Each primer has a 5' amino C6 modification. *Length represents the total length in base pairs of the amplicon derived from the primer pairs (including primers). F, forward; R, reverse. Adapted from *The Journal of Molecular Diagnostics*, 15 (4), Dames et al., *The Development of Next-Generation Sequencing Assays for the Mitochondrial Genome and 108 Nuclear Genes Associated with Mitochondrial Disorders*, pages 526-534, Copyright (2013), with permission from Elsevier.

Primers were designed to overlap in order to avoid allele drop out (ADO) as shown in Figure 4–2.

![Figure 4–2: The five overlapping amplicons used for LR-PCR amplification of the mtDNA genome.](image)

The five primers were tested to sequence genomic DNA as well as some embryonic samples. According to Dames et al. (2013), the use of 5 primer pairs reduces the chance of ADO and may improve the detection of deletions since smaller amplicons were used. Primers were diluted to 10nM. Gradient annealing for all the primers was performed to check the annealing temperature and that they worked efficiently. The PCR master mix was prepared according to the following: 12.2µL of H$_2$O, 2µL of 10x buffer, 0.4µL of DMSO, 2µL of 10x enhancer, 0.36µL of enzyme, 1µL of forward primer (10nM), 1µL of reverse primer (10nM) and 1µL of DNA to a final volume of 20µL. PCR conditions were as follows: 95°C for 2 minutes, 95°C for 15 seconds, 68°C for 10 minutes, 68°C for 20 minutes (repeated for 30 cycles from step 2), and held at 4°C on completion of the cycling reactions.

The second sequencing method followed the Human mtDNA Genome Guide sequencing protocol (Illumina, Part # 15037958, San Diego, USA). Two sets of primers were used. The master mix included: 22.5µL of dH$_2$O, 10xLA buffer 5µL, dNTPs 2µL, 2µL of each forward and reverse primer (10nM), 0.5µL of enzyme and 10µL of DNA (100pg) to a final volume of 50µL. The amplification efficiency and product size were checked by gel electrophoresis and the Bioanalyzer.

Qubit and Bioanalyzer were used to measure the concentration of amplicons in order to proceed with pooling the amplicons of each sample. Amplified amplicons of each sample were diluted to 0.2ng/µL. Samples were normalised using Resuspension buffer (RSB). Falcon tubes were used for normalisation and drops of samples were added to the wall of each tube. Samples were then all vortexed, centrifuged and put in 1.5mL tubes. The formula used to dilute the samples to 0.2ng/µL= (0.2/ concentration) x100. 100µL of each sample pool was transferred to
a 96 well PCR plate (Bio-Rad, Part # HSP-9601). The plate was sealed and covered with micro-seal adhesive seal (Bio-Rad, Part # MSB-1001).

Fragmentation and labelling of the amplified DNA samples by both PCR methods was performed using Nextera XT. The plate was left on the bench to thaw. 10µL of Tagment DNA (TD) buffer was added to each well by using the multichannel pippete in a new plate. 5 µL of sample was added. Then 5µL of Amplicon Tagment Mix (ATM) was added. All steps were followed exactly as directed in the Illumina protocol. The Illumina kit (24 indices) for 96 samples was used. This was enough for 24 samples per run. Therefore, a clean-up step was done using magnetic beads following the Illumina protocol. Samples were then left in 4°C before analysis.

4.2.1.2 Assessment of mtDNA amplicons quantity and quality

For normalisation, the Bioanalyzer and the Qubit DNA HS kit were used. Eleven samples were loaded on each Bioanalyzer chip. Therefore, three chips were used for the first 24 samples. Samples were then normalised by RSB. The reading provided by the Bioanalyzer (pmol/L) was converted to nm/L.

Qubit was used to measure the concentration of each amplified sample prior to pooling the amplicons. Using a plot of qubit concentrations against the molarity for a subset of measured samples by the Bioanalyzer the molarity of the remaining samples was determined. The molarity of samples which did not fall on the line in the scatter plot was calculated using the following formula: conc.(nM) = 10⁶ x Qubit (ng/µL) / 649 (average molecular weight of double stranded DNA pair) x size of library (16.6 KB). Following that, samples were normalised to 2pmol/µL and pooled then stored at -80°C.
4.2.1.3  Loading on the MiSeq of the 24 samples

Samples were loaded in two separate MiSeq runs. After QC checks (Bioanalyzer and Qubit) the qubit reading, and normalisation was performed based on the Bioanalyzer results. All samples showed a range of size between 150 and 200bp indicating that the fragmentation was efficient. Therefore, the concentration was normalised based on the molarity required 5µL of each sample in the tube resulting in a 0.2ng/µL. Prior to loading, the HT1+PhiX +MiSeq reagent box (1) were left to thaw on ice for one hour. PhiX was then diluted to 2nM by 10mM Tris-HCL (pH 8.5) with 0.1% Tween 20. Following that, 25µL of DNA pool library was mixed with 1.25 PhiX. For DNA denaturation, 5 µL of the DNA pool was mixed with 5 µL of freshly diluted NaOH (0.2N). This mixture was vortexed and centrifuged then incubated for 5 minutes on ice. 990µL of HT1 buffer was then added to mixture to create a final library of 6 pmoL. The sample sheet was prepared, and barcodes were entered. The flow cell was washed with ethanol and water and then placed in the MiSeq. Finally, 600µL of library mix was loaded.

4.2.1.4  Analysis

Sequencing data from Illumina instruments, including MiSeq, are streamed in real-time over the Internet to BaseSpace. BaseSpace is a powerful website where data can be stored and shared. The data is automatically converted to standard file formats for analysis. From BaseSpace (Binary Alignment/Map) BAM files, FASTQ files (text files of nucleotide sequence which can be loaded into alignment softwares for alignment) and (Variant Call Format) VCF files were produced. Raw reads were mapped back to the rCRS (NC_012920.1) using the mtDNA MiSeq Reporter Plug-In. Analysis was performed with a compatible version of the MiSeq Reporter software.
and the mtDNA Variant Analyser application. Variant sites were filtered for Base Quality 30 where the estimated error rate is 99.9%.

For haplotype analysis, variant calls were first evaluated, and all polymorphisms were checked. Nomenclature was performed as indicated by the International Union of Pure and Applied Chemistry (IUPAC) (Budowle, 1999) whether it is heteroplasmy or indels. Following that, the haplotype in an rCRS-coded format presented in rows was performed and put in an excel sheet (example shown in Appendix C 7.3.1). Haplotypes were then established by Haplogrep and EMPOP4 (Version 2.1) softwares.

The mitochondrial DNA amplification coverage was calculated according to the following equation: Total coverage (not the mapped coverage) = Total Yield * Sample% = Sample Yield, Sample yield / mitochondrial genome size = coverage, 5.6x10^9 * 0.049 = 274,400,000 (per sample yield for Sample 21 [4.9%]), 274,400,000 / 16569 = 16,561.

4.2.2 Ethnicity determined by nuclear genes

All genomic samples from the fertile and RM/RIF group of patients were genotyped using the Infinium® CoreExome-24 v1.2 SNP chip (Illumina, San Diego, USA). The HumanKaryomap-12 BeadChip (Illumina, San Diego, USA) was tested to infer the ethnicities of some couples (6 individuals) since some ancestry informative markers (AIM) were present on the chip. Since the data for aneuploidy and Karyomapping was available for some embryos, a comparison was performed between both techniques.
4.2.2.1 SNP genotyping using the Infinium® CoreExome-24 v1.2 SNP chip

the Infinium CoreExome-24 v1.1 BeadChip includes all the tag SNPs found on the Infinium Core-24 BeadChip, plus over 240,000 markers from the Infinium HumanExome BeadChip. The Chip has a total of 551,839 markers and accommodates 24 samples. The chip was chosen because it allows cost effective large-scale genotyping studies and can be used to obtain baseline sample data sets for various applications including: ancestry, common variant, mtDNA, loss of-variant, insertion/deletion (indel) detection studies and sex confirmation (Illumina, Inc).

Genomic samples were prepared and diluted to bring to 70-100 ng/µL. Following that, samples were processed at the Institute of Child’s Health (ICH) at University College London (UCL) by Dr Mark Kristiansen. All the steps were carried out using Illumina’s (Infinium CoreExom-24 v1.1 BeadChip) and standard protocols. Scanning was done using the (iScan) machine by Illumina and the Illumina's GenomeStudio v2011.1 software was used to call the genotypes. Genotyping was performed as per Illumina’s Infinium HTS Assay (15045737_A, October 2013) protocol (Illumina Inc, San Diego, USA).

4.2.2.2 Karyomapping

Karyomapping utilizes SNP array technology with the HumanKaryomap-12 BeadChip to obtain genome-wide genotypes for parents, a reference sample (genomic DNA extracted from blood) and embryos (biopsy of single/few cells). The outcome is a set of genotype-calls for each SNP on the array (almost 300,000 SNPs). In this study, the karyomap SNP chip was used to establish genotypes of three families including; parents and two embryos from each family, without the use of a sibling or other family member (one embryo in each case assigned as a
‘reference’). The selection of different sets of SNPs located on this chip enabled the assignment of parents to different population groups. The analysis of the genotypes and signal intensity of embryos was also used to compare their aneuploidy status with a-CGH (Appendix B section 7.2.3).

All pre-Karyomapping procedures were performed at UCL Centre for PGD and families were anonymised. The Karyomapping steps were performed by Kerra Pearce at the Institute of Child’s Health. Genomic DNA samples from parents and WGA embryonic samples were processed using Infinium HumanKaryomap-12 DNA Analysis Kit (Illumina, San Diego, USA).

4.2.2.3 Analysis

4.2.2.3.1 Classification of partners to population groups

The full list of total 550,601 markers & 294,602 SNPs with their locations on both chips was obtained from Illumina Inc. as Microsoft Office Excel file. From this, SNPs sets were generated. BlueFuse genotype calls are presented as AA, AB, BB or NC (no call) and therefore, GenomeStudio v1.0 (Illumina, San Diego, USA) was used to provide bases (A, T, G, C) in each locus. Raw data was exported from both BlueFuse and GenomeStudio as Excel files, then used for further analysis. The number of SNPs which gave relevant calls was counted in each region of interest. It was performed by excluding the no call (NC).

Ancestry informative markers (AIM) were selected to determine the geographical origins of ancestors of each partner as the use of these markers has been previously reported (Halder et al., 2008, Lao et al., 2010, Phillips et al., 2013, Kidd et al., 2014, Phillips et al., 2014). Markers that were available on the Infinium® CoreExome-24 v1.2 BeadChip were assessed for all individuals and the HumanKaryomap-12 v1.0
BeadChip (n=62) was used to confirm the ethnicities of six individuals who underwent PGD analysis by Karyomapping. The lists of SNPs used are available in Appendix C 7.3.2. These studies were based on analysis of the genotypes of multiple individuals from each population obtained from two datasets - 1000 Genomes Project and HGDP-CEPH Genome Diversity Panel.

Three sets of AIMs on the Infinium® CoreExome-24 v1.2 BeadChip were used. A set of 247 AIMs was used to classify samples as originated from seven population groups (Eurasians were divided into European, South Asian and Middle Eastern - the latter reference data comprising an Algerian and three non-Jewish Israeli populations). Another set was based on (n=229) 1000 most divergent AIMs and the latter set had several AIMs added that better differentiate Europeans from South Asians. However, there were insufficient markers in the exome chip to distinguish the Middle Eastern population, so certainly many Middle East vs. European/South Asian comparisons plus a South Asian vs. Middle East comparison were not enough to make a reliable distinction from the genetic data.

The set of 62 AIMs on the Karyomapping chip was used to classify samples as originating from Africa, Europe or East Asia, while another subset of 33 SNPs added another population – South Asian. European and South Asian are not extremely divergent populations, so basing the analysis on the smaller subset only could lead to classification errors. Therefore, both AIMs sets were used.

Samples were assigned to population groups with online Bayesian classifier - Snipper 2.0 (http://mathgene.usc.es/snipper/). This tool detects the underlying genetic population among a set of individuals genotyped at multiple SNPs and computes the proportion of the genome of an individual originating from each inferred population using quantitative clustering method (McNevin et al., 2013). To
determine any relationship between the difference in ethnicity and the phenotype the logistic regression test was performed.

4.2.3 Identity by State (IBS)

The identity by State (IBS) method was used for the genomic comparison that analyses variant pairs between two individuals. The results were scored as:

- No shared alleles when the genotypes were (AA or BB).
- One shared allele when the two individuals share one call only [(AB and BB) or (AA and AB)].
- Identical: both share a similar genotype call [(AA and AA) or (AB and AB) or (BB and BB)].

The real value of IBS lies not in the comparison of individual SNPs but rather comparing high density SNP data across the entire genome. By calculating the ratio of the shared alleles to the total number of alleles called for the samples, the complement of this value to 1 will be genetic difference (1-IBS). Plink (v1.9, http://pngu.mgh.harvard.edu/purcell/plink/) was used to calculate the distance, then the triangular data matrix (Figure 4-3) was analysed using a script compiled by Dr. Yoan Dieckman from the “MACE Lab” at UCL using “Python” (Appendix C section 7.3.3). After that, “R” software was used to help with the analysis and plotting of the results using a two-sample “Kolmogorov-Smirnov” test.
4.2.3.1 HLA

SNPs in nuclear genes encoding proteins important in mitochondrial function present on the human core exome SNP chip (n=959) and SNPs in the MHC were extracted (n=5632). The IBS differences between partners in each couple within the two groups were compared.

4.2.4 Aneuploidy comparison

The aneuploidy status of embryos analysed by BlueFuse Multi v4.0 (BlueGnome Ltd, Cambridge, UK) Karyomapping module was assessed and compared to the aneuploidy data previously assessed by a-CGH (GenetiSure, Agilent) (Chapter 2 section 2.4.1.7). This was possible since both techniques require the whole genome amplification of embryonic samples by MDA. Embryos from each family were assigned as a ‘reference’ one after another to check if embryo quality affects SNP calls. SNP call rates were checked against thresholds established by manufacturer. As the sex of embryos was unknown, male sex was initially selected in BlueFuse for all embryos.
4.3 Results

4.3.1 Ethnicity determined by the mitochondrial genome

4.3.1.1 Assessment of mtDNA amplicons quantity and quality

Following PCR amplification, PCR products were loaded on the gel and confirmed the fragment sizes (Figure 4–4).

![Gel electrophoresis](image)

*Figure 4–4: Gel electrophoresis (0.8%) showing the size of the PCR products following mtDNA sequencing.*

The PCR reaction worked for the genomic samples (labelled 1,2,3) for both primer sets as shown in the figure. Two negative controls (one for the single cell room and the second for the main laboratory confirmed no contamination. The gel was run for an hour at 25 volts. The 1000bp ladder was used for comparison and PCR product size was observed.

The quantity and quality of the generated mitochondrial amplicons from embryonic & genomic samples was assessed by the Bioanalyzer. To amplify the mitochondrial genome, five sets of primers were used using the first method. Only three primer sets (Mito1,2,3) generated products (Figure 4–5).
A) Primer pairs LR-MITO1 generated the expected size range 4305 bp which appeared as a single peak. B) Primer pairs LR-MITO2 generated the expected size range 5809 bp which appeared as a single peak. C) Primer pairs LR-MITO3 generated the expected size range 8136 bp which appeared as a single peak. Even though there are some artefacts.

Assessment of the mitochondrial genome amplification using the one set of primers in the second method, has only worked on genomic samples. The failure of the one primer set to give products was probably due to the use of WGA embryonic samples as when these primers were used with genomic DNA, they gave the expected products (Figure 4–6).

Figure 4–5: Electropherogram of the LR-MITO three sets of PCR Amplicon on an embryonic sample.
A) Primer pairs LR-MITO1 generated the expected size range 4305 bp which appeared as a single peak. B) Primer pairs LR-MITO2 generated the expected size range 5809 bp which appeared as a single peak. C) Primer pairs LR-MITO3 generated the expected size range 8136 bp which appeared as a single peak. Even though there are some artefacts.
Figure 4–6: Example of Electropherogram of the MTL1-MTL2 two sets of PCR Amplicon on genomic and an embryonic sample.
Primer pair MTL1 should generate a product size of 9065 bp and therefore the expected size range is between 7705-10425 bp. This appeared as a single peak 8921bp for the genomic sample. B) The MTL2 pair of primers should generate a product size of 11170 bp which is expected to appear as a single peak in the range of 9495-12845. This appeared as a single peak 9906 bp for the genomic sample. C) the same primer pairs did not work with embryonic samples.

4.3.1.2 Mitochondrial genome sequencing
Both runs worked well (Table 4-2). An average of 95% of 151bp end reads aligned to the mitochondrial genome. There was coverage of each mtDNA base with mapping quality score >30 and average coverage depth of 3710x and 9500x per samples in each run (Figure 4–7).

Table 4-2: Results of all runs used for mitochondrial genome sequencing.

<table>
<thead>
<tr>
<th>Run</th>
<th>Number of samples</th>
<th>Total number of reads</th>
<th>Average depth of read per sample</th>
<th>%&gt;=Q30*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>96</td>
<td>18,060,000</td>
<td>3424x (56.6 MB)</td>
<td>90.22%</td>
</tr>
<tr>
<td>2</td>
<td>24</td>
<td>12,090,074</td>
<td>9151x (150 MB)</td>
<td>92.9%</td>
</tr>
</tbody>
</table>

*Q30: is an indication of % of bases called at a quality of 30Q or above. The average depth of read per sample was calculated by the following formula X= (Total number of reads (million) * read length (bp) *2) / sample)/target region (0.016 Mb).
The total number of reads was 12,090,074. 93% of bases were classified at a quality of 30Q or above. Average depth of read was 9151x per sample (150 MB per sample).

4.3.1.2.1 Relationship between mitochondrial haplogroups and fertility

4.3.1.2.1.1 Comparison between the fertile (PGD) and infertile (RM/RIF) groups

Several mitochondrial haplogroups were identified reflecting the diverse ethnic population in London (Table 4-3, Figure 4–8). All the detected haplogroups descended from the ancestral sub-haplogroups N, R and M. In more detail, of the 46 individuals 39 were of European origin (K, J, T, H, U and T) and 7 were of Asian origin (F, M, N). The average female age in the two groups was; 33 in the fertile group, 35.75 in the infertile group. Overall more couples from the RM/RIF group (5/12) had identical mtDNA haplotypes compared to the fertile group (1/11).

Statistical analysis shows that repeated miscarriage and repeated implantation failure are significantly correlated with a smaller distance difference in ethnicity (p=0.04) determined by their mitochondrial genomes i.e.; not hybrid vigour. Based on the mitochondrial genetic distance between the two haplogroups in individuals from each couple, couples within the fertile group were more likely to be distantly related compared to the infertile group. According to our initial hypothesis, a similar

Figure 4–7: The QScore distribution for the second mitochondrial genome sequencing run.
The total number of reads was 12,090,074. 93% of bases were classified at a quality of 30Q or above. Average depth of read was 9151x per sample (150 MB per sample)
mitochondrial haplotype and nuclear background between parents would prevent mitochondrial nuclear mismatch. Despite our small number of samples, we were surprised to find that differences in mitochondrial haplotype between parents showed a trend towards beneficial fertility outcomes.
Table 4-3: comparison of mtDNA haplotypes between the fertile and the RM/RIF group.

<table>
<thead>
<tr>
<th>Family ID</th>
<th>mtDNA haplogroup</th>
<th>Maternal age</th>
<th>Referral reason</th>
<th>haplotype origin</th>
<th>Haplotype distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family-1</td>
<td>T2a1b1a</td>
<td>T M</td>
<td>PGD</td>
<td>Europe Asia</td>
<td>3</td>
</tr>
<tr>
<td>Family-2</td>
<td>H1e4a</td>
<td>H N</td>
<td>PGD</td>
<td>Europe Asia</td>
<td>3</td>
</tr>
<tr>
<td>Family-3</td>
<td>H31a</td>
<td>H T</td>
<td>PGD</td>
<td>Europe Eastern Europe</td>
<td>2</td>
</tr>
<tr>
<td>Family-4</td>
<td>U5b3b1</td>
<td>U T</td>
<td>PGD</td>
<td>Eastern Europe</td>
<td>2</td>
</tr>
<tr>
<td>Family-5</td>
<td>H26a1 U5a1a1</td>
<td>H U</td>
<td>PGD</td>
<td>Europe Eastern Europe</td>
<td>2</td>
</tr>
<tr>
<td>Family-6</td>
<td>H4a1a1a</td>
<td>H T</td>
<td>PGD</td>
<td>Europe Eastern Europe</td>
<td>2</td>
</tr>
<tr>
<td>Family-7</td>
<td>H4a1a3</td>
<td>H H</td>
<td>PGD</td>
<td>Europe Europe</td>
<td>0</td>
</tr>
<tr>
<td>Family-8</td>
<td>F3a1 H1+16189</td>
<td>F H</td>
<td>PGD</td>
<td>Asia Europe</td>
<td>3</td>
</tr>
<tr>
<td>Family-9</td>
<td>H26a1 T2b30</td>
<td>H T</td>
<td>PGD</td>
<td>Europe Eastern Europe</td>
<td>2</td>
</tr>
<tr>
<td>Family-10</td>
<td>H U5b2a2b</td>
<td>H U</td>
<td>PGD</td>
<td>Europe Eastern Europe</td>
<td>2</td>
</tr>
<tr>
<td>Family-11</td>
<td>U1a1a1 M22a</td>
<td>U M</td>
<td>PGD</td>
<td>Eastern Europe</td>
<td>3</td>
</tr>
<tr>
<td>Family-12</td>
<td>V U6a5b</td>
<td>V U</td>
<td>RIF</td>
<td>Europe Eastern Europe</td>
<td>2</td>
</tr>
<tr>
<td>Family-13</td>
<td>H26C H3ap</td>
<td>H H</td>
<td>RIF/ RM</td>
<td>Europe Europe</td>
<td>0</td>
</tr>
<tr>
<td>Family-14</td>
<td>M30+16234 M3c1a</td>
<td>M M</td>
<td>RIF</td>
<td>Asia Asia</td>
<td>0</td>
</tr>
<tr>
<td>Family-15</td>
<td>T2b4+152 H4a14</td>
<td>T H</td>
<td>RIF</td>
<td>Eastern Europe Europe</td>
<td>2</td>
</tr>
<tr>
<td>Family-16</td>
<td>T2b21 U4b1b1</td>
<td>T U</td>
<td>RM</td>
<td>Eastern Europe Eastern Europe</td>
<td>2</td>
</tr>
<tr>
<td>Family-17</td>
<td>H3+16311 H11a7</td>
<td>H H</td>
<td>RIF</td>
<td>Europe Europe</td>
<td>0</td>
</tr>
<tr>
<td>Family-18</td>
<td>U2C1 U7a3a</td>
<td>U U</td>
<td>RIF</td>
<td>Eastern Europe Europe</td>
<td>0</td>
</tr>
<tr>
<td>Family-19</td>
<td>M10a1a1b T2b4</td>
<td>M T</td>
<td>RIF</td>
<td>Asia Eastern Europe</td>
<td>3</td>
</tr>
<tr>
<td>Family-20</td>
<td>J1c1 H41a1a1a1</td>
<td>J H</td>
<td>RM</td>
<td>Eastern Europe Europe</td>
<td>2</td>
</tr>
<tr>
<td>Family-21</td>
<td>J1c2 U5b2a6</td>
<td>J U</td>
<td>RM</td>
<td>Europe Eastern Europe</td>
<td>2</td>
</tr>
<tr>
<td>Family-22</td>
<td>K1a4c R0a2j</td>
<td>K R</td>
<td>RIF</td>
<td>Native Europe South Asia</td>
<td>2</td>
</tr>
<tr>
<td>Family-23</td>
<td>H6a1b2 H1c1a</td>
<td>H H</td>
<td>RIF</td>
<td>Europe Europe</td>
<td>0</td>
</tr>
</tbody>
</table>

Couples referred for PGD have no history of infertility (fertile). Couples with history of infertility include RIF: repeated implantation failure, RM: repeated miscarriage.
The haplotype distance in the fertile group was significantly greater than the distance in the infertile group (p=0.04). Distance of 0: was given when they shared the same haplotype for example; both partners were H. 1: if the haplotypes were closely related – from same branch for example; one partner is J while the other is T. 2: if the haplotypes were from the same continent for example; one partner is V while the other is U. 3: if they were distantly related – from different continents for example; one partner is H while the other is M. Contrary to what we expected, Couples within the fertile group were more likely to be distantly related compared to the infertile group. The diagram on the right is adapted from MITOMAP: A Human Mitochondrial Genome Database. http://www.mitomap.org, 2017. 

Figure 4–8: Haplotype distance comparison between the fertile and RM/RIF group.
4.3.1.2.2 Correlation between mitochondrial haplogroups and blastocyst rate

4.3.1.2.2.1 Comparison within the fertile (PGD) group

Since the embryonic data was available for the fertile couples; the genetic distance in mtDNA haplogroup was tested for each couple. No correlation was found when the same test was applied comparing the mtDNA haplogroup distance based on the number of the embryos that arrested and the embryos that reached the blastocyst stage in each family. Table 4-4 shows the number of eggs that fertilised and reached the blastocyst stage in each family. The fertilisation percentage and blastocysts rate formed were compared based on the mitochondrial haplogroup.

Table 4-4: The maternal and paternal mitochondrial haplotypes with fertilisation and blastocyst formation for fertile couples

<table>
<thead>
<tr>
<th>Family ID</th>
<th>Eggs</th>
<th>Maternal age</th>
<th>mtDNA Haplogroup</th>
<th>Fertilisation %</th>
<th>Blastocyst %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9</td>
<td>42</td>
<td>T2a1b1a, M2a'b</td>
<td>3</td>
<td>88.9</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>36</td>
<td>H1e4a, N1b1b1</td>
<td>3</td>
<td>58.33</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>32</td>
<td>H31a, T1a1b</td>
<td>2</td>
<td>75</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>36</td>
<td>U5b3b1, T2b3+151</td>
<td>2</td>
<td>87.5</td>
</tr>
<tr>
<td>5</td>
<td>16</td>
<td>29</td>
<td>H26a1, U5a1a1</td>
<td>2</td>
<td>87.5</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>28</td>
<td>H4a1a1a, T2b7a1</td>
<td>2</td>
<td>70</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>37</td>
<td>H4a1a3, H1bb</td>
<td>0</td>
<td>57.143</td>
</tr>
<tr>
<td>8</td>
<td>23</td>
<td>33</td>
<td>F3a1, H1+16189</td>
<td>3</td>
<td>86.96</td>
</tr>
<tr>
<td>9</td>
<td>23</td>
<td>29</td>
<td>H26a1, T2b30</td>
<td>2</td>
<td>82.61</td>
</tr>
<tr>
<td>10</td>
<td>19</td>
<td>31</td>
<td>H, U5b2a2b</td>
<td>2</td>
<td>78.95</td>
</tr>
<tr>
<td>11</td>
<td>10</td>
<td>30</td>
<td>U1a1a1, M22a</td>
<td>3</td>
<td>70</td>
</tr>
</tbody>
</table>

Fertilisation % refers to the percentage of eggs fertilised by ICSI. Blastocyst % refers to the percentage of fertilised embryos that developed to the blastocyst stage. Haplotype distance refers to the mitochondrial haplogroup genetic distance between partners in each couple as explained in Figure 4-8.
The correlation between the blastocyst rate and the mitochondrial haplotype was reanalysed by multiple linear regression controlling for other parameters including maternal age, fertilisation rate and mitochondrial haplogroups F, T, M, N and U in relation to haplotype H (which was used as a reference since it was the most common haplotype in our group of samples. Haplogroup T showed a negative association with blastocyst rate \( p=0.03, B=-0.456 \). For most of the couples, the male partner had the T haplogroup. Also, based on the results shown (Table 4-5) the percentage of blastocysts rate was affected by the fertilisation rate \( p=0.043, B=0.018 \) as shown in Figure 4–9.

### Table 4-5: Correlation between blastocyst rate and mtDNA haplotype

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>T</th>
<th>Significance (P value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (constant)</td>
<td>-0.356</td>
<td>0.408</td>
<td>-0.873</td>
<td>0.447</td>
</tr>
<tr>
<td>Maternal age</td>
<td>-0.008</td>
<td>0.009</td>
<td>-0.152</td>
<td>Reference</td>
</tr>
<tr>
<td>Haplogroup H</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haplogroup F</td>
<td>-0.154</td>
<td>0.190</td>
<td>-0.218</td>
<td>-0.809 0.478</td>
</tr>
<tr>
<td>Haplogroup T</td>
<td>-0.456</td>
<td>0.118</td>
<td>-1.120</td>
<td>-3.859 0.031</td>
</tr>
<tr>
<td>Haplogroup M</td>
<td>-0.74</td>
<td>0.95</td>
<td>-0.141</td>
<td>-0.783 0.491</td>
</tr>
<tr>
<td>Haplogroup N</td>
<td>-0.133</td>
<td>0.143</td>
<td>-0.188</td>
<td>-0.924 0.424</td>
</tr>
<tr>
<td>Haplogroup U</td>
<td>-0.185</td>
<td>0.119</td>
<td>-0.439</td>
<td>-0.1553 0.218</td>
</tr>
<tr>
<td>Fertilisation rate</td>
<td>0.018</td>
<td>0.005</td>
<td>-0.975</td>
<td>3.388 0.043</td>
</tr>
</tbody>
</table>

\( B=\)coefficient which shows if the parameter used positively or negatively affect the outcome variable.

![Figure 4-9: Mitochondrial haplogroup effect on fertilisation and blastocyst rate within the fertile group.](image)

Female and male haplotypes of each couple are displayed on the X axis respectively. Even though couples in the fertile group had very high fertilisation rate, not all families had embryos with good quality reaching the blastocyst stage. Haplogroup T negatively affects blastocyst rate \( p=0.002 \).
4.3.1.2.2.2 Analysis of mtDNA T haplogroup SNPs

Further interactions of the T haplotype SNPs in mitochondrial genes directly interacting with the nuclear genes encoding proteins in the ETC that affect mitochondrial function were investigated (Table 4-6).

Table 4-6: The SNPs making up the mitochondrial haplogroup T.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Gene</th>
<th>SNP</th>
<th>aa change</th>
<th>Interacting nuclear gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2*JT</td>
<td>MT-ND1</td>
<td>T4216C</td>
<td>I</td>
<td>non syn: Y-H</td>
</tr>
<tr>
<td>JT</td>
<td>MT-ND4</td>
<td>A11251G</td>
<td>I</td>
<td>sy L-L</td>
</tr>
<tr>
<td>T T</td>
<td>MT-RNR1</td>
<td>G709A</td>
<td>12S</td>
<td>rRNA</td>
</tr>
<tr>
<td>T T</td>
<td>MT-RNR2</td>
<td>G1888A</td>
<td>16S</td>
<td>rRNA</td>
</tr>
<tr>
<td>T T</td>
<td>MT-ND2</td>
<td>A4917G</td>
<td>I</td>
<td>non syn: N-D</td>
</tr>
<tr>
<td>T T</td>
<td>MT-ATP6</td>
<td>G8697A</td>
<td>V</td>
<td>syn-M-M</td>
</tr>
<tr>
<td>T T</td>
<td>MT-TR</td>
<td>T10463C</td>
<td>tRNA</td>
<td></td>
</tr>
<tr>
<td>T T</td>
<td>MT-ND5</td>
<td>G13368A</td>
<td>I</td>
<td>syn: G-G</td>
</tr>
<tr>
<td>T T</td>
<td>MT-CYB</td>
<td>G14905A</td>
<td>III</td>
<td>syn: M-M</td>
</tr>
<tr>
<td>T T</td>
<td>MT-CYB</td>
<td>A15607G</td>
<td>III</td>
<td>syn: K-K</td>
</tr>
<tr>
<td>T T</td>
<td>MT-TR</td>
<td>G15928A</td>
<td>tRNA</td>
<td>Trna</td>
</tr>
<tr>
<td>T T1</td>
<td>MT-HV1</td>
<td>C12633A</td>
<td>I</td>
<td>syn: S-S</td>
</tr>
<tr>
<td>T T1a</td>
<td>MT-HV1</td>
<td>C16186T</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T T1a1</td>
<td>MT-ND5</td>
<td>C16233A</td>
<td>I</td>
<td>syn: S-S</td>
</tr>
<tr>
<td>T T1a1b</td>
<td>MT-TR</td>
<td>T10463C</td>
<td>tRNA</td>
<td></td>
</tr>
<tr>
<td>T T2</td>
<td>MT-HV1</td>
<td>C16294T</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T T2a</td>
<td>MT-HV1</td>
<td>C16294T</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T T2a1</td>
<td>MT-ND5</td>
<td>T139565</td>
<td>I</td>
<td>Syn:L-L</td>
</tr>
<tr>
<td>T T2a1b</td>
<td>MT-TR</td>
<td>T10463C</td>
<td>tRNA</td>
<td></td>
</tr>
<tr>
<td>T T2b</td>
<td>MT-HV1</td>
<td>T10463C</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T T2b1</td>
<td>MT-RNR2</td>
<td>T2141C</td>
<td>16S</td>
<td>Rrna</td>
</tr>
<tr>
<td>T T2b1b</td>
<td>MT-ND5</td>
<td>A13966G</td>
<td>I</td>
<td>non syn:T-A</td>
</tr>
<tr>
<td>T T2b1b</td>
<td>MT-TR</td>
<td>T10463C</td>
<td>tRNA</td>
<td></td>
</tr>
<tr>
<td>T T2b2</td>
<td>MT-ND4L</td>
<td>A10750G</td>
<td>I</td>
<td>non syn: N-S</td>
</tr>
<tr>
<td>T T2b2</td>
<td>MT-HV2, MT-</td>
<td>C151T</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T T2b2</td>
<td>OHR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T T2b2</td>
<td>MT-RNR1</td>
<td>G930A</td>
<td>12S</td>
<td>Rrna</td>
</tr>
<tr>
<td>T T2b2</td>
<td>MT-TR</td>
<td>T10463C</td>
<td>tRNA</td>
<td></td>
</tr>
<tr>
<td>T T2b2</td>
<td>MT-HV1</td>
<td>T10463C</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T T2b3</td>
<td>MT-TR</td>
<td>T10463C</td>
<td>tRNA</td>
<td></td>
</tr>
<tr>
<td>T T2b3</td>
<td>MT-ND5</td>
<td>T13768C</td>
<td>I</td>
<td>non-syn: F-L</td>
</tr>
<tr>
<td>T T2b3</td>
<td>MT-HV1</td>
<td>C16257d</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
As shown in Table 4-6, most of the SNPs that make up the mitochondrial T haplogroup are in complex I of the mitochondrial electron transport chain. Sequencing of the corresponding selected nuclear genes was performed in Chapter 5 before the mitochondrial nuclear mismatch investigation could be determined.

4.3.1.2.3 The effect of mtDNA haplogroup on human preimplantation embryo development and aneuploidy

The results of the chromosomal analysis were associated with the corresponding haplogroups. Based on the maternal mtDNA haplogroup, embryonic haplogroups were inferred for the un-sequenced embryos since mtDNA is maternally inherited (Appendix B section 7.2.3). Hence embryonic mitochondrial haplotypes were analysed and correlated with embryo morphology and aneuploidy status.

No association was found between embryonic mtDNA haplogroup and embryo morphology and/ or aneuploidy when multiple linear regression analysis was performed. The haplotype of the embryo does not have a significant influence on embryo morphology nor aneuploidy status (p value varied from 0.549-0.592). However, it was confirmed that aneuploidy status had a negative influence on embryo morphology (p=0.034, B= -0.787) by logistic regression.

4.3.1.2.4 The effect of embryonic mitochondrial haplogroup on embryonic mitochondrial depth of read

Applying NGS on some embryonic samples (n=23) was successful and differences in depth of read were observed according to cell number, developmental stage and embryo quality, an example is shown in (Figure 4–10). Based on these results, there was a trend showing more mtDNA in the poorly developing embryos compared to the embryos that developed to the blastocyst stage.
Figure 4–10: The difference in depth of read from 5 different embryonic samples as shown on IGV.

Difference in depth of read, inferred by integrative genomic viewer (IGV), between the genomic control sample and embryonic samples is clearly observed. Samples 2 & 3 are whole embryos from the same family which are carriers for the mitochondrial haplogroup F. Embryonic Sample 2 reached the blastocyst stage (depth of read: 263,979.0) while embryonic sample 3 was degenerated showing a higher depth of read (294,133.0).

Even though the embryonic mitochondrial genome sequencing confirmed the inheritance from the mother and showed differences in coverage compared to the genomic samples, there may be other factors that affect depth of read results from embryonic samples such as WGA, ADO and difference in the preparation process.

4.3.2 The effect of mtDNA haplogroup T and embryo mtDNA number

The mitochondrial DNA number of those embryos was already quantified by relative quantification in Chapter 3 section 3.3.2. The effect of mtDNA haplotype T on mitochondrial number was tested by the Mann Whitney U test. Embryos from individuals carrying mtDNA haplogroup T were compared with individuals carrying
the other haplogroups. However, no association was found between haplotype T and embryo mitochondrial number.

4.3.3 Ethnicity determined by nuclear genes

Since the mitochondrial haplogroup provided information about the maternal ethnicity of each individual, overall ethnicity of the couples was determined by the analysis of their nuclear genes. Ethnicity based on nuclear genes was determined using AIM which were available on the Infinium® CoreExome-24 v1.2 SNP BeadChip (n=247) for all the individuals and the HumanKaryomap-12 v1.0 BeadChip (n=62) was used to confirm the analysis for six individuals who underwent PGD analysis by karyomapping. Ethnicity of those six individuals analysed by both SNP chips confirmed the analysis. However, the final analysis was based on the Infinium® CoreExome-24 v1.2 SNP chip since it covers the whole genome. Ancestry estimates of all analysed individuals is illustrated in Figure 4-11. Individuals were mainly from a European, central south and east Asian origin. A divergent ethnicity was observed in two couples within the RM/RIF group. In family-19, the female (19-F) was clustered in East Asia while the male (19-M) was clustered in Europe. In family-12, the male partner (12-M) was clustered in Africa while his female partner (12-F) was clustered in is Central South Asia, showing different ethnicities for those two families.

Ethnicities determined by nuclear genes compared to the mitochondrial haplogroups was performed and listed in Table 4-7.
Figure 4–11: Principle component analysis (PCA) clustering for cases (RM/RIF) and controls (PGD) based on 7 populations using 229 SNPs.

As shown in the figure, individuals were mainly from a European, Central South and East Asian origin. In family-19, the female (19-F) was clustered in East Asia while the male (19-M) was clustered in Europe. In family-12, the male partner (12-M) was clustered in Africa while his female partner (12-F) was clustered in is Central South Asia, which suggests a divergent ethnicity for those two families.
Table 4-7: Ethnicities determined by nuclear genes compared to the mitochondrial haplogroups.

<table>
<thead>
<tr>
<th>Family ID</th>
<th>Maternal age</th>
<th>Ethnicity by nDNA</th>
<th>Ethnicity by mtDNA</th>
<th>Family ID</th>
<th>Maternal age</th>
<th>Ethnicity by nDNA</th>
<th>Ethnicity by mtDNA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Hap</td>
<td></td>
<td></td>
<td></td>
<td>Hap</td>
</tr>
<tr>
<td>Family-1-Female ♦</td>
<td>42</td>
<td>European</td>
<td>T</td>
<td>Family-12-Female*</td>
<td>35</td>
<td>Middle east</td>
<td>V</td>
</tr>
<tr>
<td>Family-1-Male ♦</td>
<td></td>
<td>Central South Asian</td>
<td>M</td>
<td>Family-12-Male*</td>
<td></td>
<td>African</td>
<td>U</td>
</tr>
<tr>
<td>Family-2-Female ♦</td>
<td>36</td>
<td>European or Middle Eastern</td>
<td>H</td>
<td>Family-13-Female</td>
<td>35</td>
<td>European or Middle east</td>
<td>H</td>
</tr>
<tr>
<td>Family-2-Male ♦</td>
<td></td>
<td>European or Middle Eastern</td>
<td>N</td>
<td>Family-13-Male</td>
<td></td>
<td>European</td>
<td>H</td>
</tr>
<tr>
<td>Family-3-Female ♦</td>
<td>32</td>
<td>European</td>
<td>H</td>
<td>Family-14-Female</td>
<td></td>
<td>Central Asian</td>
<td>M</td>
</tr>
<tr>
<td>Family-3-Male</td>
<td></td>
<td>European or Middle Eastern</td>
<td>T</td>
<td>Family-14-Male</td>
<td>37</td>
<td>Central Asian</td>
<td>M</td>
</tr>
<tr>
<td>Family-4-Female</td>
<td>36</td>
<td>European</td>
<td>U</td>
<td>Family-15-Female</td>
<td>35</td>
<td>European</td>
<td>T</td>
</tr>
<tr>
<td>Family-4-Male</td>
<td></td>
<td>European</td>
<td>T</td>
<td>Family-15-Male</td>
<td></td>
<td>European</td>
<td>H</td>
</tr>
<tr>
<td>Family-5-Female</td>
<td>29</td>
<td>European</td>
<td>H</td>
<td>Family-16-Female</td>
<td>33</td>
<td>European or Middle east</td>
<td>T</td>
</tr>
<tr>
<td>Family-5-Male</td>
<td></td>
<td>European</td>
<td>U</td>
<td>Family-16-Male</td>
<td></td>
<td>European</td>
<td>U</td>
</tr>
<tr>
<td>Family-6-Female</td>
<td>28</td>
<td>European</td>
<td>H</td>
<td>Family-17-Female</td>
<td>36</td>
<td>European</td>
<td>H</td>
</tr>
<tr>
<td>Family-6-Male</td>
<td></td>
<td>European or Middle Eastern</td>
<td>T</td>
<td>Family-17-Male</td>
<td></td>
<td>European</td>
<td>H</td>
</tr>
<tr>
<td>Family-7-Female</td>
<td>37</td>
<td>European</td>
<td>H</td>
<td>Family-18-Female</td>
<td>37</td>
<td>Central Asian</td>
<td>S</td>
</tr>
<tr>
<td>Family-7-Male</td>
<td></td>
<td>European</td>
<td>H</td>
<td>Family-18-Male</td>
<td></td>
<td>Central Asian</td>
<td>S</td>
</tr>
</tbody>
</table>

HAP: refers to haplotype. Mismatched couples are coloured in red. *Couples with different ethnicities determined by nDNA. ♦Couples distantly related based on mitochondrial haplotypes.
<table>
<thead>
<tr>
<th>Family ID</th>
<th>Maternal age</th>
<th>Ethnicity by nDNA</th>
<th>Ethnicity by mtDNA</th>
<th>Family ID</th>
<th>Maternal age</th>
<th>Ethnicity by nDNA</th>
<th>Ethnicity by mtDNA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hap</td>
<td>origin</td>
<td></td>
<td>Hap</td>
<td>Origin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family-8-Female ♦</td>
<td>33</td>
<td>Central South Asian</td>
<td>F</td>
<td>Asia</td>
<td>Family-19-Female*♦</td>
<td>33</td>
<td>East Asian</td>
</tr>
<tr>
<td>Family-8-Male ♦</td>
<td></td>
<td>European</td>
<td>H</td>
<td>Europe</td>
<td>Family-19-Male*♦</td>
<td></td>
<td>European or Middle Eastern</td>
</tr>
<tr>
<td>Family-9-Female</td>
<td>29</td>
<td>European</td>
<td>H</td>
<td>Europe</td>
<td>Family-20-Female</td>
<td>37</td>
<td>European</td>
</tr>
<tr>
<td>Family-9-Male</td>
<td></td>
<td>European</td>
<td>T</td>
<td>Eastern Europe</td>
<td>Family-20-Male</td>
<td></td>
<td>European</td>
</tr>
<tr>
<td>Family-10-Female</td>
<td>31</td>
<td>European</td>
<td>H</td>
<td>Europe</td>
<td>Family-21-Female</td>
<td>37</td>
<td>European</td>
</tr>
<tr>
<td>Family-10-Male</td>
<td></td>
<td>European</td>
<td>U</td>
<td>Eastern Europe</td>
<td>Family-21-Male</td>
<td></td>
<td>European</td>
</tr>
<tr>
<td>Family-11-Female ♦</td>
<td>30</td>
<td>Central South Asian</td>
<td>U</td>
<td>Eastern Europe</td>
<td>Family-22-Female</td>
<td>37</td>
<td>Middle East</td>
</tr>
<tr>
<td>Family-11-Male ♦</td>
<td></td>
<td>Central South Asian</td>
<td>M</td>
<td>Asia</td>
<td>Family-22-Male</td>
<td></td>
<td>Central South Asian or Middle Eastern</td>
</tr>
<tr>
<td>Family-8-E8</td>
<td></td>
<td>Central South Asian or European</td>
<td>F</td>
<td>Asia</td>
<td>Family-23-Female</td>
<td>37</td>
<td>European</td>
</tr>
<tr>
<td>Family-6-Relative-SJ</td>
<td></td>
<td>European</td>
<td>K</td>
<td>Native Europe</td>
<td>Family-23-Male</td>
<td></td>
<td>European</td>
</tr>
</tbody>
</table>

HAP: refers to haplotype. Mismatched couples are coloured in red. *Couples with different ethnicities determined by nDNA. ♦Couples distantly related based on mitochondrial haplotypes.
4.3.4 Nuclear genome template difference between individuals within the fertile and the RM/RIF couples

For these analyses, we compared the genetic similarity measured by the proportion of alleles that are identical by state (IBS), and the genetic distance measured by the proportion of alleles that are different by state (here abbreviated by 1-IBS). High density SNPs across the entire nuclear genome covered by the Infinium CoreExome-24 v1.2 Bead Chip were compared between individuals in each couple. The results for each couple were extracted from a triangular matrix (explained in Chapter 4 section 4.2.3). Cases (RM/RIF group) and controls (fertile group) genetic differences were measured. Statistical differences of the distributions were assessed via a two-sample “Kolmogorov-Smirnov” test. No significant genetic differences were observed between the fertile and RM/RIF groups (P-value = 0.1493) as shown in Figure 4-12.
Figure 4–12: A comparison of IBS between the fertile (PGD) and the infertile (RM/RIF) group of patients against all the SNPs present on the chip.

No significant differences were observed. A histogram of absolute allele frequency difference in SNPs between the fertile (PGD) and infertile (RM/RIF) groups is presented on the right. The number of SNPs extracted with absolute allele frequencies ≥ 0.4 were approximately 1000 SNPs.

Further analysis of these nuclear SNPs, showed that most of the SNPs with allele frequency difference greater than 0.4 between the males of the fertile and infertile group were on the X chromosome. Other SNPs were also present including one on NDUFA9 on chromosome 12 encoding a protein in complex I within the mitochondrial ETC. A total of 327 mtDNA SNPs are included in the HumanCore exome SNP chip. Differences between the fertile and infertile couples according to those SNPs were analysed (Figure 4-13). Only the top 20 SNPs with a frequency difference > 0.2 were extracted. Most of those SNPs were in mitochondrial complex I (Table 4-8).
Table 4-8: Mitochondrial SNPs that are different between the fertile and RM group.

<table>
<thead>
<tr>
<th>rCRS position</th>
<th>Diff %</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>9899</td>
<td>0.2</td>
<td>MT-CO3</td>
</tr>
<tr>
<td>4917</td>
<td>0.4</td>
<td>MT-ND2</td>
</tr>
<tr>
<td>12308</td>
<td>0.2</td>
<td>MT-TL2</td>
</tr>
<tr>
<td>15928</td>
<td>0.3</td>
<td>MT-TT</td>
</tr>
<tr>
<td>4917</td>
<td>0.4</td>
<td>MT-ND2</td>
</tr>
<tr>
<td>5004</td>
<td>0.2</td>
<td>MT-ND2</td>
</tr>
<tr>
<td>15452</td>
<td>0.3</td>
<td>MT-CYB</td>
</tr>
<tr>
<td>15928</td>
<td>0.3</td>
<td>MT-TT</td>
</tr>
<tr>
<td>10463</td>
<td>0.3</td>
<td>MT-TR</td>
</tr>
<tr>
<td>11251</td>
<td>0.3</td>
<td>MT-ND4</td>
</tr>
<tr>
<td>1888</td>
<td>0.3</td>
<td>MT-RNR2</td>
</tr>
<tr>
<td>12705</td>
<td>0.2</td>
<td>MT-ND5</td>
</tr>
<tr>
<td>11467</td>
<td>0.2</td>
<td>MT-ND4</td>
</tr>
<tr>
<td>2706</td>
<td>0.235</td>
<td>MT-RNR2</td>
</tr>
<tr>
<td>1811</td>
<td>0.2</td>
<td>MT-RNR2</td>
</tr>
<tr>
<td>15924</td>
<td>0.2</td>
<td>MT-TT</td>
</tr>
<tr>
<td>4024</td>
<td>0.2</td>
<td>MT-ND1</td>
</tr>
</tbody>
</table>

Figure 4–13: Mitochondrial SNPs differences between couples within the fertile and infertile groups.

327 mitochondrial SNPs were present on the Infinium CoreExome-24 SNP chip. 17 SNPs with a difference ≥ 0.2% have shown a difference between individuals in couples within the fertile and RM/RIF group. Details of the probe IDs (Appendix C section 7.3.3).

4.3.5 Differences between partners in couples in nuclear encoded mitochondrial SNPs and the MHC region

To further investigate the differences between individuals in couples, SNPs in nuclear genes encoding proteins important in mitochondrial function present on the human core exome SNP chip (n=959) were compared between individuals in couples of each category. Also, SNPs in the MHC region were extracted (n=5632) and the IBS differences between individuals in each couple were compared (Figure 4–14). No significant differences were found between couples within the fertile group.
compared to the infertile group for either the nuclear encoded mtDNA SNPs or the MHC.

![Box plots comparing IBS distances for fertile and infertile groups.](image)

**Figure 4-14:** IBS difference comparison between the PGD (fertile) and the infertile (RM/RIF) groups based on the nuclear encoded mitochondrial selected genes and MHC SNPs.

No significant difference was observed in both comparisons.

### 4.3.6 Karyomapping analysis with BlueFuse

All genomic samples presented SNP call rates above 98% and embryonic samples gave SNP call rates of around 85% and above (Table 4-9). These results exceeded threshold levels set up by manufacturer (80% for non-embryo and 60% for embryo samples) and met recommended values (95-99% blood, 75-95% blastomere biopsy, 85-99% trophectoderm biopsy) presented in Infinium Karyomapping Protocol Guide (Illumina, San Diego, USA).
Table 4-9: Karyomapping SNP call rates.

<table>
<thead>
<tr>
<th>sample ID</th>
<th>Origin</th>
<th>Call Rate %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family-4-Male</td>
<td>Genomic</td>
<td>98.25%</td>
</tr>
<tr>
<td>Family-4-Female</td>
<td>Genomic</td>
<td>98.25%</td>
</tr>
<tr>
<td>Family-4-E2</td>
<td>WGA product</td>
<td>96.37%</td>
</tr>
<tr>
<td>Family-4-E1</td>
<td>WGA product</td>
<td>91.90%</td>
</tr>
<tr>
<td>Family-6-Male</td>
<td>Genomic</td>
<td>98.22%</td>
</tr>
<tr>
<td>Family-6-Female</td>
<td>Genomic</td>
<td>98.24%</td>
</tr>
<tr>
<td>Family-6-E5</td>
<td>WGA product</td>
<td>84.93%</td>
</tr>
<tr>
<td>Family-6-E4</td>
<td>WGA product</td>
<td>85.20%</td>
</tr>
<tr>
<td>Family-11-Male</td>
<td>Genomic</td>
<td>98.21%</td>
</tr>
<tr>
<td>Family-11-Female</td>
<td>Genomic</td>
<td>98.24%</td>
</tr>
<tr>
<td>Family-11-E5</td>
<td>WGA product</td>
<td>93.58%</td>
</tr>
<tr>
<td>Family-11-E6</td>
<td>WGA product</td>
<td>87.48%</td>
</tr>
</tbody>
</table>

Before this study the sex of the embryos was unknown, so all were initially classified as males. Whenever this was inconsistent with the karyomapping data, BlueFuse showed a warning. As BlueFuse generates warning messages and full case reports only for samples defined as a ‘reference’, embryos from each family were classified as a reference one after another; this did not change SNP calls, neither did the quality of embryo chosen as a reference affect SNP calls. Additional confirmation of sexing was performed by analysis of B-Allele Frequency (BAF) chart of the sex chromosomes. The BAF charts for genomic samples presented more distinctive clusters than embryonic samples, which reflects differences in DNA quality. The BAF is a normalised measure of the allelic intensity ratio of two alleles (A and B). A BAF of 1 indicates the complete absence of A allele (genotype B/B or B/-), a BAF of 0.5 indicates the equal presence of both alleles (A/B) and a BAF of 0 represents absence of B allele (A/A or A/-). Homozygous deletions result in a failure of the BAF to cluster (Alkan et al., 2011), which was observed in Y chromosome region in female embryos. In male samples a cluster at BAF of 0.5
(A/B) was observed at Xq, which refers to a pseudo-autosomal region 2 (PAR2) located at Xq28. As a final result of the combined analysis, all embryos were sexed Family-4-E1 and Family-4-E2 as males and the rest as females.

4.3.7 Comparison between HumanKaryomap SNP chip & a-CGH for aneuploidy detection

Embryos which were WGA by MDA and diagnosed using both array CGH and the HumanKaryomap SNP chip showed high levels of discordance in the chromosomal aneuploidies detected. However, this was mainly due to segmental losses and gains detected using the aCGH platform which is sensitive (Appendix A section 7.1.1). It may also suggest that WGA by MDA requires further optimisation for this platform (Table 4-10, Figure 4–15).
Table 4-10: Comparison between HumanKaryomap SNP chip and a-CGH.

<table>
<thead>
<tr>
<th>Embryo ID</th>
<th>Genetisure (aCGH)</th>
<th>HumanKaryomap (SNP chip)</th>
<th>Chromosomal imbalances</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GAINS</td>
<td>LOSSES</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chr</td>
<td>Size (Mbp)</td>
<td>Position</td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td>7.5</td>
<td>17p13.3-16p12.3</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>38.9</td>
<td>8q22.3-8q24.3</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>33</td>
<td>0</td>
</tr>
<tr>
<td>Family-4</td>
<td>14</td>
<td>19q13.13-19q19.32</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>29.9</td>
<td>16q12.1-16q23.2</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>9</td>
<td>19q13.13-19q19.32</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>25.1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>26</td>
<td>22q11.1-22q13.2</td>
<td>0</td>
</tr>
</tbody>
</table>

Cells highlighted in red: discordant chromosomal imbalances. Cells highlighted in green: concordant chromosomal imbalances. Chr: chromosome. T: total number different chromosome imbalances detected by either platform. Concordant: number of identical chromosomal imbalances detected by both platforms/T. Discordant: number of different chromosomal imbalances detected by only one platform/T.
<table>
<thead>
<tr>
<th>Embryo ID</th>
<th>Genetisure (aCGH)</th>
<th>HumanKaryomap (SNP chip)</th>
<th>Chromosomal imbalances</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GAINS</td>
<td>LOSSES</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chr</td>
<td>Size (Mbp)</td>
<td>Position</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>199</td>
<td>2q24.3-2q31.2</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>123</td>
<td>8p23.1-8q24.22</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>73.5</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>19</td>
<td>21.1</td>
<td>19p13.3-19p12</td>
</tr>
<tr>
<td>Family-6-E4</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>17</td>
<td>7p13.1-7p11.2, 7q22-7q25.3</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>18</td>
<td>8q11.1-8q23</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>21</td>
<td>21q11.2-21q22.3</td>
</tr>
<tr>
<td>Family-6-E5</td>
<td>1</td>
<td>68</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>35</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>31</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>25</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>34</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>23</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>41</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>24</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>31</td>
<td>0</td>
</tr>
</tbody>
</table>

Cells highlighted in red: discordant chromosomal imbalances. Cells highlighted in green: concordant chromosomal imbalances. Chr: chromosome. T: total number different chromosome imbalances detected by either platform. Concordant: number of identical chromosomal imbalances detected by both platforms/T. Discordant: number of different chromosomal imbalances detected by only one platform/T.
<table>
<thead>
<tr>
<th>Embryo ID</th>
<th>Genetisure (aCGH)</th>
<th>HumanKaryomap (SNP chip)</th>
<th>Chromosomal imbalances</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GAINS</td>
<td>LOSSES</td>
<td>GAINS</td>
</tr>
<tr>
<td>Chr</td>
<td>Size (Mbp)</td>
<td>Position</td>
<td>Chr</td>
</tr>
<tr>
<td>1</td>
<td>122</td>
<td>5</td>
<td>99</td>
</tr>
<tr>
<td>2</td>
<td>209</td>
<td>2p25.2-2q37.3</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>3p26.3-3p12.3</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>43</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>114</td>
<td>10p15.3-10q26.3</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>40</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>22</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>9.7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>29</td>
<td>20p13-20q13.33</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>29</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>42</td>
<td>9</td>
<td>89</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>12</td>
<td>55</td>
</tr>
</tbody>
</table>

Cells highlighted in red: discordant chromosomal imbalances. Cells highlighted in green: concordant chromosomal imbalances. Chr: chromosome. T: total number different chromosome imbalances detected by either platform. Concordant: number of identical chromosomal imbalances detected by both platforms/T. Discordant: number of different chromosomal imbalances detected by only one platform/T.
A) Karyomapping analysis of E1 from family 4.

B) Array CGH analysis of E1 from family 4

Figure 4–15: Comparison between (A) Karyomapping and (B) Array CGH by GenetiSure.

The detected gain on chromosome 8 in Family 4-E1 as shown on BlueFuse software in (A) in comparison to Agilent Cytogenomics software (B). A whole gain for chromosome 8 was identified by the HumanKaryomap SNP chip while a segmental gain was reported by aCGH.
As shown in Figure 4-15 (A), the karyomapping analysis of E1 in family 4, has shown a whole gain in chromosome 8 and losses in chromosomes 16, 19 and 22. Array CGH analysis on the other hand, has detected the loss in chromosome 16 and segmental gains in chromosomes 8, 17, 20 and 22.

4.4 Discussion

One proposed theory for the cause of infertility is mitochondrial-nuclear genomes mismatch. According to Lane (Lane, 2011) as four out of five complexes of the OXPHOS system includes subunits encoded by both mitochondrial and nuclear genomes, it is likely that any mismatch between the two genomes slows down the electron transfer, resulting in an increase of reactivity of electron with oxygen, which results in a rise in free-radical leak (Lane, 2011). From that simple concept, fertility, fitness and lifespan differences would be explained in principle. The process of aging and the onset of age-related diseases within individuals can also be predicted based on that theory. A literature review was done to select nuclear genes that may be affected by the mismatch and/or have a role in mitochondrial function. This chapter will discuss the results obtained from sequencing the mitochondrial genome. This provides some information about the ethnic background from the mother since mitochondria are maternally inherited. We also explored the effect of the ethnicity determined by the nuclear genome to determine if differences in ethnicities between individuals in a couple affect fertility. All the samples were genotyped using the HumanCore exome SNP chip. As the HumanKaryomap SNP chip is routinely used in clinical analysis of embryos, it was also used to evaluate if the patients’ ethnicities could be determined. The results obtained from sequencing the selected nuclear genes and mitochondrial genome will be discussed in the context of the
mitochondrial-nuclear mismatch and the effect on human infertility as a driver of human evolution in Chapter 5.

4.4.1 Couples within the fertile group were more likely to be distantly related compared to the infertile group

Several mitochondrial haplogroups were identified reflecting the diverse ethnic population in London. Haplogroups included haplotypes: T, H, M, N, U, J, K & R0. According to the phylogenetic tree; most of those haplotypes were from a European origin (H: west and east of Europe, T: eastern European, Caucasian, J: Western Europe, middle eastern, N: Australian, M: East Asia, K: East and Western Europe, U: Eastern European and Caucasian). Overall more couples from the RM/RIF group (5/12) had identical mtDNA haplotypes compared to the fertile group. No clear evidence of association between certain mtDNA haplotypes and RM/RIF was found, however, partners in the RM/RIF group were more closely related in their mitochondrial haplotype compared to the fertile group (p=0.04). Despite our small number of samples there was a trend suggesting that incompatibilities in mitochondrial haplotypes are beneficial for fertility. This is consistent with a previous study showing benefits in unmatched nDNA/mtDNA haplotypes in mice (Latorre-Pellicer et al., 2016).

Several groups have performed a comparison between fertile and RM females by examining the variations and heteroplasmy levels within the mitochondrial genome in female partners without considering the male mitochondrial factor effect. It is believed that if the male mitochondria had an effect, it will only be during fertilisation since the bulk of the sperm mitochondria does not enter the oocyte.

The role of mtDNA mutations in miscarriages was studied by screening women with RM for heteroplasmic mtDNA variations. Nineteen different variations were
found within the RM group but no evidence for accumulation of pathogenic mtDNA variations was found (Kaare et al., 2009). Therefore, the detected variations were thought to be polymorphisms, and not likely to contribute to the miscarriages experienced by the studied women.

Previously the presence a missense mutation T4216C was observed in 9% of women with RM (mainly homoplasmic) while the fertile women with proven fertility showed this mutation in a mainly heteroplasmic condition (Vanniarajan et al., 2011). The heteroplasmic nature of the mutation in the patients showed the significance of the mutation in the pathogenesis of RM. This missense mutation T4216C has been reported as a haplogroup JT-specific marker in some populations (Herrnstadt et al., 2002).

The unique approach of this study, is that we considered the male mitochondrial genome as a surrogate marker for the possible combinations of nuclear genes (important in mitochondrial function) leading to a mismatch that may have an effect on embryo development. The comparison of haplotype distance between partners in fertile couples and couples experiencing RM is an indirect approach to reveal the possibility of a mitochondrial nuclear mismatch effect on fertility if present. If embryos from couples with proven fertility that were age matched with RM couples were available, this would have made the analysis more robust.

4.4.2 Mitochondrial haplogroup T is associated with poor blastocyst rate

Another type of analysis was performed for the fertile group based on the mtDNA haplotype and embryo development affecting blastocyst rate. The presence of the mitochondrial T haplogroup was associated with a decline in blastocyst rate formation by multiple linear regression analysis.
Several polymorphic markers that differentiate the T haplogroup are located in genes that encode complex I subunits (T4216C at ND1, A 4917G at ND2, and G13368A at ND5) which promote either amino acid substitutions (YrH at ND1 and DrN at ND2) or show no change (ND5). It was suggested that these changes could explain the apparently lower performance of complex I in T mtDNA carrying individuals (Ruiz-Pesini et al., 2000). However, these could not explain the more significant reduction in complex IV activity observed in some T individuals (Ruiz-Pesini et al., 2000). The most important mitochondrial functions, oxidative phosphorylation and apoptosis were found to be involved in the RM (Vanniarajan A et al., 2011). But, based on our data, no correlation was found between the presence of the T haplogroup and RM/RIF.

The mitochondrial haplogroup effect was not only related to RM and diseases, a study on drosophila have suggested that naturally occurring male haplotypes affect male fertility in vivo (Yee et al., 2013). In this regard, Ruiz-Pesini et al. (Ruiz-Pesini et al., 2000) reported an association between asthenozpermia and haplogroup T. Other studies however, such as Pereira et al. (Pereira et al., 2005) did not find mitochondrial haplogroup association in 101 southern Portugal oligozoospermic males using geographically well-matched controls suggesting that population stratification problems in the study of Ruiz-Pesini et al., might be the reason behind their findings highlighting the importance of using adequate geographic matched controls to avoid spurious associations. Whilst most of the mitochondrial haplogroups T were present in the male partners in our study, all fertilisation was achieved following ICSI, therefore sperm motility was not a factor in our observation.

Interestingly, in our analysis, the T haplotype was mainly from the male partner (Table 4-3). Although the mitochondrial component of male partner is not present in
the embryo since mitochondria are maternally inherited, we have considered the male mitochondrial haplogroup as a surrogate marker for possible combinations of nuclear genes that can be transmitted to the embryo leading to mitochondrial nuclear mismatch. This analysis suggests that the nuclear background where one partner has mitochondrial T haplogroup may significantly contribute to poor embryo progression to the blastocyst stage. This finding is in agreement with other reports supporting an association between some haplogroups and the susceptibility to some pathologies. For example, haplogroup J is associated with an increased risk of developing different degenerative diseases, while haplogroup H seems to act as a protective haplogroup against the development of age related macular generation (Kenney et al., 2013). A similar trend was observed in the incidence of aneuploidy, with the sister haplogroups J and T presenting a significantly higher incidence of chromosome errors in oocytes when compared with haplogroup H (Gianaroli et al., 2015). Accordingly, the efficiency of the electron transport chain and ATP production have been reported to be diminished for haplogroup J in comparison with haplogroup H (Marcuello et al., 2009). This could explain the highest incidence of prezygotic meiotic errors reported by Gianaroli et al. It is plausible that if a cell produces a reduced amount of ATP, not enough energy could be available to pull chromatids apart to the corresponding PB during the first meiotic division (Gianaroli et al., 2015). On the other hand, aneuploidy status analysis in our samples, showed that aneuploid embryos had poor morphology regardless of the presence of the mitochondrial T haplogroup. This could be due to the small samples size since only one female partner in the fertile group in our analysis was a carrier for the haplogroup T and the rest were males. Or it could be that the male partner effect is not related to aneuploidy and it is related to ATP and embryo development only.
Based on our hypothesis and findings, when the male partner is a carrier for mtDNA haplogroup T, 50% of his nuclear component which will be passed to the offspring, can affect post zygotic development of the embryo. In other words, since ATP is crucial for embryo development, certain haplogroups may affect the development process through differential levels of ATP production which is determined by the mitochondrial nuclear matching in the mitochondrial ETC.

Based on that, the T haplogroup seems to be a marker for detrimental embryo development. If it was carried by the mother, it will have an effect on prezygotic meiotic errors as well as post zygotic events leading to a higher rate of aneuploidy. Instead, if the father was a carrier for the T haplogroup, the alleles transmitted to the embryo will affect complex I activity in the embryo which will have a different mitochondrial background (inherited from the mother). Therefore, nuclear background is very important, but it is mainly based on what was the initial mitochondrial haplogroup of both parents. Clearly, more patients and a further analysis, including other infertility factors, are necessary to support this hypothesis that, if confirmed, would contribute relevant information not only on cell physiology but also at the clinical level.

Furthermore, in our study the association to haplogroup T was with the fertilisation rate based on the embryos that reached the blastocyst stage. So, the common factor with all what has been reported so far, is that the T haplogroup has a detrimental effect either causing a disease or affecting fertility outcome in some way. It could be hypothesised that mtDNA-nDNA mismatch would reduce the ATP production because of an inefficient respiratory chain, affecting embryo development. If the RNA was available from those embryos, we would have tested
this hypothesis by sequencing the RNA for the selected nuclear gene to test if the lowering of transcripts is associated with mitochondrial function and biogenesis.

4.4.2.1 Embryonic mtDNA haplogroup &embryo quality

Several mitochondrial embryonic haplogroups confirmed the complete inheritance from the mother. The mother’s haplogroup was used to infer the haplogroup for the embryos since it is maternally inherited. Therefore, analysis of embryos was also performed based on the inferred mtDNA haplogroup in relation to their aneuploidy status, morphology and mtDNA number. We aimed to explore the effect of the level of the polymorphisms in the sequenced embryos since the depth of read data showed inconsistency between the runs so we decided to focus on the haplotypes only. Heteroplasmy and mitochondrial quantification by NGS from WGA embryonic samples is very tricky to detect and what has been used in previous studies was Restriction Fragment Length Polymorphism (RFLP) which is very sensitive (Gigarel et al., 2005, Steffann et al., 2006) especially that those embryos were from families with no reported mitochondrial mutations.

In this study, further analysis was performed by linking the aneuploidy status and the mtDNA haplotype of the parents. No effect was found when the same test was applied to compare the effect on the proportion of the embryos that arrested and the embryos that reached the blastocyst stage in each family. While aneuploidy status confirmed a negative influence on embryo morphology (p=0.03), the mitochondrial haplogroup of the embryo did not have significant effect on embryo morphology when multiple linear regression analysis was performed. This finding does not necessarily conflict with what has been reported previously by Gianoroli et al (Gianoroli et al., 2014) who reported differences in susceptibility to aneuploidy by
some haplogroups especially T and J. Our data do not contradict those results but maybe makes it clearer for us to understand the different effect of mitochondrial-nuclear interaction on prezygotic and post zygotic meiotic errors. Fragouli (Fragouli et al., 2015) suggested that the greater the mtDNA number the more aneuploid is the embryo at the blastocyst stage. According to our data, we did not find an association between mtDNA number and aneuploidy in embryos tested at the blastocyst stage (as discussed in Chapter 3). However, we found that elevated mtDNA is significantly present in euploid embryos arrested at cleavage compared to embryos that developed to the blastocyst stage. When the same comparison was performed based on T haplotype carriers, no positive or negative associations were determined. This suggests that the mitochondrial haplogroup effect on embryo developmental stage is not correlated with mitochondrial number or aneuploidy in our cohort.

4.4.3 Nuclear genotype differences between partners (IBS) had no significant effect on fertility

In humans, consanguinity has been reported to increase the risk of complex adult diseases (Bener et al., 2012), the rates of neonatal and reproductive mortality and morbidity (Kerkeni et al., 2007, Bittles and Black, 2010, Lisa et al., 2015), and the recurrent risk of early deaths, infant mortality, and birth defects (Stoltenberg et al., 1999). However, its role in fertility remains controversial: in some studies inbreeding was associated with reduced fertility (Ober et al., 1999; Robert et al., 2009), while in others no significant association or even a positive association was found (Edmond and De Braekeleer, 1993, Saad and Jauniaux, 2002, Blanco Villegas and Fuster, 2006, Helgason et al., 2008, Labouriau and Amorim, 2008, Weller and Santos, 2013).
A positive association between consanguinity and late fertility has been reported reducing the negative impact of inbreeding (due to the increase in homozygosity of autosomal recessive detrimental mutations) and advanced maternal age in a Sardinian population (Lisa et al., 2015). Based on our results, differences in nuclear genotypes between partners did not have a significant effect on their fertility. However, couples with RM/RIF in our cohort showed more similarity between individuals of each couple compared with fertile couples. In general, this suggest that the interplay between mitochondrial and nuclear genomes may be important in terms of fertility.

4.4.4 Ancestry informative markers (AIMs) had no significant effect on fertility

AIMs were analysed using two SNP chips. Initially, the HumanKaryomap SNP chip was used since it is widely used for PGD screening and could serve for both purposes then by the Infinium Core Exome chip. Ethnicity based on the mitochondrial genome was not always confirmed by the nuclear genome suggesting that parents of individuals were of a different ethnic origin. The ethnicity of E8 in family 8 confirmed the inheritance of nuclear ethnic backgrounds from both parents. The mitochondrial haplotype for the female partner was Central south Asian and the male partner was European. The nuclear DNA confirmed these findings in the embryo (E8) which was classified as Central South or European by the nuclear SNPs.

Looking at PCA clustering of samples shows that partners in families \( n=2 \) within the RM/RIF group, were more distantly related in their nuclear genes compared with the fertile group. Even though it was not significant, it should be noted that this is
different to the mitochondrial data obtained earlier. Therefore, more samples are required to confirm this analysis.

When karyomapping is offered to a couple for PGD, additional information could be obtained and used for research purposes. Even though we were able to pull out the AIM SNPs, the allele changes were difficult to obtain, and we had to use Genome Studio to infer the actual DNA sequence. If this feature could be altered only for the AIMs and if the SNPs for the mitochondrial genome could be included in the Karyomap SNP chip, the ethnicity and mitochondrial haplotypes of the couple could be determined. This would generate useful data to establish if ethnicity and relatedness (based on mitochondrial haplogroups) could be used as additional predictors for the likelihood of achieving a successful pregnancy during the PGD work-up stage. A comparison of alleles based on SNPs in the nuclear genome on the whole Karyomap SNP chip to determine relatedness between the parents by IBS could also be used for predictive analysis.

4.4.5 Immunological factors have no effect on fertility outcome

Several studies associating the HLA compatibility and outcome following ART treatment were reported and contradictory evidence has accumulated. Some studies reported that maternal-fetal compatibility for HLA region genes may be a contributing factor in some couples with unexplained infertility due to repeated implantation failure. The first study of HLA sharing on RIF couples following ART was reported by Weckstein et al., 1991 where 7 of 10 couples suffering RIF shared 3 or more HLA antigens compared with 4 of 10 control subjects (Weckstein et al., 1991). Following that, two studies reported similar results (Balasch et al., 1993, Ho et al., 1994), suggesting that the increased HLA sharing among couples experiencing RIF following ART could be the reason behind their infertility. Compatibility for class II
region genes influence implantation failure (Coulam et al., 1987, Ober et al., 1992, Jin et al., 1995) as well as the loss of clinically recognised pregnancies (reviewed in (L.B.Olding, 1997). Compatibility for class I region genes may also influence recognised fetal losses, at least in an inbred population (Ober et al., 1992). Therefore, based on all those studies, it was hypothesised that incompatibility for HLA genes specifically class II region genes, is important for successful pregnancy outcome among both infertile and fertile couples.

On the other hand, inbred communities do not suffer from repeated miscarriage and infertility (Hamamy, 2012). Due to this contradictory evidence, we wanted to rule out if those difference between couples in their mtDNA haplotypes are not actually due to differences in the MHC if that was the case. The IBS analysis in the MHC region present in the human core exome SNP chip were analysed between partners in each couple. No significant differences were observed between the fertile and infertile group. If embryos were available for the RM/RIF group, we would have tested the maternal KIR AA genotype with a fetal HLA-C2. It has been reported that SNPs combinations of Maternal KIR and trophoblast HLA-C2 genes influence the risk of Preeclampsia and reproductive success through a physiological role related to placental development and not due to an immune function (Hiby et al., 2004).

4.5 Results summary

1. Couples within the fertile group were more likely to be distantly related in their mtDNA haplotypes compared to the infertile group (p=0.04).

2. Poor blastocyst formation was associated with the mitochondrial T haplotype in one partner of each couple p=0.002. In our cohort, most of the mitochondrial haplogroup T carriers were males.
3. Differences between partners in nuclear alleles in the RM/RIF group compared to the fertile group showed no significant effect on fertility.

4. No significant difference was observed when the analysis of the HLA SNP between individuals in a couple was conducted.

5. It is possible to obtain information for ethnicity and aneuploidy from karyomapping data. However, aneuploidy and ethnicity have no significant effect on embryo development based on our cohort.

### 4.6 Limitations and future work

The availability of samples from patients with RM and RIF was limited in this study. The analysis would have been more robust if more samples including embryos from the RM and RIF group of patients were available. If embryos from couples with proven fertility that were age matched with RM and RIF couples were available, rather than couples undergoing PGD with no known infertility, a better comparison could have been made.

With regards to haplotype analysis, even though London is a mixed population, due to the small sample size, most of the identified haplotypes were from Europe. Replicating this pilot study in the future by including couples from a range of ethnicities and population groups may identify other mitochondrial haplotypes associated with mitochondrial dysfunction and infertility.
Chapter 5 Incompatibilities between the mitochondrial and nuclear genomes that affect preimplantation embryo development

5.1 Aims summary

Chromosomal aneuploidies are a major cause of early reproductive losses and embryonic implantation failure. In addition to chromosomal abnormalities, poor preimplantation embryo development can be caused by maternal factors which include mitochondrial function. Proteins required for mitochondrial function are encoded by both mitochondrial DNA (mtDNA) and the nuclear DNA (nDNA) necessitating the coordination between the two genomes. Moreover, mtDNA transcription, replication, and translation are completely directed by factors encoded by nDNA (Woodson and Chory, 2008). Nuclear-encoded proteins must recognize and bind to regulatory motifs in mtDNA for proper function. Because of these close interactions, nuclear and mitochondrial genomes undergo adaptive co-evolution to maintain fitness in aerobic metabolism (Bayona-Bafaluy et al., 2005, Dowling et al., 2008, Pichaud et al., 2012, Wolff et al., 2014, Camus et al., 2015, Ma et al., 2016). The impact of nDNA-mtDNA incompatibility on the organismal level is less obvious since reproductive barriers prevent the birth of live offspring in most hybrids. For this reason, it has been proposed that mitochondrial-nuclear mismatch may cause embryonic cell death. This study explores the effect of nDNA genotypes of 53 nuclear encoded genes involved in mitochondrial function, by assessing preimplantation embryo development in relation to aneuploidy status and the mitochondrial haplogroup. Chapter 4 showed that couples where one partner had mitochondrial haplogroup T had embryos with poor development even when
fertilisation rates were high. No association between the mitochondrial haplotype of the embryo and aneuploidy was found.

In this study, the aim was to identify evidence of mitochondrial nuclear mismatch by assessing differences in nuclear genotypes from selected nuclear genes important in mitochondrial function by comparing variants between

- Embryos that degenerated/arrested at cleavage and embryos that developed to the blastocyst stage from fertile couples.
- A fertile group and a group of patients suffering repeated miscarriage (RM) or repeated implantation failure (RIF) for unknown reasons.

We aimed to find the location of key positions of interaction between proteins that are encoded by the mitochondrial genes with proteins encoded by the nuclear genes. We also aimed to find if aneuploidy has an effect or is affected by the mito-nuclear mismatch if present.

### 5.2 Materials & Methods

#### 5.2.1 Samples

In this cohort study the DNA of two groups was tested. The first group included 11 fertile couples undergoing PGD for single gene disorders and their corresponding embryos (n=57) which were diagnosed as affected following PGD or unsuitable for transfer following IVF. The embryos were collected on day 6-7 post fertilisation. The other group was composed of 12 couples suffering RM or RIF for unknown reasons. A detailed summary of the fertile couples and their embryos is shown in the Appendix B 7.2.3.
5.2.2 Nuclear encoded mitochondrial genes selection and assay design

Fifty-three nuclear genes important in mitochondrial function were selected from the literature. Other genes were also selected and used as controls.

The selected nuclear genes were submitted to Agilent to design a targeted nuclear genome sequencing assay that would provide the best coverage (design ID: 0719611). The assay was performed using specific online tools provided by the company together with the bioinformatics support. This included a total of 59 selected nuclear genes involved in mitochondrial function (n=53) (Figure 5-4) and other genes (n=6) which were important in other nuclear processes were used as a control. In this panel design it was decided to perform the assay in two probe groups which were then combined. The first probe group included the sequencing of the entire transcribed region of 8 genes with a moderately stringent masking for repetitive sequences and 5X tiling capacity. The second probe group included the sequencing of the exons and 3’,5’ UTRs of the genes with a more relaxed stringency. The target regions for each gene were defined and consisted of all the exons identified in NCBI RefSeq + / − 50 intronic flanking bp. The main characteristics of the assay is shown in Table 5-1.
Table 5-1: The characteristics of the sequencing assay designed by Agilent

<table>
<thead>
<tr>
<th>probe group</th>
<th>1</th>
<th>2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of genes</strong></td>
<td>8</td>
<td>51</td>
<td>59</td>
</tr>
<tr>
<td><strong>Target size</strong></td>
<td>283.838 kbp</td>
<td>254.947 kbp</td>
<td>538.785 kbp</td>
</tr>
<tr>
<td><strong>Online tool</strong></td>
<td>Sure design</td>
<td>Sure design</td>
<td>-</td>
</tr>
<tr>
<td><strong>Probe size</strong></td>
<td>226.939 kbp</td>
<td>289.436 kbp</td>
<td>516.375 kbp</td>
</tr>
<tr>
<td><strong>Number of probes</strong></td>
<td>10914</td>
<td>19812</td>
<td>30726</td>
</tr>
<tr>
<td><strong>Region</strong></td>
<td>Entire transcribed region</td>
<td>Coding Exons + UTRs + 5′ UTR + 3′ UTR</td>
<td>-</td>
</tr>
<tr>
<td><strong>Region extension</strong></td>
<td>25 bases from each 3′/5′ end</td>
<td>25 bases from each 3′/5′ end</td>
<td>-</td>
</tr>
<tr>
<td><strong>Mean probes per base</strong></td>
<td>5</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td><strong>Boosting</strong></td>
<td>maximum performance</td>
<td>maximum performance</td>
<td>-</td>
</tr>
<tr>
<td><strong>extension into repeats</strong></td>
<td>20</td>
<td>20</td>
<td>-</td>
</tr>
<tr>
<td><strong>Design stringency</strong></td>
<td>Moderate masking</td>
<td>Least</td>
<td>-</td>
</tr>
<tr>
<td><strong>Strand</strong></td>
<td>Sense</td>
<td>Sense</td>
<td>-</td>
</tr>
</tbody>
</table>

The selected genes were sequenced from parental genomic DNA samples. Also, the embryonic samples were sequenced following whole genome amplification by MDA using Sure Select QXT (Agilent, UK). The sequencing was performed over four independent sequencing runs (of 30 patients for each method including repeats and replicates). All samples were normalised to 25ng/µl prior to sequencing. Samples with low concentrations were vacuum centrifuged. Some embryos had a low 230/260 ratio (less than 2) on NanoDrop. A clean-up step using the GE healthcare kit was attempted but this did not improve yields, therefore this step was abandoned. Following the initial fragmentation and adaptors tagging steps, samples were tested by the Bioanalyzer and the TapeStation to ensure the experiments had worked and to measure the DNA concentration for the following steps. Replicates and repeats were used to ensure that there was no contamination and that the results on 28 samples were reproducible.

The analysis of the fertile group was linked to the previously identified mitochondrial haplogroups and embryonic aneuploidy analysis (Appendix B section...
7.2.3). Analysis was performed by identifying nDNA genotypes that only occurred in the embryos with poor morphology (arrested/ degenerated at cleavage). Bioinformatics data clean-up included several filtering steps to avoid the effect of whole genome amplification removing missing variants, uninformative variants and inconsistent variants between the mother and father.

As mitochondria in embryos are maternally derived, differences in the embryonic combination of SNPs and indels of nuclear genes were assessed and compared to the maternal haplotype per family to identify evidence of mitochondrial mismatch in arrested embryos. A summary of the work flow is presented in Figure 5–1.

![Figure 5–1: An illustration of the workflow applied for sequencing the targeted nuclear encoded mitochondrial genes of parental DNA samples as well as their embryos. The nuclear genome of the fertile group, RM couples as well as all the WGA embryos were sequenced.](image-url)
5.2.3 Nuclear genome sequencing and haplotyping

5.2.3.1 Enrichment methods
The libraries were prepared following the instructions from the manufacturer “SureSelect QXT “Target Enrichment for Illumina Multiplexed Sequencing, Version C0, January 2015”. All experiments were prepared following the provided protocol SureSelectQXT (Agilent Technologies, Santa Clara, CA, USA). Firstly, DNA samples from couples and whole genome amplified products from embryos (MDA products) were quantified by Qubit high sensitivity assay kit and diluted accordingly to 25 ng/µL. Then, DNA was fragmented, and adaptors were added in a single enzymatic step. The adaptor-tagged DNA libraries were purified by magnetic beads and then amplified. Following that, 750ng/µL of each library was hybridised using SureSelectQXT capture library for 90 minutes. The resulting libraries were recovered using streptavidin magnetics beads, and a post-capture PCR amplification was carried out.

5.2.3.2 Assessment of DNA library quality and quantity
Final libraries were quantified by the Qubit High Sensitivity kit (Invitrogen, Carlsbad, CA). The quality of the library was assessed on a Bioanalyzer High Sensitivity DNA chips or the Tape Station as described in Chapter 2 section 2.4.1.5.1.

5.2.3.3 Sequencing
Thirty samples were included in each run. Sequencing of the libraries was carried out on the MiSeq and NextSeq instruments (Illumina). The library was diluted to a final concentration of 4nM for samples loaded on the MiSeq and 2nM when samples were loaded on the NextSeq. According to the manufacturer’s protocol “Preparing libraries for sequencing on the MiSeq”, a final library concentration should be
ranging from 8 to 10pM to carry out cluster generation and sequencing on a standard Flow Cell and MiSeq Reagent Kit v2 300 cycles (2 × 150 cycles) or v3 150 cycles. Loading was performed as described in Chapter 4 section 4.2.1.3.

5.2.3.4 Data analysis

The de-multiplexing of the samples, generation of FASTQ files, and alignment against the reference with Burrows-Wheeler Aligner (BWA) software was performed using the SureCall software (Agilent). Resulting primary BAM files were then treated with Picard (http://broadinstitute.github.io/picard) to mark duplicates, genome analysis tool kit (GATK) to perform realignment around indels and SAMtools for merging, sorting and indexing BAM files.

5.2.3.4.1 Variants filtering strategy

1. Removal of all the positions for which there was at least one value missing. The analysis was about the respective proportions of the different variants, and for this a full set per position was required (meaning variants should be called for all individuals (Mother, father and all embryos).

2. Exclusion of inconsistent variations. The consistency of the embryo's genotypes and the parents was checked. In some cases, some embryos were not consistent. All variants showing this were excluded from the analysis because removing only non-consistent embryos would bias the analysis.

3. Filtering of non-informative alleles in the parents (e.g. mother AA father AA).

4. Pooling variant positions per genotype to be able to quantify the differences in genotypes. For example: mother and father both heterozygous in the same way (e.g. mother AC father AC) n=70. Embryos were counted (E1: AA, E2:CC and E3:AC/CA, p as in possible embryo from the genotypes of the
parents). A Chi-squared test was performed to see whether there was a difference in the frequencies of the genotypes between blastocysts (good) and arrested at cleavage (bad) embryos.

5. Disregarding Mendelian frequencies. For example: Mother AC father CC was considered the same as mother CC and father CA.

A lot of uninformative variants in the final set were found. A script was written with the help of Dr Andrews at the Babraham institute, to filter these out and count how many of each different type of failure was obtained (Appendix D section 7.4.1). The analysis pipeline is presented in the following diagram (Figure 5-2).

### 5.2.3.5 Analysis of SNPs predicted function

Identified SNPs in introns and 3’UTR regions were subjected to further analysis with Ensemble variant effect predictor for the biological consequence prediction. The list of SNP rsIDs for each gene was copied into the variant effect predictor and the outcome of analysis was exported as Excel file for further analysis.
Reordered BAM files (rBAM) are BAM files where any non-informative SNP between parents and embryos was excluded. The genome analysis toolkit (GATK) was used to call variants. BWA: Burrows-Wheeler Aligner (BWA) was used to rebuilds local consistent reads to local regions.

Figure 5–2: Target Enrichment Analysis pipe
5.3 Results

5.3.1 Candidate genes selection for mitochondrial-nuclear mismatch investigation

Six genes encoding mitochondrial subunits within the mitochondrial electron transport chain were chosen to aid in the selection of directly interacting nuclear genes encoding proteins which may be affected by the mismatch between the two genomes leading to embryonic developmental arrest. The mitochondrial encoded subunit chosen were the following: ND2, ND4 and ND5 in complex I, MTCYB in complex III and the two subunits involved in gating proton channels in complex IV (CO1 and CO3). The nDNA-encoded subunits constitute excellent candidates because they interact with selected mtDNA-encoded subunits are: NDUFC2 and NDUFA1 in complex 1, UQCRB in complex III and Cytochrome C which interacts with complex IV (Figure 5-3).

![Figure 5-3](image)

**Figure 5–3:** A graphic representation of the mitochondrial and nuclear encoded mitochondrial proteins including the genes that were focused on. Nuclear encoded subunits are coloured in blue while the mitochondrial encoded subunits are coloured in red. Adapted from The New England Journal of Medicine, 348, DiMauro and Schon, Mitochondrial respiratory-chain diseases, pages 2656-2668, Copyright (2003), with permission from Massachusetts Medical Society.
Other genes were also selected and added to the sequencing panel since they encode proteins which are important in mitochondrial function. Thus, genes were classified into three groups. The main group included the genes encoding directly interacting proteins within the ETC which are in complex I, III & complex IV. These genes encode proteins which directly interact with the mitochondrial encoded subunits. The second group included proteins encoded by nuclear genes which are important for mitochondrial function. These are mainly proteins and transcription factors encoded by genes involved in mitochondrial biogenesis, transcription and replication and other mitochondrial related cellular processes. Finally, the last group included proteins encoded by nuclear genes that are important in other cellular processes which were used as control genes. A summary of all selected proteins in represented in Figure 5–4.

Figure 5–4: The selected nuclear genes for mitochondrial-nuclear mismatch investigation.
A total of 59 genes were selected. Those genes were classified into three groups. The main group included genes encoding directly interacting proteins within the electron transport chain (ETC) shown in the middle between both circles. Those are in complex I, III & complex IV which directly interact with the mitochondrial encoded proteins stated. The second group includes nuclear genes encoding proteins which are important for mitochondrial function shown on the red circle. Finally, the last group included nuclear genes encoding proteins which have other cellular processes.
5.3.2 Targeted next generation sequencing

5.3.2.1 Assessment of DNA library quantity and quality

Following the first fragmentation and amplification of the adaptor tagged DNA samples, all samples concentrations were measured by the Bioanalyzer and TapeStation. This was also performed to verify that the fragmentation step had worked efficiently on the whole genome amplified (WGA) embryonic samples by MDA. Other samples were verified using the TapeStation which confirmed the same results. Figure 5-5 shows an example of the pre-capture analysis of amplified DNA library using the Bioanalyzer. The Bioanalyzer provided two outputs: Illustrative sample electropherograms showing the size and quality of the amplified products and gels are graphically presented to determine the quantity.

Other samples tested by the TapeStation showed similar results as shown in Figure 5-6. Based on the Bioanalyzer and TapeStation peak heights, the pre-capture DNA libraries were diluted to 750ng/µL based on the size of designed custom capture library (<=3 Mb).
Figure 5–5: The pre-capture analysis of amplified DNA library amplicons using the Bioanalyzer

As illustrated, the peak of DNA fragment size is positioned between 245 to 325 base pairs. This confirmed that the library had amplified with the expected size range. The table aligns with the gel tracks of the amplified libraries and identifies the sample concentrations in ng/µL.
Figure 5–6: The pre-capture analysis of amplified DNA library amplicons using the TapeStation.

As illustrated, the peak of DNA fragment size is positioned between 245 to 325 bp. The table aligns with the gel tracks of the amplified libraries and identifies the sample concentrations in nM.
5.3.2.2  Assessment of indexed library DNA quantity and quality

Following the hybridisation and capture steps libraries were then assessed again using the TapeStation. Similarly, the amplified indexed DNA libraries were analysed using the High Sensitivity DNA kit. The peak of DNA fragment size was verified to be positioned between 325 and 450 bp for the genomic samples and embryos respectively.

![Graph](image)

**Figure 5–7:** Post capture analysis of amplified indexed library DNA using the TapeStation. The peak DNA fragment size was verified by the electropherogram to be positioned between 325 and 450 base pairs. The table aligns with the gel tracks of the amplified libraries and identifies the sample concentrations in pg/µL.
According to the peak heights, samples were diluted and pooled to a final concentration of 4nM for multiplexed sequencing by the MiSeq. Following that, samples were sequenced using the MiSeq and NextSeq platforms. A total of four runs were performed. The details of all the sequencing runs are shown in (Table 5-2).

Table 5-2: Details of all the runs performed for sequencing the nuclear encoded genes

<table>
<thead>
<tr>
<th>Run</th>
<th>Platform</th>
<th>Number of samples</th>
<th>Total number of reads</th>
<th>Average depth of read per sample</th>
<th>%&gt;=Q30*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MiSeq</td>
<td>5</td>
<td>15,488,541</td>
<td>1240X (626 Mb)</td>
<td>94.61%</td>
</tr>
<tr>
<td>2</td>
<td>MiSeq</td>
<td>30</td>
<td>21,117,652</td>
<td>206.6X (107.4 Mb)</td>
<td>95.5%</td>
</tr>
<tr>
<td>3</td>
<td>MiSeq</td>
<td>34</td>
<td>18,405,832</td>
<td>158.3 X (82.3Mb)</td>
<td>96.44%</td>
</tr>
<tr>
<td>4</td>
<td>NextSeq</td>
<td>64</td>
<td>121,379,674</td>
<td>552.6 X (287.4 Mb)</td>
<td>83.94%</td>
</tr>
</tbody>
</table>

*Q30: is an indication of % of bases called at a quality of 30Q or above. The average depth of read per sample was calculated by the following formula: \( X = \frac{\text{Total number of reads (million)} \times \text{read length(bp)} \times 2}{\text{sample number}}\) / target region (0.52Mb).

5.3.3 Analysis of variants compared between blastocysts & arrested or degenerated embryos

Several variants were filtered based on the quality control checks described earlier in 5.2.3.4.1. Data cleaning removed the variants with missing, uninformative, inconsistent alleles and/or showing the possibility of allele drop out (ADO). For some families, there were virtually no variants remaining (possibly because the coverage in one sample was very poor, but in other hundreds of variants were available for analysis. Analysis of SNPs that gave sequencing results that were not suspected of having ADO in embryos were compared (n=4204). SNPs associated with all arrested embryos were compared with the SNPs present in the embryos that developed to the blastocyst stage. SNPs were identified within families rather than between families. This could be because of the small numbers of samples. Haplotypes per gene were established based on those SNPs. This analysis highlighted certain genes within four families.
5.3.3.1 Identified SNPs Families 3 and 4

Analysis of sequences in embryos from the selected nuclear genes identified 4 SNPs in the COQ9 gene in embryos with poor preimplantation development in two different families; Family-3 and Family-4 (Figure 5-8) SNP rs223861 is a C/T single-nucleotide variation on human chromosome 16 which has been reported to have a regulatory function in the promoter region of COQ9 gene.

Other SNPs with regulatory effects were also identified in family 2 (Table 5-3). rs28673064 is a T/A single nucleotide variation on human chromosome 3 which has a regulatory function on TP63 gene promoter flanking region. Furthermore, rs11574853 in NFKB2 is a single nucleotide variation on human chromosome 10 with a regulatory feature on open chromatin region allele A. Interestingly In both families, mitochondrial haplogroup T was identified in the male partner.

Table 5-3: summary of the genes associated with poor embryo development in Family-3.

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Gene</th>
<th>Variant position</th>
<th>SNP ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr3</td>
<td>TP63</td>
<td>189349247</td>
<td>rs28673064*</td>
</tr>
<tr>
<td>chr3</td>
<td>OPA1</td>
<td>193365939</td>
<td>rs9831772</td>
</tr>
<tr>
<td>chr3</td>
<td>OPA1</td>
<td>193386657</td>
<td>rs12494647</td>
</tr>
<tr>
<td>chr7</td>
<td>PRKAG2</td>
<td>151478602</td>
<td>rs12374732</td>
</tr>
<tr>
<td>chr10</td>
<td>NFKB2</td>
<td>104157727</td>
<td>rs4919632</td>
</tr>
<tr>
<td>chr10</td>
<td>NFKB2</td>
<td>104161796</td>
<td>rs11574853*</td>
</tr>
<tr>
<td>chr16</td>
<td>COQ9</td>
<td>57481670</td>
<td>rs223861*</td>
</tr>
<tr>
<td>chr16</td>
<td>COQ9</td>
<td>57488613</td>
<td>rs686402</td>
</tr>
<tr>
<td>chr16</td>
<td>COQ9</td>
<td>57490721</td>
<td>rs223865</td>
</tr>
<tr>
<td>chr16</td>
<td>COQ9</td>
<td>57494178</td>
<td>rs223868</td>
</tr>
</tbody>
</table>

*Variants with regulatory function. TP63 – regulates mitochondrial dependent apoptosis pathway, NFKB2-transcription factor, COQ9- electron transporter in the ETC.

Aneuploidy analysis with the position of the genes confirmed that a loss of a copy of the relevant chromosomes is not the reason behind the homozygosity. In Family-3, losses or gains in chromosome 16 were not identified in any embryos. Some gains and losses were present in E1 and E3 on the other genes. In E1 a gain on chromosome 3 and chromosome 7 and loss on chromosome 10, while on E3 a gain
on chromosome 10. This confirms that at least in COQ9 the aneuploidy was not the cause of homozygosity. On the other hand, in Family-4: there was a gain on the region identified by a-CGH on COQ9 in E1 while none on the other genes were included within the duplications or deletions. Embryo 2 and E4 showed no gains or losses since they were diagnosed as normal, while in E7 gains were identified in all genes except a loss in CoQ9. Since it is not concordant in all the samples then the homozygosity is not an issue.
Figure 5–8: Analysis of SNPs associated with embryos arrested at the cleavage stage in Family 3 and Family 4.

As shown in the pedigrees above, several SNPs have been identified. Four of those SNPs have shown regulatory effects such as; rs28673064 (promoter flanking region on TP63), rs11574853 & rs4919632 (open chromatin region in NFKB2), rs223861 (promoter C on COQ9). SNPs in COQ9 have been associated with poor embryo morphology in both families.
5.3.3.2 Identified SNPs in Family-9

Several SNPs were associated with the poorly developing embryo in Family 9 (Figure 5-9). The associated haplotypes spanned a large region on chromosome 4. This region mainly included PPARGC1A gene which is important in mitochondrial biogenesis. Other SNPs were identified within non-mitochondrial related genes such as CHRNA1 and NFE2L2 on chromosome 2. SNP rs7560774 has a regulatory effect on CHRNA1 gene.

Table 5-4: Summary of the genes associated with poor embryo development in Family-9.

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Gene</th>
<th>Variant position</th>
<th>SNP ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr2</td>
<td>CHRNA1</td>
<td>175629021</td>
<td>rs7560774*</td>
</tr>
<tr>
<td>chr2</td>
<td>NFE2L2</td>
<td>178257214</td>
<td>rs3813259*</td>
</tr>
<tr>
<td>chr5</td>
<td>PPARGGC1B</td>
<td>149232307</td>
<td>rs26119*</td>
</tr>
<tr>
<td>chr4</td>
<td>PPARGC1A</td>
<td>23856492</td>
<td>rs7672915</td>
</tr>
<tr>
<td>chr4</td>
<td>PPARGC1A</td>
<td>23875834</td>
<td>rs10028665</td>
</tr>
<tr>
<td>chr4</td>
<td>PPARGC1A</td>
<td>23876186</td>
<td>rs10006687</td>
</tr>
<tr>
<td>chr4</td>
<td>PPARGC1A</td>
<td>23886323</td>
<td>rs2946385</td>
</tr>
<tr>
<td>chr4</td>
<td>PPARGC1A</td>
<td>23887454</td>
<td>rs2970877</td>
</tr>
<tr>
<td>chr4</td>
<td>PPARGC1A</td>
<td>23893017</td>
<td>rs2970876</td>
</tr>
<tr>
<td>chr4</td>
<td>PPARGC1A</td>
<td>23893274</td>
<td>rs17576121</td>
</tr>
<tr>
<td>chr4</td>
<td>PPARGC1A</td>
<td>23895209</td>
<td>rs4383605</td>
</tr>
<tr>
<td>chr4</td>
<td>PPARGC1A</td>
<td>23897314</td>
<td>rs61795680</td>
</tr>
<tr>
<td>chr4</td>
<td>PPARGC1A</td>
<td>23897809</td>
<td>rs4536908</td>
</tr>
<tr>
<td>chr4</td>
<td>PPARGC1A</td>
<td>23899510</td>
<td>rs13127297</td>
</tr>
<tr>
<td>chr4</td>
<td>PPARGC1A</td>
<td>23901275</td>
<td>rs2970860</td>
</tr>
</tbody>
</table>

*Variants with regulatory function. CHRNA1 - non-mitochondrial related, NFE2L2- transcription factor, PPARGC1B- mitochondrial biogenesis, PPARGC1A - mitochondrial biogenesis.

Comparison of the aneuploidy data with the position of the genes has shown some gains and losses on those regions in embryos which showed multiple aneuploidies. However, a loss in chromosome 4 and the short arm of chromosome 5 were only identified in E5. Thus, a lost copy of chromosome 4 might have been the reason behind the homozygosity on that embryo only. E11, has shown a gain in chr5 and E12 has shown losses in chromosomes two and five. Thus, a loss of homozygosity in chromosome 4 was only identified in one E5.
Figure 5-9: Pedigree showing the analysis of SNPs associated with the arrested embryos in Family-9.

Embryos E5, E8, E11 arrested at day 6-7 post fertilisation therefore, were poorly developed embryos. Embryos E10, E12 were classified as better-quality embryos since they developed to the blastocyst stage.
5.3.3.3 Identified SNPs in Family-10

Only two SNPs have been associated with the poorly developing embryo in Family 10. The associated SNPs are shown in Table 5-5 and Figure 5-10. Homozygosity in E4 is not due to aneuploidy on those genes.

Table 5-5: Summary of the genes associated with poor embryo development in Family-10

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>mtDNA haplotype</th>
<th>Embryo grade</th>
<th>Aneuploidy status</th>
<th>Chr</th>
<th>chr2</th>
<th>chr3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family-10-Female</td>
<td>H</td>
<td>-</td>
<td>-</td>
<td>C/A</td>
<td>G/C</td>
<td></td>
</tr>
<tr>
<td>Family-10-Male</td>
<td>U5b2a2b</td>
<td>-</td>
<td>-</td>
<td>C/A</td>
<td>G/C</td>
<td></td>
</tr>
<tr>
<td>Family-10-E1</td>
<td>H</td>
<td>Developed to blastocyst</td>
<td>N/A</td>
<td>C/C</td>
<td>C/C</td>
<td></td>
</tr>
<tr>
<td>Family-10-E4</td>
<td>H</td>
<td>Arrested at cleavage</td>
<td>Aneuploid</td>
<td>A/A</td>
<td>G/G</td>
<td></td>
</tr>
<tr>
<td>Family-10-E7</td>
<td>H</td>
<td>Degenerated</td>
<td>N/A</td>
<td>A/A</td>
<td>G/G</td>
<td></td>
</tr>
</tbody>
</table>

RNASEH1 - mitochondrial replication, MFN1 - mediator of mitochondrial fusion. Aneuploidy in E4 was a gain in chromosome 17.

Figure 5–10: The associated SNPs with family 10 as displayed on IGV
5.3.3.3.1 Analysis of Indels

The analysis was first based on the percentage of similarity between embryos to their mother since they have the same mitochondrial haplotype as well. However, with such type of analysis a slight difference was observed by logistic random regression analysis (OR = 1.03 (95% CI 1.00 to 1.06), p = 0.049). As the indel percentage difference from the mother increases by one, the odds of an embryo to develop to the blastocyst stage increases by 3%. The same analysis was performed for a set of nuclear non-mitochondrial related genes (n=6) which showed no significant effect, odds ratio (OR) = 0.998 (95% CI 0.97 to 1.02), p = 0.84 especially when a comparison based on indels was performed. The informativity is higher in indels and it is difficult to do such analysis when there is a lot of filtering performed.

On the other hand, we focused on the analysis of indels associated with the poorly developing embryos in each family. All the indels that were associated with poorly developing embryos per family are shown (Table 5-6).

Further analysis of these indels was performed. Indels in regulatory regions were present in \textit{RNASEH1} that are important in mitochondrial replication, and \textit{CHRNA1}, which is not mitochondrial related but a mutation in that gene was identified to cause recurrent fetal loss and reveal the diagnosis of a lethal human phenotype (Shamseldin et al., 2013).

The other indels were mainly in \textit{PPARGC1A}, which is important in mitochondrial biogenesis; \textit{TFB1M}, which is important in mtDNA transcription; \textit{UQCRB}, which is important in ETC; and \textit{PRKAB1}, which is involved in kinase phosphorylation. Indels in COQ9- electron transporter in the ETC were present in the same embryos from the two families mentioned earlier. These indels were in LD with the previously mentioned SNPs.
Table 5-6: The indels associated with the poorly developing embryos

<table>
<thead>
<tr>
<th>Family ID</th>
<th>Sample ID</th>
<th>mtDNA haplotype</th>
<th>Embryo grade</th>
<th>Aneuploidy status</th>
<th>Chromosome</th>
<th>chr2</th>
<th>chr2</th>
<th>chr2</th>
<th>chr4</th>
<th>chr5</th>
<th>chr6</th>
<th>chr8</th>
<th>chr8</th>
<th>chr12</th>
<th>chr16</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Gene</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Mother</td>
<td>H</td>
<td>Developed to blastocyst</td>
<td>N/A</td>
<td>RNASEH1</td>
<td>3592674*</td>
<td>3592368</td>
<td>175624107*</td>
<td>23780272</td>
<td>149226206</td>
<td>155578324</td>
<td>97242133</td>
<td>972450061</td>
<td>120110459</td>
<td>57493802</td>
</tr>
<tr>
<td></td>
<td>Father</td>
<td>U5b2a2b</td>
<td></td>
<td></td>
<td></td>
<td>CT/C</td>
<td>CT/C</td>
<td>CT/C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E1</td>
<td>H</td>
<td>Arrested at cleavage</td>
<td></td>
<td>ANEUPLOID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E4</td>
<td>H</td>
<td></td>
<td></td>
<td>N/A</td>
<td></td>
<td>CT/CT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E7</td>
<td>H</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CT/CT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Mother</td>
<td>H31a</td>
<td>Arrested at cleavage</td>
<td>ANEUPLOID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACTA</td>
<td>CAAAT/C</td>
</tr>
<tr>
<td></td>
<td>Father</td>
<td>T1a1b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACTA</td>
<td>CAAAT/C</td>
</tr>
<tr>
<td>E1</td>
<td>H31a</td>
<td>Arrested at cleavage</td>
<td></td>
<td>ANEUPLOID &gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACTA</td>
<td>CAAAT/C</td>
</tr>
<tr>
<td>E2</td>
<td>H31a</td>
<td>Arrested at cleavage</td>
<td></td>
<td>MULTIPLE ANEUPLOID &gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACTA</td>
<td>CAAAT/C</td>
</tr>
<tr>
<td>E3</td>
<td>H31a</td>
<td>Arrested at cleavage</td>
<td></td>
<td>ANEUPLOID &gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACTA</td>
<td>CAAAT/C</td>
</tr>
<tr>
<td>4</td>
<td>Mother</td>
<td>U5b3b1</td>
<td>Arrested at cleavage</td>
<td>ANEUPLOID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACTA</td>
<td>CAAAT/C</td>
</tr>
<tr>
<td></td>
<td>Father</td>
<td>T2b3+151</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACTA</td>
<td>CAAAT/C</td>
</tr>
<tr>
<td>E1</td>
<td>U5b3b1</td>
<td>Arrested at cleavage</td>
<td></td>
<td>ANEUPLOID &gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACTA</td>
<td>CAAAT/C</td>
</tr>
<tr>
<td>E2</td>
<td>U5b3b1</td>
<td>Arrested at cleavage</td>
<td></td>
<td>ANEUPLOID &gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACTA</td>
<td>CAAAT/C</td>
</tr>
<tr>
<td>E3</td>
<td>U5b3b1</td>
<td>Arrested at cleavage</td>
<td></td>
<td>ANEUPLOID &gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACTA</td>
<td>CAAAT/C</td>
</tr>
<tr>
<td>E4</td>
<td>U5b3b1</td>
<td>Arrested at cleavage</td>
<td></td>
<td>ANEUPLOID &gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACTA</td>
<td>CAAAT/C</td>
</tr>
<tr>
<td>E7</td>
<td>U5b3b1</td>
<td>Arrested at cleavage</td>
<td></td>
<td>ANEUPLOID &gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ACTA</td>
<td>CAAAT/C</td>
</tr>
<tr>
<td>9</td>
<td>Mother</td>
<td>H2b8a1</td>
<td>Developed to blastocyst</td>
<td>EUPLOID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Father</td>
<td>T2b30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E10</td>
<td>H2b8a1</td>
<td>Arrested at cleavage</td>
<td></td>
<td>ANEUPLOID &gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E11</td>
<td>H2b8a1</td>
<td>Arrested at cleavage</td>
<td></td>
<td>ANEUPLOID &gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E12</td>
<td>H2b8a1</td>
<td>Arrested at cleavage</td>
<td></td>
<td>ANEUPLOID &gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E5</td>
<td>H2b8a1</td>
<td>Arrested at cleavage</td>
<td></td>
<td>ANEUPLOID &gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E8</td>
<td>H2b8a1</td>
<td>Arrested at cleavage</td>
<td></td>
<td>ANEUPLOID &gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Mother</td>
<td>H4a1a3</td>
<td>Developed to blastocyst</td>
<td>EUPLOID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Father</td>
<td>H4b1b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E1</td>
<td>H4a1a3</td>
<td>Arrested at cleavage</td>
<td></td>
<td>ANEUPLOID &gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E3</td>
<td>H4a1a3</td>
<td>Arrested at cleavage</td>
<td></td>
<td>ANEUPLOID &gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E5</td>
<td>H4a1a3</td>
<td>Arrested at cleavage</td>
<td></td>
<td>ANEUPLOID &gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E6</td>
<td>H4a1a3</td>
<td>Arrested at cleavage</td>
<td></td>
<td>ANEUPLOID &gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E9</td>
<td>H4a1a3</td>
<td>Arrested at cleavage</td>
<td></td>
<td>ANEUPLOID &gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Variants in regulatory regions. RNASEH1 - mitochondrial replication, CHRNA1 - non-mitochondrial related, PPARGC1A - mitochondrial biogenesis, TFB1M - mtDNA transcription, UQCRB - important in ETC, PRKAB1- kinase phosphorylation, COQ9 - electron transporter in the ETC.
5.3.4 Analysis of SNPs and indels between the fertile and RM group

A total of 1476 SNPs and 235 indels were identified as different between the fertile and the infertile groups. However, differences on those variants between the two groups were not significant since they were present in the majority of the infertile group of individuals but not all. Two SNPs and two indels were associated with the majority of the infertile group of individuals in comparison with the fertile (Table 5-7).

Table 5-7: Comparison of variants between the fertile and RM/RIF group

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Position</th>
<th>SNP ID</th>
<th>Variant change in RM group</th>
<th>Gene</th>
<th>VEP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chr7</td>
<td>129297332</td>
<td>rs1882094</td>
<td>G/G</td>
<td>NRF1</td>
<td>Protein coding-modifier</td>
</tr>
<tr>
<td>Ch16</td>
<td>57494992</td>
<td>rs223869</td>
<td>C/A</td>
<td>COQ9</td>
<td>COQ10 Deficiency, fatal neonatal</td>
</tr>
<tr>
<td>chr3</td>
<td>179103007</td>
<td>CT/CT</td>
<td></td>
<td>MFN1</td>
<td>Intronic variant</td>
</tr>
<tr>
<td>chr4</td>
<td>23891840</td>
<td>TAA/TA  A</td>
<td></td>
<td>PPARGC1A</td>
<td>Upstream gene variant</td>
</tr>
</tbody>
</table>

* NRF1—nuclear respiratory factor 1, MFN1—mediator of mitochondrial fusion, PPARGC1A—mitochondrial biogenesis, COQ9—important in mitochondrial ETC.

The homozygous SNPs and indels in COQ9, NRF1, MFN1 and PPARGC1A were present in 70% of the couples within the RM/RIF in both partners and not present in the fertile group, suggesting they might have had an effect on their fertility. NRF1 encodes a protein that homodimerizes and functions as a transcription factor which activates the expression of some key metabolic genes regulating cellular growth and nuclear genes required for respiration, heme biosynthesis, and mtDNA transcription and replication. MFN1 is a mediator of mitochondrial fusion. PPARGC1A is important in mitochondrial biogenesis with conditions of increased energy demands inducing its expression. To perform this task, PPARGC1A regulates the transcriptional activities of a large number of transcription factors, including, among others, and

5.3.5 mtDNA haplogroup specific comparison based on nuclear encoded mitochondrial genes associated SNPs in genomics

A comparison between individuals in their nuclear SNPs according to their mitochondrial haplotypes was performed. Individuals were categorised into several groups. Previously we identified that carriers for mitochondrial haplogroup T (especially male partners), had embryos with poor development even when the fertilisation rates were high (Chapter 4 section 4.4.2). Since our concern was on the male haplogroup T carriers, we focused our comparisons on the carriers of this haplogroup and compared them to the others in the study.

The only identified SNPs which were only present in the male T carriers compared with all the other groups was when the comparison was made with M mitochondrial haplogroup carriers. 80% of the M mitochondrial haplogroup carriers had heterozygous SNPs in TFB2M and PPARGC1A while all the male carriers for mtDNA haplogroup T were homozygous for SNPs on those gene (Figure 5–11). No significant differences in indels were present.

<table>
<thead>
<tr>
<th>Chr</th>
<th>Pos</th>
<th>SNP ID</th>
<th>Gene</th>
<th>Impact</th>
<th>VEP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chr1</td>
<td>4246711789</td>
<td>rs3124131</td>
<td>TFB2M</td>
<td>Modifier</td>
<td>Intronic variant</td>
</tr>
<tr>
<td>Chr4</td>
<td>23817409</td>
<td>rs2932968</td>
<td>PPARGC1A</td>
<td>Modifier</td>
<td>Intronic variant</td>
</tr>
<tr>
<td>chr4</td>
<td>23829862</td>
<td>Rs3774907</td>
<td>PPARGC1A</td>
<td>Modifier</td>
<td>Intronic variant</td>
</tr>
</tbody>
</table>

Figure 5–11: Nuclear SNPs comparisons between male carriers for mitochondrial haplogroup T with carriers for M mitochondrial haplogroups. T=5, M=5. The identified SNPs were on TFB2M, which is important in mtDNA transcription; and PPARGC1A, which is important in mitochondrial biogenesis.
5.4 Discussion

5.4.1 Mitochondrial-nuclear mismatch investigation in embryos

Natural human mtDNA sequence differences, generally denoted as mtDNA haplotypes, are often associated with regional migration and adaptation to climate and are measured by SNP changes (Wallace et al., 1999). Mitochondrial DNA haplotypes can modulate the pathological effects of mutated nuclear encoded genes (Strauss et al., 2013), and the same variation in nDNA can be beneficial or deleterious depending on its mtDNA background and environment (Ji et al., 2012, Burgstaller et al., 2015). Mitochondrial DNA is actively transcribed from the 2-cell stage to produce respiratory chain subunits, in which paternal nDNA encoded proteins are incorporated, even though mitochondrial biogenesis is absent at this stage (Piko and Taylor, 1987). Shortly after activation of the embryonic genome, the assessment of mitochondrial function at the 4-cell stage embryos have shown differences in the generation of a MMP; which drives energy production by enabling oxidative phosphorylation of pyruvate, and has been used to reflect variations in mitochondrial function in the embryo (Van Blerkom et al., 2002, Thouas et al., 2004, Binder et al., 2012).

In the embryo, the mtDNA is inherited with a haploid maternal genome. Energy production by the OXPHOS system in mitochondria depends on an extensive cross-talk between genes from the nucleus (inherited from mother and father) and genes from mitochondria (inherited from mother) (Johnson et al., 2001, Reinhardt et al., 2013). Half of the nuclear genes participating in the production of proteins involved in OXPHOS are from the father. Proteins encoded by this set of genes interact with the proteins encoded by the mitochondrial encoded set of proteins in the embryo.
Effective interaction between these subunits is essential to avoid the significant slow-
down in electron transfer leading to low ATP, high leak of ROS, loss of cytochrome
C and eventually apoptosis. This can be seen as functional selection against mito-
nuclear mismatch (Lane, 2011). If certain variations cause a mismatch between
mtDNA and the nDNA the implications of this nDNA-mtDNA combination are
unknown. They could either be beneficial or detrimental or have no effect at all. For
this reason, following the mitochondrial genome sequencing, we aimed to further
investigate the mitochondrial-nuclear mismatch effect on embryo development by
sequencing the nuclear genes important for mitochondrial function of the parents
and the embryos produced from those families.

5.4.2 Selection of candidate nuclear genes

Nuclear genes were selected based on the importance of the proteins they encode
which are directly interacting with the mitochondrial encoded proteins in the
mitochondrial ETC. Even though the whole mitochondrial genome was sequenced in
Chapter 4, six mitochondrial-encoded subunits in the ETC have been highlighted to
aid in the selection of nuclear genes.

The capacity for dynamic interactions between subunits was used as a basis for
the selection of candidate genes. This was determined based on the hydrophobicity
(HYD) and Ser/Thr content (STC) of Mitochondrial Membrane Proteins (MMPS).
Strong hydrophobic interactions between membrane proteins restrict conformational
freedom, as the energetic barrier in separating the surfaces of hydrophobic helices is
high. According to Kitazoe et al (Kitazoe and Tanaka, 2014) selection for
mitochondrial power across metazoans, from sessile sponges to vertebrates (with
high aerobic capacity), involved striking reduction in HYD and tightly correlated to
increases Threonine content (TC) of MMPs after a systemic analysis of metazoan mitochondrial genomes incorporating 1700 sequences. Lower HYD may increase the conformational freedom of respiratory proteins, while moderately polar Thr and Ser residues help to stabilize dynamic conformational changes via cooperative networks of hydrogen bonds within and between trans-membrane helices (Figure 5-12).

This relationship is mainly noticed in the 3-large proton pumping subunits of complex I; ND2, ND4 and ND5 and the two subunits involved in gating proton channels in complex IV (CO1 and CO3) since they appear to be significantly enriched in TC, with 3-6 folds increase in metazoan. Thus, it has been proposed that the HYD-TC correlation reflects dynamic inter-helical hydrogen-bonding between and within respiratory subunits so that the low HYD and high TC are linked to increases in conformational freedom and dynamic stability of MMPs. This relates to the energetic requirements and the longevity of organisms.

![Diagram explaining the Hydrophobicity and Threonine content effect on the energy release.](Note: Image of diagram showing the relationship between low HYD and high TC with increased conformational freedom and energy release, versus high HYD and low TC with decreased conformational freedom and energy release.)
5.4.2.1 Complex I genes

ND1, ND4, and ND6 seem to be fundamental to Complex I (CI) assembly, while ND3 and ND5 are important for its activity but not for assembly. ND2 is the core subunit in complex I. Mutations in ND2 alter CI assembly, with abnormal intermediate accumulation (Perales-Clemente et al., 2010). ND4 codes for NADH dehydrogenase 4 which is active in complex I. SNPs in *MT-ND2* and *MT-ND4* have been associated with recurrent miscarriage and pregnancy loss in Indian women (Vanniarajan et al., 2011). To date, genetic defects have been reported in all seven mtDNA encoded complex I subunits (Fassone and Rahman, 2012). *MT-ND5* codes for NADH dehydrogenase 5 which is active in complex I. The nDNA-encoded subunits in complex I which constitute excellent candidates for binding highlighted mtDNA-encoded subunits are: NDUFC2 and NDUFA1. These two subunits underwent accelerated amino acid replacement (Mishmar et al., 2006), suggesting their adjustment to the elevated mtDNA rate of change.

NDUFC2 and NDUFA1 are likely to interact with the ND5/ND4 and ND5/ND4/ND1 mtDNA-encoded subunits respectively. Using the yeast two-hybrid system, the predicted interactions were experimentally established suggesting the interaction of NDUFC2 with ND4, the interactions ND1 and ND4 with NDUFA1, and the absence of interaction between NDUFC2 with ND3 and NDUFA1 in humans, thus providing a proof of concept (Gershoni et al., 2010).
Figure 5–13: A proposed schematic model for the interaction network of NDUFC2, NDUFA1, and mtDNA-encoded complex I subunits.

The predicted interactions among subunits in complex I. The thickness of the arrows resembles the strength of the predicted interactions. Thick arrows represent the best predicted interactions (based on the top-scoring correlated pairs of residues). Question mark indicates a lack of clear predictions for the interaction. Reproduced from Journal of Molecular Biology, 404, Gershoni et al., Coevolution Predicts Direct Interactions between mtDNA-Encoded and nDNA-Encoded Subunits of Oxidative Phosphorylation Complex I, pages 158-171, Copyright (2010), with permission from Elsevier.

5.4.2.2 COMPLEX III and complex IV

Complex III accepts electron from coenzyme Q$_{10}$ and transfers them onto cytochrome c in the ETC. The process of electron transfer from coenzyme Q$_{10}$ to cytochrome c is known as the Q cycle, participating indirectly in ATP synthesis. Only one mitochondrial encoded subunit is present in complex III known as MTCYB, Defects in MTCYB have been associated with mitochondrial complex III deficiency. Bioinformatics analysis has shown that the UQCRB nuclear encoded subunit in complex III is involved in proton pump (http://www.uniprot.org/uniprot/P14927). This gene encodes a protein that plays a vital role in hypoxia induced angiogenesis through ROS mediated signalling, has been linked to redox linked proton pump. Expression of that gene is highly up regulated in GV oocytes (Cui and Kim, 2007).

In complex IV cytochrome c was selected. This gene encodes a small heme protein that functions as a central component of the ETC in mitochondria. The protein encoded by this gene accepts electrons from cytochrome b and transfers
them to the cytochrome oxidase complex and is also involved in the initiation of apoptosis.

5.4.3 Nuclear DNA variations associated with poor embryo development

Contrary to what was anticipated, following the analysis of the genetic variants within the selected genes using different computational tools, our results have shown that the variations that were associated with poor embryo development were from genes important in mitochondrial function but not necessarily in genes involved in direct protein-protein interaction processes within the electron transport chain. Also, they were associated with arrested embryos in specific families but not among the whole group of poorly developed embryos.

5.4.3.1 CoQ9

A single variant in the 5 UTR region upstream of the gene at position 57481670 of coenzyme Q9 (COQ9) (T/C; rs223861) was associated with genetic variation in preimplantation embryo development, with the T allele associated with embryonic developmental arrest in two families. The potential predicted effect of the SNPs in COQ9 showed that the identified SNP in the promotor binding site of COQ9 gene might change protein expression level when the associated allele is C. Other identified SNPs on COQ9 located in intronic regions, downstream or upstream of the gene might have had a modifying effect on gene expression. However, this could not be confirmed in this study due to the lack of RNA samples.

COQ9 is necessary for the synthesis of Ubiquinone (Coenzyme Q₁₀); a lipid binding protein found in most biological membranes and important in many redox processes including mitochondrial ETC. Coenzyme Q₁₀ is a co-factor in the function of the respiratory chain in transporting electrons from Complexes I and II to Complex
III. While the electrons are transferred to complexes III and IV, a proton gradient is created which ends in the generation of ATP by ATP synthase. Coenzyme Q₁₀ is a major cellular antioxidant, altering various signalling pathways, controlling cellular redox, influencing transcriptional activity of cells and important for the activity of succinate dehydrogenase (Crane, 2001, Quinzii et al., 2010, Ben-Meir et al., 2015). Point mutations in enzymes responsible for Coenzyme Q₁₀ synthesis, such as Pdss2 and Coq6, are characterised by phenotypes involving high energy-consuming tissues such as skeletal muscle, the central nervous system and the kidneys (Lopez et al., 2006, Peng et al., 2008, Heeringa et al., 2011). Coenzyme Q₁₀ has been implicated in protection from oxidative stress and may play a role in the aging process (Chong-Han, 2010). Accordingly, it has been used to treat mitochondrial diseases and as a dietary supplement to improve mitochondrial function. Coenzyme Q₁₀, alpha-lipoic acid and resveratrol have been used as endogenous vitamins known as “mitochondrial nutrients” and shown to directly or indirectly protect mitochondria from oxidative damage (Bentov et al., 2014). They were also reported to increase mitochondrial function, including energy production and have been shown to have several benefits in neurodegenerative and cardiovascular diseases (Al Ghafli et al., 2004, Hager et al., 2007). Changes in the sequence of COQ9 may affect ubiquinone synthesis disturbing mitochondrial ATP production, increasing oxidative stress and apoptosis affecting many processes including embryogenesis.

Several studies investigating the effect of Coenzyme Q₁₀ to improve fertility outcome has been conducted. In animals, it was demonstrated that Coenzyme Q₁₀ supplementation resulted in improving oocyte mitochondrial function in older mice resulting in a significant increase in oocyte mitochondrial ATP production, reduced
ROS levels in aged controls and increased the number of ovulated oocytes to levels comparable to those detected in the oocytes of younger mice (Burstein, 2009). Another study examined the use of Coenzyme Q₁₀ in the in vitro culture of bovine embryos and found an increase in ATP content in the group of embryos cultured with Coenzyme Q₁₀. These changes were associated with a greater rate of early embryo cleavage, blastocyst formation, percentage of expanding blastocysts, and a larger ICM (Marriage et al., 2004).

In humans, Coenzyme Q₁₀ effect on fertility was assessed among patients above 35 years of age undergoing IVF–ICSI. Patients were treated with Coenzyme Q₁₀ and placebo which lasted for two months before and until oocyte retrieval. A comparison of the rate of oocyte aneuploidy between the two groups was performed by polar body biopsy and aneuploidy screening by a-CGH. As a result, a lower rate of aneuploidy in the group treated with Coenzyme Q₁₀ was observed. However, the data was not significant due to the small sample size and the study was terminated for safety concerns regarding the effects of polar body biopsy on embryo quality and implantation (Bentov et al., 2014). High levels of Coenzyme Q₁₀ in follicular fluid is associated with higher pregnancy rates and optimal embryo morphokinetic parameters (Akarsu et al., 2017). Based on all this evidence it seems that the COQ9 homozygous SNPs present in the poorly developed embryos, could be the reason behind their poor development by the increase in ROS production.

5.4.4 How do those SNPs interact with the mitochondrial genome?

Interestingly, COQ9 homozygous SNPs shared between the poorly developing embryos were apparent in couples when the father had a mitochondrial T haplotype. Haplotype T is mainly determined by a combination of SNPs in complex I.
5.4.4.1 Mitochondrial T haplogroup and ROS

Emerging evidence suggests that different mtDNA haplogroups are associated with subtle differences in the generation of ROS and OXPHOS capacity (Esposito et al., 1999, Ruiz-Pesini et al., 2000, Kofler et al., 2009). Recent evidence indicates that mtDNA haplogroups have functional effects on sperm motility, longevity, certain types of cancer and affecting the risk of individuals developing specific late-onset neurodegenerative diseases. Haplogroup-specific differences, in free radical production during OXPHOS, may be one such mechanism (Moreno-Loshuertos et al., 2006, Kofler et al., 2009).

In chapter 4, we have demonstrated that haplogroup T individuals (specifically males), had poor blastocyst formation rate compared to the other individuals with different haplogroups. Since it is the male partner, and the embryo does not inherit the father’s mtDNA, we assume that this effect could be due the co-adapted nuclear genes passed from the father. It has been suggested that Individuals carrying haplogroup T might be more vulnerable to oxidative damage than carriers of other haplogroups (Kofler et al., 2009) which could be due to complex I mutations generating more ROS.

Our theory explaining the results obtained suggests that the combination of effects, affected Coenzyme Q₁₀ in mitochondrial T haplogroup background, both affect the transfer of electrons from complex I to complex III thus generating more ROS. This does not seem to have a major effect since many individuals carrying this haplogroup are alive while some may experience some health issues as has been reported. In our cohort, it appears that the nuclear background associated with the mitochondrial haplogroup T was detrimental to embryo development when this nuclear background was associated with different mitochondrial haplotypes in the
embryos. Haplogroup T is determined by complex I SNPs. Certain SNPs in complex I could lead to slowed electron flow to Coenzyme Q$_{10}$. This combined with low Coenzyme Q$_{10}$ would cause a dual problem with electron transfer to CIII. As a consequence, this would lead to additional electron leakage from complex I than usual potentially triggering apoptosis and eventually embryonic developmental arrest (Figure 5–14).
We propose that:
A.) Male T haplogroup carriers may have a reduced electron flow. T haplogroup SNPs (mainly encoding complex I proteins) acts as a 'modifier' for the nuclear genome via constant interactions mediating a 'baseline' of bioenergetics (e.g. energy production, ATP turnover) and nDNA expression levels for specific cellular pathways (e.g. apoptosis). This does not seem to have a major effect since many individuals carrying this haplogroup are presumably co-adapted.
B) Embryos from male T haplogroup carriers inherit the nuclear component from the mother and father while inheriting mitochondrial genome from the mother. If the embryo inherited a different mitochondrial haplogroup (for example; H), the nuclear component will face a different background. This might affect complex I activity interrupting the shuttling of electrons. This may result in fewer electrons passing to Coenzyme Q₁₀ while more are present in CI. More ROS will be generated from complex I while less will be transferred to complex III exacerbated by Coenzyme Q₁₀ deficiency.

Figure 5–14: Schematic diagram for mtDNA–nDNA mismatch theory.
5.5 Results summary

1. No association was found between SNPs on the selected nuclear genes important in mitochondrial function and the poorly developing embryos among all the families.

2. Analysis comparing the developed with arrested embryos in each family, showed homozygosity in SNPs in COQ9 and PPARGC1A. Those were shared between poorly developing embryos in two families (Family-3, Family-4 & Family-9). A SNP in COQ9 was of regulatory function.

3. Homozygosity at other SNPs affecting regulatory functions (TP63, NFKB2, CHRNA1) were also identified.

4. Homozygosity in most cases was not due to aneuploidy or segmental chromosomal losses in COQ9 and PPARGC1A.

5. T haplogroup was present from the father in all those families suggesting an indirect mitochondrial-nuclear mismatch effect.

6. Aneuploidy status analysis showed that chromosomally multiple aneuploidy >5 embryos had poor morphology regardless of the presence of the mitochondrial T haplogroup.

7. Comparison of the percentage of similarities and differences of embryonic genotypes to the maternal haplotype of the selected 53 nuclear encoded mitochondrial genes showed a marginal significant effect on embryo morphology when the indels percentages were different between the embryo and maternal haplotype (p = 0.049). This effect was independent of the chromosomal status.
5.6 Limitations and future work

In this study, the analysis included the filtering of SNPs which were not present in the parents to minimise sequencing errors. However, this approach excluded de novo SNPs that may have been present in the embryos. Even though some SNPs were identified in the arrested embryos compared with the embryos that developed to the blastocyst stage, without the associated expression analysis we cannot predict exactly why SNPs on these genes were identified. If the RNA was available from these embryos, we would test this hypothesis by comparing the expression levels of these nuclear genes associated with mitochondrial function and biogenesis between arrested embryos and blastocysts.

Another limitation, is that we only tested one generation. Analysis of multiple generations has been useful in the identification of co-adaption in other species. As technologies such as exome and whole genome sequencing are more commonly used in studies of infertility, it is likely that more mitochondrial variants associated with infertility will be identified and the pathogenesis more fully understood.

For future work, the risk factors and genes identified in this study—especially COQ9 and PPARGC1A—should be further investigated by sequencing more samples from a broader population group, and in particular, gene expression profiling of extracted RNA should be performed. If a lowering in the expression level of COQ9 is associated with embryonic development, mitochondria-targeted therapeutics such as MitoQ could be investigated further.
Chapter 6 General discussion

Infertility is genetically heterogeneous. Aneuploidy and maternal age are not the only mechanisms by which this complex phenotype occurs. Energy production by mitochondria has been suggested to be an important determinant of early embryonic development, affecting the regulation of cell cycle, spindle organisation and chromosomal segregation (Thundathil et al., 2005). The most important way to produce energy in the cell is via the OXPHOS in mitochondria. The OXPHOS complexes are encoded by both the nDNA and mtDNA. OXPHOS subunits have to maintain specific protein-protein interactions (Zhang and Broughton, 2013) which corresponds to the synchronisation between the proton pump, leakage of ROS and ATP production. A mismatch between the two genomes may be one of the causes of infertility (Lane, 2011) and the developmental arrest of the embryo. The central importance and genetic co-dependency of mitochondria suggests that for the cell to grow and divide, any significant deficiency in mitochondrial function might trigger an adaptive nuclear response (Yun and Finkel, 2014). Earlier studies showed that mitochondrial genetic deficiencies induce a coordinated and complex nuclear response that alters the expression of forty or more genes in Saccharomyces cerevisiae (Epstein et al., 2001, Yun and Finkel, 2014) and more than 1000 nuclear genes were up- or down-regulated in drosophila (Innocenti et al., 2011).

This could be by increasing the number of mitochondria to produce more ATP or undergo apoptosis. Since mitochondrial replication does not occur until after implantation, some groups (Monnot et al., 2013) proposed that embryos with low OXPHOS activity might undergo mtDNA replication in early embryogenesis to compensate for the energy deficiency. Mitochondrial DNA number and ATP levels in
the embryo could provide an indication of cellular stress level which might contribute to embryonic developmental arrest.

To our knowledge, the mito-nuclear mismatch effect in humans has not been tested in terms of human fertility. We wanted to investigate this effect on human fertility and if this is related to embryonic developmental arrest. A literature review was done to select nuclear genes that may be affected by a mismatch and/or have a role in mitochondrial function. We aimed to sequence the mitochondrial genome and selected nuclear genes of couples and their embryos to see if there are any groups or populations that struggle to conceive based on incompatibilities between their mtDNA and nDNA backgrounds affecting mitochondrial function. Sequences from couples attending the ART clinic for PGD with no history of fertility were compared with couples suffering RM/RIF for unknown reasons. Since the mitochondria are maternally inherited, it provides some information about the ethnic background from the mother. We also explored the effect of the ethnicity determined by the nuclear genome to determine if differences in ethnicities between individuals in a couple affect fertility.

The mitochondrial haplogroup of the embryos was inferred from the maternal haplotype on the basis of maternal transmission. Nuclear genes in embryos at various developmental stages were sequenced to test the mito-nuclear mismatch effect. Recognised determinants of embryo quality are morphology and aneuploidy status. Therefore, the embryonic aneuploidy status was considered in our mito-nuclear mismatch analysis.

Our study has shown a significant difference in mtDNA number between euploid and aneuploid arrested embryos at various stages from fertile couples following PGD treatment. The mtDNA number was significantly elevated in euploid compared with
aneuploid embryos arrested at various stages suggesting a possible link to mitochondrial function. Analysis of ATP level in blastocysts showed that ATP in aneuploid blastocysts was highly variable, showing a greater variance until it showed a massive decline in embryos with more than 10 aneuploidies.

Two retrospective studies measuring mtDNA copy number in human preimplantation embryonic samples have been performed as a measure for implantation potential. The MitoScore test suggests that fewer mtDNA copies are present in euploid cleavage stage embryos (tested on day 3) resulting in better implantation outcomes (Diez-Juan et al., 2015). Others, have introduced the Mito grade test, which used mtDNA copy number for the same purpose by measuring the mtDNA copies in trophectoderm samples from blastocysts and found that more mtDNA copies were present in aneuploid blastocysts (Fragouli et al., 2015, Fragouli et al., 2017, Ravichandran et al., 2017, Wells et al., 2017). They have also reported better implantation outcomes when mtDNA was lower in the transferred blastocysts. Our study confirms both findings indirectly. In our study embryos with high levels of mtDNA were euploid, however these embryos were arrested and unlikely to implant.

Based on these observations, it seems that higher levels of mtDNA are associated with poor implantation potential. The increased amount of mtDNA in euploid arrested embryos, aneuploid blastocysts and embryos with poor implantation potential (Diez-Juan et al., 2015, Bayram et al., 2017) may be an indication of a reduction in metabolic energy during oocyte maturation that is a consequence of a compensating mechanism to normalise ATP generation. Mitochondria in the oocytes of older hamsters and mice have been shown to generate higher levels of ROS and produce less ATP, indicating less energy support for dynamic processes, such as preimplantation development (Duran et al., 2011). In humans, if this scenario occurs,
an increase in mitochondrial number may be needed in the embryos of older women or poorly developing embryos, to maintain the required ATP levels. Thus, based on all these observations, mtDNA number is not fixed and changes in response to stress in embryos. It is also possible that these differences in mtDNA content could reflect different mtDNA content in the oocyte, perhaps also as a result of energy stress; but if so that could not reflect mito-nuclear incompatibilities, as pre-fertilisation the paternal chromosomes could not have contributed to the phenotype.

Mitochondrial number and replication are controlled by nuclear genes which show elevated expression in mice and bovine embryos at different stages (Aiken et al., 2008, Wai et al., 2010). There seems to be a balance between degradation and mtDNA synthesis. Therefore, the timing at which mitochondrial number changes in embryos is affected by developmental stage, expression of specific genes and possibly aneuploidy status. From this, we hypothesise, that the elevated levels of mtDNA in some embryos suggests that mtDNA biogenesis might possibly have occurred. Further investigation of the possible mechanisms will require gene expression analysis, determination of aneuploidy and levels of mtDNA at different stages of embryonic development. Availability of appropriate human samples may prove difficult due to the preponderance of blastocyst biopsy in clinical practice.

For future mtDNA quantification tests, a non-human template (NHT) could be used for normalisation in real time PCR analysis. This NHT should be normalised to several nuclear genes as well as a mitochondrial gene to compare the ratios of NHT: mtDNA and NHT: nuclear genes measured in embryos. This may give a better indication of the overall amount of mtDNA.

The major part of this study tested the mito-nuclear mismatch hypothesis (Lane, 2011, Burton et al., 2013) and was based on the analysis of mitochondrial and
nuclear genome sequences from fertile couples as well as couples suffering from idiopathic RM/RIF. Recently arguments have been brought forward regarding mito-nuclear mismatch in terms of fitness, reproductive health and mitochondrial replacement for the avoidance of mitochondrial disease transmission as a major safety issue however no one has investigated its effect on human fertility. Mitochondrial dysfunction and mtDNA mutations affecting ATP production may cause developmental arrest before the pregnancy is clinically recognised (Van Blerkom et al., 1998; Van Blerkom, 2004). Embryos have the same mitochondrial component (haplogroup) as their mothers but have a different nuclear genome as this is inherited from both mother and father. The nuclear genome includes many genes important in mitochondrial function. Based on the assumption that the mtDNA haplotype could be used as a surrogate marker for the nuclear background for parental genes important for mitochondrial function, we identified differences in mtDNA haplotypes of male and female partners and compared these differences between fertile couples and couples with RM/RIF.

Contrary to what we expected, couples within the fertile group were more likely to be distantly related by mitochondrial haplotype compared with the RM/RIF groups showing more differences in mitochondrial haplotypes between individuals, compared with the RM/RIF groups. Despite our small number of samples there was a trend suggesting that incompatibilities in mitochondrial haplotypes can be beneficial. This is consistent with a previous study showing benefits in unmatched nDNA/mtDNA haplotypes in mice having a longer mean lifespan (Latorre-Pellicer et al., 2016). Our study also agrees with the mitohormesis theory, which argues that “Any of a number of endogenous or exogenous stress can perturb mitochondrial function, there perturbations are relayed to the cytosol through, at present, poorly
understood mechanisms that may involve mitochondrial ROS as well as other mediators. These cytoplasmic signalling pathways and subsequent nuclear transcriptional changes induce various long lasting cyto-protective pathways. This augmented stress resistance allows for protection from a wide array of subsequent stresses.” (Yun and Finkel, 2014). It could be that cellular ATP is influenced by certain mitochondrial haplotypes, whose variants mediate not only cell growth through the production of energy, but also cell signalling for the major molecular pathways, including those regulating chromosome segregation (Kenney et al., 2013). Studies in macaques and mouse models support the view that mito-nuclear mismatch can have a limited effect, if any at all, on mitochondrial transfer techniques in human populations (Chinnery et al., 2014). In terms of evolution, adult trans-mitochondrial non-human primates have been produced carrying wild-type mtDNA haplotypes from different rhesus macaque subpopulations and in which longitudinal studies demonstrate normal growth and development to adults with unaffected ATP levels and membrane potential (Tachibana et al., 2013).

On the other hand, the mito-nuclear mismatch effect can cause infertility influence lifespan in other species for example; in Drosophila (Camus et al., 2012) and may also have dramatic effects on fitness and fertility in other species such as copepods (Barreto et al., 2015). Depending upon the model and phenotypes examined different haplogroup combinations could have different effects on fertility outcome. We have demonstrated that mtDNA haplogroup T carriers (specifically in males), are significantly associated with an increased chance of embryo developmental arrest in Middle European Caucasian populations compared with the other individuals with different mtDNA haplogroups. Since the T haplotype was present in the male partner, and therefore this mitochondrial haplotype was not transmitted to the
embryo, we postulate that this effect could be due the co-adapted nuclear genes passed from the father. Homozygous SNPs in nuclear genes that were associated with embryonic developmental arrest were identified in some families where the father was a carrier for mitochondrial haplogroup T. These homozygous SNPs were identified in *COQ9* and *PPARGC1A*. Haplogroup T is mainly determined by a combination of SNPs in complex I in the ETC. Carriers for mitochondrial haplogroup T might be more vulnerable to oxidative damage compared with carriers of other haplogroups (Kofler et al., 2009), which could be due to complex I mutations lowering respiratory flux and generating more ROS.

*COQ9* is necessary for the synthesis of COQ₁₀ which is important in transporting electrons from complex I and complex II to complex III. Explaining possible reasons for our results is that affected COQ₁₀ results in a reduced electron flow in mitochondrial T haplogroup (complex I SNPs). Lower flux would of course lower ATP availability but could also affect flux through the Krebs cycle (Le et al., 2012), with possible effects on epigenetics, as well as potentially generating more ROS. This does not seem to have a major effect since many individuals carrying this haplogroup are alive while some may experience some health issues as has been reported in some diseases. However, in the embryo this leads to a slowed electron flow from complex I to COQ₁₀, presumably causing a build-up in electrons that are not passed to complex III. The additional electron leakage from complex I than usual could potentially trigger apoptosis and eventually embryonic developmental arrest.

*PPARGC1A* is a major stress protein and is essential for mitochondrial biogenesis. It has been identified along with some transcription factors including NRF1, NRF2 and YY1, as candidate regulators of proteins related to the OXPHOS system (van Waveren and Moraes, 2008, Leigh-Brown et al., 2010). It controls many
processes; however, without the associated expression analysis we cannot predict exactly why SNPs on this gene were identified in the arrested embryos compared with the developed. However we can hypothesize that in arrested embryos, whether euploid or aneuploid, transcripts are degraded (reviewed by (Menezo et al., 2010). Embryos may have arrested due to changes in gene expression. It may be that PPARGC1A is not activated due to a SNP inherited from a T haplogroup father leading to a mito-nuclear mismatch signalling stress. Once this stress reaches a certain level as in the case when the nuclear component of T haplogroup male carriers is passed to the embryo with the H mitochondrial background from the mother, it signals the need for mitochondrial biogenesis. Therefore, it seems that certain incompatible combinations of nuclear and mitochondrial components signal stress. However, no correlation was found between mtDNA copy number and mitochondrial haplogroup T in those arrested embryos. This could be because the sample size was small and did not reach statistical significance; or it could be that the SNPs in PPARGC1A interfered with mitochondrial biogenesis; something similar might also have happened with the SNPs in COQ9. If a larger combination of mtDNA haplotypes were available, haplogroups which are more fertile compared to others may have been identified. Also, if two partners for a mitochondrial haplogroup T were present, we would be able to confirm that incompatibility is the reason behind the observations seen in our study.

In the future, enlarging this study to include couples with different mitochondrial haplotypes from a range of ethnicities such as Asia and population groups where consanguinity is common such as Saudi Arabia may identify rare nuclear variants associated with mitochondrial dysfunction and infertility. In the present study analysis, SNPs not present in the parents were excluded to minimise sequencing
errors however this approach excluded *de novo* SNPs that may have been present in the embryos. Another limitation, is that we only tested one generation. Analysis of multiple generations has been useful in the identification of co-adaption in other species (Camus et al., 2012). As technologies such as exome and whole genome sequencing are more commonly used in studies of infertility, it is likely that more mitochondrial variants associated with infertility will be identified and the pathogenesis more fully understood.

Our study showed that limited information regarding parental ethnicity and genotyping of nuclear genes associated with mitochondrial function can be determined from karyomapping data. Karyomapping can also give information regarding the aneuploidy status of the embryo. Collating this information from routine clinical PGD cycles may be a useful source of research data for evidence of mitonuclear mismatch.

For future research, the risk factors and genes identified here-especially COQ9 and PPARGC1A- should be further investigated by sequencing more samples from a broader population group, and in particular, gene expression profiling of extracted RNA should be performed. If COQ9 expression deficiency is confirmed to affect embryonic development, mitochondria-targeted therapeutics which have been used for treatment in processes involving oxidative stress and metabolic failure could be investigated. The recently developed, Mitochondria-targeted antioxidant MitoQ (mitoquinone mesylate: dimethoxy-2-methyl-3,6-dioxo-1,4-cyclohexadienyl) could be a good example of a specific antioxidant that only accumulates within mitochondria compared with other antioxidants which are nonspecific and has shown to interact with ROS when used to treat some pathologies (Smith et al., 2003, James et al., 2005, Rodriguez-Cuenca et al., 2010). It has recently been reported that MitoQ
reduced the effect on the progression of AD-like pathologies prevented in the brains of young female mice in vivo (McManus et al., 2011). It also has antioxidant efficiency in a number of tissues in vivo (Gane et al., 2010; reviewed by (Smith and Murphy, 2010)). However, it has also been used as an effective pharmaceutical and has been used in clinical trial to treat parkinson’s and hepatitis C diseases. In human clinical trials, treatment of parkinson’s by MitoQ had no effect (Snow et al., 2010). On the other hand, Hepatitis C patients who were treated with MitoQ showed a decrease in liver damage, but it was not sufficient for full treatment (Gane et al., 2010). Therefore, in the near future, if the clinical relevance of treatment with MitoQ tablets was proven for many more disease, it could be used for treatment of infertility.

To test embryo viability, the use of the OROBOROS Oxygraph-2k is suggested. This is a unique instrument used for high-resolution respirometry (HRR) analysis. Respirometry reflects the function of mitochondria as structurally intact organelles. It provides a dynamic measurement of metabolic flux (rates), in contrast to the static determination (states) of molecular components, such as metabolite and enzyme levels, redox states and membrane potential, concentrations of signalling molecules, or RNA and DNA levels. Such measurements have to be performed on the fresh samples with delicate handling procedures to preserve the structure and function of mitochondria. Measurement of respiratory flux in different metabolic states is essential for evaluating the effect of oxidative phosphorylation changes in membrane permeability, metabolite levels, or activity of individual enzymes. Small imbalances in metabolic flux can result in large cumulative changes of the metabolic state. However, this should be done on live embryos immediately after tubing therefore, it can only be used for research purposes to better understand mitochondrial function.
at the early stages of human embryo development. Frozen-thawed embryos that are scheduled for discarding could be used for this type of analysis immediately after thawing. With regards to ATP, ATP could be a reflection of ATP produced by glycolysis or the ETC. Therefore, the ADP/ATP ratios would give a better clearer understanding.

6.1 Conclusion

In conclusion, our data support the findings by other groups suggesting mtDNA content could be an indicator of energetic stress in the embryo and not a reflection of energetic competence. Therefore, it could be used to predict their embryonic implantation potential.

The results from this study, point to the relevance of mitochondrial haplotypes of both parents in diagnosis. Couples within the fertile group were more likely to be distantly related in their mtDNA haplotypes compared to the infertile group. Some combinations of nuclear and mitochondrial genomes may lead to a mismatch affecting fertility. In this study, the combination of a nuclear component from mitochondrial haplogroup T male carriers with female partners who had mitochondrial haplotype H was suggestive of mitonuclear mismatch. These findings may allow investigators in the future to define a mtDNA (haplogroup-specific) tests to assess susceptibility to infertility for couples. Studying the risk factors for mitochondrial haplogroup T can help identify a group of individuals where other molecular mechanisms that affect embryo development are identified.
Chapter 7

Appendix

7.1 Appendix A

7.1.1 Validation of GenetiSure kit for aneuploidy screening

Before using the GenetiSure arrays for aneuploidy screening, the arrays were validated by comparing the results obtained with previously analysed samples. Two validation steps were performed. The first validation step was by comparing GenetiSure to BlueGnome arrays results of embryos from families (n=2) who underwent PGD for reciprocal translocations. The other validation step was comparing the results obtained by FISH and Karyomapping from embryos from families who underwent PGS for aneuploidy screening.

7.1.1.1 Array CGH BlueGnome versus GenetiSure

The results of some embryos from two families who underwent PGD for reciprocal translocations (Family-24, Family-25) were analysed and compared without knowing the original reports of the cases. Only six embryos that have passed the quality control (QC) metrics in the GenetiSure array were used in this validation step as embryos with evaluate results have shown extra abnormalities that could affect the validity of the results. Moreover, because the samples in GenetiSure are compared to two references (male and female), the only counted abnormalities are the ones present with both references.

As shown in Table 7-1, most of the aneuploidies detected were in the chromosomes involved in the translocations. When the calculation of the similarity percentage between BlueGnome and Agilent arrays was performed, all the recorded abnormalities were counted in the measurement showing 62.5% similarity. However,
if the calculation includes only the chromosomes with translocations, the similarity has increased to 75%.

**Table 7-1: A comparison between BlueGnome and Genetisure array results.**

<table>
<thead>
<tr>
<th>Embryo No</th>
<th>BlueGnome (24 sureplex +)</th>
<th>Genetisure vs female</th>
<th>Genetisure vs male</th>
<th>XY</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Family-24-E4</strong></td>
<td>Loss of p arm in Chr.18</td>
<td>Gain of p arm in Chr.9 Loss of p arm in Chr.18</td>
<td>Gain of p arm in Chr.9 Loss of p arm in Chr.18</td>
<td>XX</td>
</tr>
<tr>
<td><strong>Family-24-E5</strong></td>
<td>Gain of q arm in Chr.9</td>
<td>Gain of q arm in Chr.9 Loss of q arm in Chr.18</td>
<td>Gain of q arm in Chr.9 Loss of q arm in Chr.18</td>
<td>XX</td>
</tr>
<tr>
<td><strong>Family-24-E8</strong></td>
<td>Loss of p arm in Chr. 9</td>
<td>Loss of p arm in Chr.9 Gain of p arm in Chr.18</td>
<td>Loss of p arm in Chr.9 Gain of p arm in Chr.18</td>
<td>XXY</td>
</tr>
<tr>
<td><strong>Family-24-E12</strong></td>
<td>Loss of p arm in Chr. 9</td>
<td>Loss of p arm in Chr.9 Gain of p arm in Chr.18</td>
<td>Loss of p arm in Chr.9 Gain of p arm in Chr.18</td>
<td>XY</td>
</tr>
<tr>
<td><strong>Family-25-E6</strong></td>
<td>Gain of Chr.16 Loss of Chr.22</td>
<td>Loss in P arm of Chr. 1 Gain in q arm of Chr. 5 Gain in q arm of Chr. 7 Loss in arm of Chr. 12 Gain of Chr.16 Loss in p arm of Chr.19 Loss of Chr.22</td>
<td>Loss in P arm of Chr. 1 Gain in q arm of Chr. 5 Loss in arm of Chr. 12 Gain of Chr.16 Loss of Chr.22</td>
<td>XX</td>
</tr>
<tr>
<td><strong>Family-25-E7</strong></td>
<td>Gain of Chr.16</td>
<td>Gain of Chr.16</td>
<td>Complete Gain of Chr.16 Loss of p arm in Chr.19</td>
<td>XX</td>
</tr>
</tbody>
</table>

The differences in calls between the two arrays could be explained by the fact that the BlueGnome arrays are BAC arrays so the probes are much less densely distributed. On the other hand, GenetSure arrays have many more oligonucleotide probes which are more evenly distributed. The oligonucleotide arrays can be more precise and detect smaller aberrations and translocations, however the appearance of the arrays is more sensitive to noise. Therefore, for this reason only losses and gains that were observed in both female and male reference and were 50 Mbp or above were scored.
7.1.1.2 Array-CGH (GenetiSure) versus FISH & Karyomapping

The same analysis standards used in the validation part for GenetiSure array were used to analyse the embryos from two PGS families (Family-20 and Family-26). Two array runs were performed; the first one included all the available embryos from both cases (Table 7-2).

The contradictory results of FISH to GenetiSure array and karyomapping, that demonstrate clear losses (monosomies) of chromosome 22, in both embryos E1 and E2 in Family-20, raises many queries about the validity of FISH and GenetiSure tests, the source and the mechanism of the chromosomal errors occurred and the different conditions and possibilities that could generate such outcomes. Because chromosome 12 was not included in the FISH analysis, which is one of the main drawbacks of this approach, the chromosome loss in Family-20-E5 that was identified by the arrays was missed by FISH. One possible explanation of the contradictory outcomes could be the less reliable results of FISH since it was performed on single cells that obtained at the cleavage stage, which did not represent the whole chromosomal status of the embryo.
Table 7-2: Embryo results using different techniques (FISH, array-CGH and karyomapping)

<table>
<thead>
<tr>
<th>Embryo</th>
<th>FISH</th>
<th>Genetisure (whole chromosome)</th>
<th>Genetisure (segmental)</th>
<th>Karyomapping (whole chromosome)</th>
<th>Karyomapping (segmental)</th>
<th>Gender</th>
<th>Embryo fate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Family-20-E1</strong></td>
<td>Diploid for chromosomes 13, 18, 21, 15 and 16. Trisomy 22.</td>
<td>Loss of Chr. 22</td>
<td>Loss in p arm of Chr. 19</td>
<td>Loss of Chr. 22</td>
<td>Loss in Chr. 19</td>
<td>XY</td>
<td>Whole</td>
</tr>
<tr>
<td><strong>Family-20-E2</strong></td>
<td>Monosomy 18, 21. nullisomy 15. Diploid for chromosomes 13, 16 and 22.</td>
<td>Mosaic 19 (loss)</td>
<td>No segmental abnormalities</td>
<td>No whole chromosome abnormalities</td>
<td>Loss in Chr. 19 and 22</td>
<td>XY</td>
<td>Whole</td>
</tr>
<tr>
<td><strong>Family-20-E4</strong></td>
<td>Mosaic. A. Trisomy 13, 15. Monosomy 22. Diploid for chromosomes 16, 18 &amp; 21. B. Monosomy 21, Trisomy 22. Diploid for chromosomes 13, 18, 15 and 16.</td>
<td>Loss of Chr. 22</td>
<td>Loss in p arm of Chr. 19</td>
<td>Loss of Chr. 22</td>
<td>Loss in Chr. 17 and 19 gain in Chr. X</td>
<td>XX</td>
<td>Whole</td>
</tr>
<tr>
<td><strong>Family-20-E5</strong></td>
<td>Mosaic. A. Monosomy 13, 18. Monosomy 15. Diploid for chromosomes 16, 21 and 22. B. Monosomy 18, nullisomy 13, 15, 22. Diploid for chromosomes 16 and 21.</td>
<td>Mosaic 12, 18, 19 (loss)</td>
<td>No segmental abnormalities</td>
<td>Loss of Chr.12</td>
<td>Loss in Chr. 17, 18, 19 and 22</td>
<td>XX</td>
<td>Whole</td>
</tr>
<tr>
<td><strong>Family-20-E6</strong></td>
<td>Monosomy 16, trisomy 22 (different source)</td>
<td>Mosaic 19 (loss)</td>
<td>No whole chromosome abnormalities</td>
<td>No whole chromosome abnormalities</td>
<td>Loss in Chr. 17, 19 and 22 gain in Chr. X</td>
<td>XX</td>
<td>Whole</td>
</tr>
<tr>
<td><strong>Family-26-E2</strong></td>
<td>Trisomy 21, nullisomy 13, 16 monosomy 15, 18</td>
<td>Gain of Ch. 16</td>
<td>No segmental abnormalities</td>
<td>Gain of Chr.16</td>
<td>Losses in Chr. 17, 19, 22</td>
<td>XY</td>
<td>Whole</td>
</tr>
<tr>
<td><strong>Family-26-E9</strong></td>
<td>Nullisomy 13, 15, 16</td>
<td>Loss of Chr’. 17 (p13.3-123.2) (contains the centromere), mosaic 10(loss) 14 (gain)</td>
<td>No segmental abnormalities</td>
<td>Loss of Chr. 17 (includes the centromere)</td>
<td>Losses in Chr. 10, 19, 22</td>
<td>XX</td>
<td>1 cell</td>
</tr>
</tbody>
</table>
7.2 Appendix B

7.2.1 Calculation used to dilute the standards (oligonucleotides) used in this study

According to literature, the amount of mtDNA molecules in an oocyte is approximately between \(10^5\) and \(10^6\) molecules. Since mtDNA replication does not occur until after implantation, each blastomere derived from a day 3 dividing embryo will contain:

\[
\frac{10^6}{8} = 125000 = 1.25 \times 10^5 \text{ mtDNA molecules}
\]

Each oligonucleotide was diluted to a 50\(\mu\)M concentration and the following calculation was performed in order to establish a dilution that would contain a similar number of oligonucleotide template as mtDNA molecules in a single blastomere based on what have been reported in literature.

\[
50 \, \mu\text{M} = 50\mu\text{mol/L} = 50 \times 10^{-6} \times 10^{-6} \text{ moles/\mu l} = 50 \times 10^{-12} \text{ moles/\mu l.}
\]

To convert that to molecules:

\[
50 \times 10^{-12} \times 6.02 \times 10^{23} = 3.01 \times 10^{13} \text{ molecules/\mu l}
\]

So, there are \(3.01\times10^{13}\) mtDNA nucleotide templates in every 1\(\mu\)l of a 50\(\mu\)M stock (normal or mutant) oligonucleotide. Thus, to establish the factor by which the stock oligonucleotide would need to be diluted to represent the number of mtDNA molecules in a single cell from a cleavage stage embryo the following calculation was performed:

\[
3.01 \times 10^{13} / 1.25 \times 10^5 = 2.41 \times 10^8 \text{ dilution factor}
\]

This indicates that \(2.41 \times 10^8\mu\)l of stock would contain a similar amount of mtDNA molecules in a blastomere derived from a day 3 embryo. Therefore, a 1/10 serial dilution was performed eight times for both normal and mutant oligonucleotides:
Dilution 1: $3.01 \times 10^{12}$
Dilution 2: $3.01 \times 10^{11}$
Dilution 3: $3.01 \times 10^{10}$
Dilution 4: $3.01 \times 10^9$
Dilution 5: $3.01 \times 10^8$
Dilution 6: $3.01 \times 10^7$
Dilution 7: $3.01 \times 10^6$
Dilution 8: $3.01 \times 10^5$

Every 1µl from dilution 8 consists of $3.01 \times 10^5$ mtDNA molecules.

**7.2.2 Calculation used to dilute the ATP standards**

- According to the kit the ATP standard concentration is 1 mg/10ml = 0.1 mg/1ml=100ng/µL (before processing).

*After processing:*

**1:4 dilution** is performed since 50µL of sample +50µL of dH₂O +100µL of ATP releasing agent (A mixture).

*Another 1:2 dilution* is performed since 100µL of sample (A mixture) is mixed with 100µL of ATP assay mix containing luciferase. So, the samples and the standard are diluted while processing 1:8.

A 1:2.5 serial dilution was performed for the standard:

- Dilution 1= $100/ (2.5) \times 8 =5 \text{ ng/µL}$
- Dilution 2: $100/ (2.5)^2 \times 8 =2 \text{ ng/µL}$
- Dilution 3: $100/ (2.5)^3 \times 8 =0.8 \text{ ng/µL}$
- Dilution 4: $100/ (2.5)^4 \times 8 =0.32 \text{ ng/µL}$
- Dilution 5: $100/ (2.5)^5 \times 8=0.128 \text{ ng/µL}$
- Dilution 6: $100/ (2.5)^6 \times 8=0.0512 \text{ ng/µL}$
### Table 7-3: Embryonic aneuploidy status, morphology & mitochondrial haplogroups.

<table>
<thead>
<tr>
<th>Family ID</th>
<th>Maternal age</th>
<th>Mitochondrial haplotype</th>
<th>ID</th>
<th>Morphology</th>
<th>Aneuploidy</th>
<th>Whole chromosome or &gt;50Mbp</th>
<th>Whole chromosome aneuploides</th>
<th>segmental losses/gains</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>losses</td>
<td>gains</td>
<td>losses</td>
</tr>
<tr>
<td>1</td>
<td>42</td>
<td>T2a1b1a E1</td>
<td>Arrested at Cleavage</td>
<td>Multiple Aneuploid &gt;5</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>References</td>
<td>E3</td>
<td>Blastocyst</td>
<td>Aneuploid ≤ 5 (-chr12, -chr16, -chr 22)</td>
<td>12,16</td>
<td>0</td>
<td>22q</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Blastosct</td>
<td>Euploid</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Blastosct</td>
<td>Multiple Aneuploid &gt;5</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Blastosct</td>
<td>Multiple Aneuploid &gt;5</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Blastosct</td>
<td>Multiple Aneuploid &gt;5</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Blastosct</td>
<td>Multiple Aneuploid &gt;5</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>2</td>
<td>36</td>
<td>H1e4a E4</td>
<td>Arrested at Cleavage</td>
<td>Aneuploid ≤ 5 (+chr9, +chr17, +chr 10q)</td>
<td>0</td>
<td>9,17</td>
<td>0</td>
<td>10q, 16q, 19p, 19q, 20q</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Blastocyst</td>
<td>Multiple Aneuploid &gt;5</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>E4</td>
<td>Multiple Aneuploid &gt;5</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>E7</td>
<td>Arrested at Cleavage</td>
<td>Multiple Aneuploid &gt;5</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>E8</td>
<td>Arrested at Cleavage</td>
<td>Multiple Aneuploid &gt;5</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>E9</td>
<td>Arrested at Cleavage</td>
<td>Multiple Aneuploid &gt;5</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>E10</td>
<td>Arrested at Cleavage</td>
<td>Multiple Aneuploid &gt;5</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>E11</td>
<td>Arrested at Cleavage</td>
<td>Multiple Aneuploid &gt;5</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>E12</td>
<td>Arrested at Cleavage</td>
<td>Multiple Aneuploid &gt;5</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>E13</td>
<td>Arrested at Cleavage</td>
<td>Multiple Aneuploid &gt;5</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>E14</td>
<td>Arrested at Cleavage</td>
<td>Multiple Aneuploid &gt;5</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>E15</td>
<td>Arrested at Cleavage</td>
<td>Multiple Aneuploid &gt;5</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

A total of 11 fertile families were collected. The first column indicates the Family ID. The second column shows the maternal age. The third column shows the mitochondrial haplogroup which was inferred from the mother in the case of the embryonic samples since mtDNA is maternally inherited. Column 5 shows the embryo developmental stage at day 6-7 post fertilisation. Aneuploidy status is represented in the last column. E: refers to the embryo followed by its number. Embryo morphology: the morphology taken day 6-7 post fertilisation. Euploid: refers to normal chromosomal number, Aneuploid: refers to single monosomy or single trisomy. Aneuploid ≤ 5 refers to more than one aneuploidy, Multiple Aneuploid >5: refers to more than 5 aneuploidies. * : refers to multiple chromosomal aneuploidy which were difficult to score. N/A: not available.
<table>
<thead>
<tr>
<th>Family ID</th>
<th>Maternal age</th>
<th>Mitochondrial haplogroup ID</th>
<th>Morphology</th>
<th>Aneuploidy status</th>
<th>Whole chromosome or &gt;50Mbp</th>
<th>Segmental losses/gains</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 28</td>
<td></td>
<td>H4a1a1a</td>
<td>Arrested at Cleavage</td>
<td>Euploid</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E2</td>
<td>Arrested at Cleavage</td>
<td>Euploid</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E3</td>
<td>Arrested at Cleavage</td>
<td>Multiple Aneuploid &gt;5</td>
<td>6,7</td>
<td>2p,5p,5q,1q,7q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E4</td>
<td>Arrested at Cleavage</td>
<td>Multiple Aneuploid &gt;5</td>
<td>1,7,18</td>
<td>5q,6q,21q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E5</td>
<td>Blastocyst</td>
<td>Aneuploid (+chr 19)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E6</td>
<td>Arrested at Cleavage</td>
<td>Aneuploid ≤ 5 (+chr 10, +chr 19)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7 37</td>
<td></td>
<td>H4a1a3</td>
<td>Arrested at Cleavage</td>
<td>Aneuploid ≤ 5 (+chr 19, - chr 2p, - chr 3q)</td>
<td>0</td>
<td>1p,2p,3q,6p,17q,22q,20q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E5</td>
<td>Arrested at Cleavage</td>
<td>Aneuploid ≤ 5 (+chr 19, +chr 17, - chr 7q)</td>
<td>13</td>
<td>1q,2p,3p,6p,20q,2q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E3</td>
<td>Arrested at Cleavage</td>
<td>Multiple Aneuploid &gt;5</td>
<td>1,3,5,7,8,10,14,16</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E9</td>
<td>Arrested at Cleavage</td>
<td>Euploid</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E1</td>
<td>Blastocyst</td>
<td>Euploid</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8 33</td>
<td></td>
<td>F3a1</td>
<td>Arrested at Cleavage</td>
<td>Multiple Aneuploid &gt;5</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E10</td>
<td>Blastocyst</td>
<td>Aneuploid (-chr 13)</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E15</td>
<td>Arrested at Cleavage</td>
<td>Euploid</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E2</td>
<td>Blastocyst</td>
<td>Euploid</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E3</td>
<td>Arrested at Cleavage</td>
<td>Aneuploid ≤ 5 (+chr 14, +chr 15, +chr 17, -chr 4q)</td>
<td>0</td>
<td>1q,4q,9q,9p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E13</td>
<td>Arrested at Cleavage</td>
<td>Euploid</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E16</td>
<td>Blastocyst</td>
<td>Multiple Aneuploid &gt;5</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E19</td>
<td>Arrested at Cleavage</td>
<td>Aneuploid ≤ 5 (+chr 6, +chr 8, +chr 10, +chr 16, - chr 2p)</td>
<td>6</td>
<td>1q,2p,9q,5q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E17</td>
<td>Blastocyst</td>
<td>Aneuploid ≤ 5 (+chr 14, +chr 17)</td>
<td>0</td>
<td>1p,8p,16p,20q,22q</td>
</tr>
</tbody>
</table>

A total of 11 fertile families were collected. The first column indicates the Family ID. The second column shows the maternal age. The third column shows the mitochondrial haplogroup which was inferred from the mother in the case of the embryonic samples since mtDNA is maternally inherited. Column 5 shows the embryo developmental stage at day 6-7 post fertilisation. Aneuploidy status is represented in the last column. E: refers to the embryo followed by its number. Embryo morphology: the morphology taken day 6-7 post fertilisation. Euploid: refers to normal chromosomal number, Aneuploid: refers to single monosomy or single trisomy. Aneuploid ≤ 5 refers to more than one aneuploidy, Multiple Aneuploid >5: refers to more than 5 aneuploidies. *: refers to multiple chromosomal aneuploidies which were difficult to score. N/A: not available.
<table>
<thead>
<tr>
<th>Family ID</th>
<th>Maternal age</th>
<th>Mitochondrial haplotype ID</th>
<th>Morphology</th>
<th>Aneuploidy</th>
<th>Whole chromosome or &gt;50Mb</th>
<th>Whole aneuploidies</th>
<th>chromosome</th>
<th>segmental losses/gains</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>29</td>
<td>H26a1</td>
<td>Arrested at Cleavage</td>
<td>Aneuploid ≤ 5 (+chr6, chr17, - chr 1q, -16q)</td>
<td>0</td>
<td>6,17</td>
<td>1q, 16q</td>
<td>5q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H26a1</td>
<td>Blastocyst</td>
<td>Multiple Aneuploid &gt;5</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H26a1</td>
<td>Arrested at Cleavage</td>
<td>Euploid</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H26a1</td>
<td>Blastocyst</td>
<td>Euploid</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H26a1</td>
<td>Arrested at Cleavage</td>
<td>Multiple Aneuploid &gt;5</td>
<td>4, 5, 16, 18, 19, 20</td>
<td>11, 21</td>
<td>7q</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>31</td>
<td>H</td>
<td>Arrested at Cleavage</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H</td>
<td>Blastocyst</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H</td>
<td>Arrested at Cleavage</td>
<td>Aneuploid (+chr 17)</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>30</td>
<td>U1a1a1</td>
<td>Arrested at Cleavage</td>
<td>Multiple Aneuploid &gt;5 (+2, 10, 20, + chr 1p, + chr 3p, + chr 14q, -chr 5q)</td>
<td>0</td>
<td>2, 10, 20</td>
<td>5p, 5q, 9p, 9q</td>
<td>1p, 3p, 6p, 14q, 15q, 22q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>U1a1a1</td>
<td>Arrested at Cleavage</td>
<td>Aneuploid ≤ 5 (-chr5, -chr10, -chr20)</td>
<td>5, 10, 20</td>
<td>0</td>
<td>21q</td>
<td>19p, 19q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>U1a1a1</td>
<td>Arrested at Cleavage</td>
<td>Aneuploid ≤ 5 (- chr5, + chr 1p, + chr 3p, - chr 7q)</td>
<td>5</td>
<td>0</td>
<td>7q, 9p</td>
<td>1p, 3p, 4q, 17q</td>
</tr>
</tbody>
</table>

A total of 11 fertile families were collected. The first column indicates the Family ID. The second column shows the maternal age. The third column shows the mitochondrial haplogroup which was inferred from the mother in the case of the embryonic samples since mtDNA is maternally inherited. Column 5 shows the embryo developmental stage at day 6-7 post fertilisation. Aneuploidy status is represented in the last column. E: refers to the embryo followed by its number. Embryo morphology: the morphology taken day 6-7 post fertilisation. Euploid: refers to normal chromosomal number. Aneuploid: refers to single monosomy or single trisomy. Aneuploid ≤ 5 refers to more than one aneuploidy. Multiple Aneuploid >5: refers to more than 5 aneuploidies. *: refers to multiple chromosomal aneuploidies which were difficult to score. N/A: not available.
7.2.4 Validation of mtDNA template quantification technique on small number of cells:

Real-Time PCR conditions were first optimised using the diluted oligonucleotides and single cells (lymphocytes). Following that, the same conditions were applied on whole embryos and single blastomeres in order to detect the efficiency of the PCR conditions. The efficacy of ALB and PK lysis buffers was tested indicating that both lysis methods work well when real time PCR by Resolight or TaqMan assays was applied. However, PK lysis method was preferred since one round of PCR is needed. Standards were diluted in a 1/10 serial dilution and were loaded with each run. A standard curve plotting the concentration of the diluted standards and the corresponding Cq value was performed with each run. The concentration of each sample was calculated using the formula derived from the linear regression of the standard curve performed (Figure 7-1).

![Figure 7-1: Standard curve used in mtDNA template analysis for samples lysed by PK.](image)

The X-axis represents the mtDNA template while the Y-axis represents the Cq value. The equation from the standard curve on the right was used to calculate the mtDNA template of each sample. The PK lysed embryos mtDNA template was calculated from the equation derived from the above standard curve. The log of each value was taken which represents the mtDNA template concentration. Since PK lysis method was used when samples were treated by Resolight and TaqMan the same equation was used for those samples while ALB lysed embryos were analysed based on the equation derived from the standards curve of standards run with ALB.
7.2.5 Validation of standard curves used for mtDNA template quantification:

In order to validate the standard curves used in establishing the mtDNA template, a statistical method was used for validation. This was performed by creating appropriate interaction terms to check if the slopes are different than their curves in case the interaction gave a significant coefficient. If a significant interaction between the curves and the value of X was observed, then there is an indication of a difference between the slopes. Thus, the standard curves are different with each run. However, if no significant difference was indicated, this would suggest that the standard curves are the same. Hence, a standard curve with each run is not necessary. No significant interaction between the repeats for mtDNA template was found. Hence, the same equation can be used for mtDNA template concentration (Figure 7-2).
Figure 7–2: Standard curve for testing the distribution accuracy of the standards dilution used for mtDNA template analysis.

Regression of Y (Cq value) against X (mtDNA template), including interaction term to test whether the slopes of the repeat samples are parallel. No significant interaction between the repeats for mtDNA template was found.
7.3 Appendix C

7.3.1 Identification of mitochondrial haplogroups by EMPOP

Figure 7–3: Mitochondrial haplogroups assignment using EMPOP.

D, a mix of A, G, T (IUB code); d, a mix of A,G,T and deletion. changes were made to some SNPs: 514d to 514del, 515d to 515del.
### 7.3.2 List of AIM available on the Karyomap SNP chip

<table>
<thead>
<tr>
<th>FULL SET OF 62 AIMs: CLASSIFICATION AS AFRICAN/EUROPEAN/EAST ASIAN</th>
<th>SUBSET OF 33 AIMs: ADDS SOUTH ASIAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1039630</td>
<td>rs1079597</td>
</tr>
<tr>
<td>rs1040577</td>
<td>rs12498138</td>
</tr>
<tr>
<td>rs10413216</td>
<td>rs1426654</td>
</tr>
<tr>
<td>rs10416569</td>
<td>rs2024566</td>
</tr>
<tr>
<td>rs10768085</td>
<td>rs260690</td>
</tr>
<tr>
<td>rs11052203</td>
<td>rs310644</td>
</tr>
<tr>
<td>rs1113337</td>
<td>rs3737576</td>
</tr>
<tr>
<td>rs1125425</td>
<td>rs3823159</td>
</tr>
<tr>
<td>rs1533677</td>
<td>rs3827760</td>
</tr>
<tr>
<td>rs1546541</td>
<td>rs4471745</td>
</tr>
<tr>
<td>rs173537</td>
<td>rs7251928</td>
</tr>
<tr>
<td>rs1888952</td>
<td>rs7722456</td>
</tr>
<tr>
<td>rs1937147</td>
<td>rs798443</td>
</tr>
<tr>
<td>rs275837</td>
<td>rs870347</td>
</tr>
<tr>
<td>rs282496</td>
<td>rs9522149</td>
</tr>
<tr>
<td>rs2851060</td>
<td>rs10183022</td>
</tr>
<tr>
<td>rs2965282</td>
<td>rs1557553</td>
</tr>
<tr>
<td>rs595961</td>
<td>rs1592672</td>
</tr>
<tr>
<td>rs727878</td>
<td>rs17130385</td>
</tr>
<tr>
<td>rs877783</td>
<td>rs2180052</td>
</tr>
<tr>
<td>rs892263</td>
<td>rs2471552</td>
</tr>
<tr>
<td>rs915056</td>
<td>rs4780476</td>
</tr>
<tr>
<td>rs959858</td>
<td>rs4979274</td>
</tr>
<tr>
<td>rs984654</td>
<td>rs6088466</td>
</tr>
<tr>
<td>rs987284</td>
<td>rs6875659</td>
</tr>
<tr>
<td>rs1048610</td>
<td>rs7151991</td>
</tr>
<tr>
<td>rs1907702</td>
<td>rs7307862</td>
</tr>
<tr>
<td>rs7213527</td>
<td>rs7630522</td>
</tr>
<tr>
<td>rs926774</td>
<td>rs820371</td>
</tr>
<tr>
<td>-</td>
<td>rs881929</td>
</tr>
<tr>
<td>-</td>
<td>rs11779571</td>
</tr>
<tr>
<td>-</td>
<td>rs7354930</td>
</tr>
<tr>
<td>-</td>
<td>rs9487258</td>
</tr>
</tbody>
</table>
7.3.3 Script used in the IBS analysis

Python script

# plink --bfile Razan --distance 1-ibs --out Razan_160809 (THE ONE WE USED) # # ids=[]

with open(RES + 'Razan_160809.mdist.id') as to_parse: for line in to_parse:
	_tmp, _id = line.strip().split('t')
	ns.append(_id) #print(ids)

# # # id_to_id_to_distance = defaultdict(dict) with open(RES + 'Razan_160809.mdist') as to_parse:
for i, line in enumerate(to_parse): line = line.strip().split('t') for j, dist in enumerate(line):
	# if ids[i+1] in ['New-E8', 'SJ'] or ids[j] in ['New-E8', 'SJ']:
	continue #

id_to_id_to_distance[ids[i+1]][ids[j]] = float(dist)

id_to_id_to_distance[ids[j]][ids[i+1]] = float(dist) #print(id_to_id_to_distance['AP-Mother']['AP-Father']) # # # pairs_controls = [(‘AP-Mother’, ‘AP-Father’), (‘JC-Mother’, ‘JC-Father’), (‘JT-Mother’, ‘JT-Father’),
#

#
The probe IDs of the identified mitochondrial SNPs showing differences between fertile and RM/RIF groups analysed by the CoreExome SNP chip.

<table>
<thead>
<tr>
<th>Probe ID</th>
<th>Diff %</th>
<th>rCRs position</th>
<th>position</th>
</tr>
</thead>
<tbody>
<tr>
<td>MitoT9900C</td>
<td>0.2</td>
<td>9899</td>
<td>MT-CO3</td>
</tr>
<tr>
<td>MitoA4918G</td>
<td>0.4</td>
<td>4917</td>
<td>MT-ND2</td>
</tr>
<tr>
<td>exm-rs2853498</td>
<td>0.2</td>
<td>12308</td>
<td>MT-TL2</td>
</tr>
<tr>
<td>exm2263337</td>
<td>0.3</td>
<td>15928</td>
<td>MT-TT</td>
</tr>
<tr>
<td>exm2263338</td>
<td>0.4</td>
<td>4917</td>
<td>MT-ND2</td>
</tr>
<tr>
<td>MitoT5005C</td>
<td>0.2</td>
<td>5004</td>
<td>MT-ND2</td>
</tr>
<tr>
<td>exm-rs3088309</td>
<td>0.3</td>
<td>15452</td>
<td>MT-ND2</td>
</tr>
<tr>
<td>MitoG15929A</td>
<td>0.3</td>
<td>15928</td>
<td>MT-TT</td>
</tr>
<tr>
<td>exm-rs28358279</td>
<td>0.3</td>
<td>10463</td>
<td>MT-TR</td>
</tr>
<tr>
<td>MitoA11252G</td>
<td>0.3</td>
<td>11251</td>
<td>MT-ND4</td>
</tr>
<tr>
<td>exm2216232</td>
<td>0.3</td>
<td>1888</td>
<td>MT-RNR2</td>
</tr>
<tr>
<td>MitoT12706C</td>
<td>0.2</td>
<td>12705</td>
<td>MT-ND5</td>
</tr>
<tr>
<td>MitoA11468G</td>
<td>0.2</td>
<td>11467</td>
<td>MT-ND4</td>
</tr>
<tr>
<td>exm2216249</td>
<td>0.235</td>
<td>2706</td>
<td>MT-RNR2</td>
</tr>
<tr>
<td>exm2216228</td>
<td>0.2</td>
<td>1811</td>
<td>MT-RNR2</td>
</tr>
<tr>
<td>MitoA15925G</td>
<td>0.2</td>
<td>15924</td>
<td>MT-TT</td>
</tr>
<tr>
<td>MitoA4025G</td>
<td>0.2</td>
<td>4024</td>
<td>MT-ND1</td>
</tr>
</tbody>
</table>
7.4 Appendix D

7.4.1 Script to filter out valid SNPs for the analysis

- The level of coverage in the enriched vs unenriched genes was about a $2^8$ (250X) enrichment of reads in the selected regions. Thus, the capture was considered to have worked well.

- There was one embryonic sample which had an unusual profile but this was disregarded since a replicate was available for that sample.

- Variants were called, and BAM files were generated to include the variants of parents and their for each family separately.

- Homozygous SNPs were excluded from the analysis as they will always produce the same type of variant at these particular positions.

- After merging the FASTQ files of each family separately (parents and their embryos). A script was compiled by Dr. Andrews at the Babraham Institute to filter out the uninformative SNPs in each family.

7.4.1.1 Script to filter-out un informative SNPs

#!/usr/bin/perl
use warnings;
use strict;

my @files = <*SNPS*.txt>;
foreach my $file (@files) {
    process_file($file);
}

sub process_file {
    my ($file) = @_;  
    open (IN,$file) or die "Can't read $file: $!";
    open (OUT,'>',"../Valid_Family_Variant_Tables/$file") or die "Can't write outfile: $!";
    my $header = <IN>;
    print OUT $header;
}
chomp $header;

my @names = split(/\t/, $header);

# Find the mother
my $mother_index;
my $father_index;

foreach my $index (0..$#names) {
    if ($names[$index] =~ /MOTHER/) {
        $mother_index = $index;
    }
    if ($names[$index] =~ /FATHER/) {
        $father_index = $index;
    }
}

unless ($mother_index and $father_index) {
    die "Couldn't find mother and father from $header";
}

my %counts;

VARIANT: while (<IN>) {
    ++$counts{total};
    chomp;
    my @variants = split(/\t/);
    my @mother_alleles;
    my @father_alleles;

    # Find which are the mother and the father
    for my $index (2..$#variants) {
        if ($variants[$index] !~ /^[GATC]/[GATC]$/) {
            # warn
            "$variants[$index] isn't a valid variant
";
            ++$counts{unobserved};
            next VARIANT;
        }
        if ($index == $mother_index) {
            @mother_alleles = split(/\//, $variants[$index]);
        }
        if ($index == $father_index) {
            @father_alleles = split(/\//, $variants[$index]);
        }
    }
}

Razan Jawdat / PhD thesis
# Find variants which aren't compatible with
# or where all of the genotypes are the same
my $found_multiple_genotypes = 0;

for my $index (2..$#variants) {
    if ($variants[$index] ne $variants[2]) {
        $found_multiple_genotypes = 1;
    }
}

next if ($index == $mother_index or $index == $father_index);

my $allele_valid = 0;

ALLELE: foreach my $mother_allele (@mother_alleles) {
    foreach my $father_allele (@father_alleles) {
        if ($variants[$index] eq "$mother_allele/$father_allele" or $variants[$index] eq "$father_allele/$mother_allele") {
            $allele_valid = 1;
            last ALLELE;
        }
    }
}

unless ($allele_valid) {
    #warn "Allele $variants[$index] wasn't compatible with @mother_alleles and @father_alleles\n";
    ++$counts{incompatible};
    next VARIANT;
}

unless ($found_multiple_genotypes) {
    #warn "Single genotype in @variants\n";
    ++$counts{no_variation};
    next;
}

++$counts{valid};

print OUT $_,"\n";

print "Summary for $file\n";

foreach my $key (sort { $counts{$b} <=> $counts{$a} } keys %counts) {
    print "\t$key":$counts{$key}."\n";
}

print "\n\n";
The codes in the previous file are as follows:
unobserved = one or more samples did not have a genotype call for that position
incompatible = the genotype of one of the embryos was not compatible with the
genotypes reported for the parents
no variation = all of the genotypes were the same in all children and parents (but
were all different to the reference genome)
valid = the subset of SNPs for analysis.

• genotypes which only occur in the arrested embryos and never in the
blastocysts or in either of the parents were selected for further analysis.

7.4.1.1.2 Script used to pull out the variants that were only present in poor morphology
embryos in each family and not present in blastocysts.

#!/usr/bin/perl
use warnings;
use strict;

# This script goes through a variant table to find genotypes which only
ever
# appear in bad embryos

my $morphologies;
read_morphologies();

my @files = <*>table.txt>;

open (OUT,'>', "only_bad_morphology_genotypes.txt") or die $!;
print OUT
$chr","pos","family","genotype","bad_with_this_genotype","bad_with_different_count","good_count"),"n";

foreach my $file (@files) {
    process_file($file);
}

sub process_file {
    my ($file) = @_;

    my $family;
    if ($file =~ /^(\w+)-FAMILY/) {
        $family = $1;
        warn "Looking at $family family\n";
    }
    else { 
        die "Couldn't get family from $file";
    }
open (IN,$file) or die $!

$_ = <IN>;
chomp;

my @header = split(/\t/);

my %embryo_indices;
my %bad_morphology_indices;

# We need to work out the indices of each of the Embryos
# so we can later identify which of these have bad
# morphology

foreach my $index (0..$#header) {
  if ($header[$index] =~ /-(E\d+).GT/) {
    $embryo_indices{$1} = $index;
    warn "Index for $1 is $index\n";
    unless (exists $morphologies->${family}-{$1}) {
      die "No morphology for $1 in $family\n";
    }
    if ($morphologies->${family}-${1}-{{morphology} eq 'BAD'}) {
      $bad_morphology_indices{$index} = 1;
    }
  }
}

if ((scalar keys %embryo_indices) - (scalar keys %bad_morphology_indices) == 0) {
  warn "Skipping family $family as all embryos are bad\n";
  close IN;
  return;
}

warn "There are ".scalar (keys %bad_morphology_indices)." bad morphology embryos in $family\n";

# Now we can go through the variants. In each case we'll
# look at the genotypes found in the
# bad embryos and try to find those which only occur in a
# bad embryo and never in a good one
# or in one of the parents.

while (<IN>) {
  chomp;
  my @genotypes = split(/\t/);
  my %genotype_counts;
  for my $index (2..$#genotypes) {

if (exists $bad_morphology_indices{$index}) {
    ++$genotype_counts{$genotypes[$index]}->{BAD};
} else {
    ++$genotype_counts{$genotypes[$index]}->{GOOD};
}

# Now go through the genotypes and see if there are any
# which are exclusively in BAD morphology samples
foreach my $genotype (keys %genotype_counts) {
    next unless ($genotype_counts{$genotype}->{BAD}); # We don't care about genotypes which are only good
    next if ($genotype_counts{$genotype}->{GOOD}); # We don't care about genotypes which can be both good and bad

    # We're going to report
    # # Chr
    # # Pos
    # # Family
    # # Genotype
    # # No of bad with this genotype
    # # No of bad with a different genotype
    # # No of good embryos with a different genotype
    my $chr = $genotypes[0];
    my $pos = $genotypes[1];
    my $bad_with_count = $genotype_counts{$genotype}->{BAD};
    my $bad_different_count = (scalar keys %bad_morphology_indices) - $bad_with_count;
    my $good_count = (scalar keys %embryo_indices) - (scalar keys %bad_morphology_indices);

    print OUT join("	",($chr,$pos,$family,$genotype,$bad_with_count,$bad_different_count,$good_count)),"\n";
}
}
sub read_morphologies {
    open (IN,"../embryo_morphologies.txt") or die $!
    $_ = <IN>
    while (<IN>) {
        chomp;
        next unless ($_);
        my ($family,$embryo,$haplotype,$goodbad) = split(/\t/);
        $morphologies-{$family}-{$embryo} = {morphology => $goodbad,haplotype=>$haplotype};
    }
    close IN;
}

• The predicted effect of the selected variants was checked by the variant effect predictor on ensemble. Input files for VEP were first generates using a script.

7.4.1.1.3 Script used to generate files to be further used by VEP.

#!/usr/bin/perl
use warnings;
use strict;

# This script reformats our genotype data to the format required by the Ensembl Variant Effect Predictor (VEP).
# Our starting format looks like:
# chr    pos    family    genotype    bad_with_this_genotype    bad_with_different_count    go
# chr1   57111169 AP       G/G     1       2       3
# The Ensembl format is:
# chr    start    end    allele    strand    identifier
# For SNPS the start/end are the same.
# The allele format is ref/alt so we'll need to pull the ref from the genome sequence
# the strand is going to be forward for all of these - hopefully that doesn't matter
# the identifier is optional and we won't have this.
my ($file) = @ARGV;

# Read the data, split different genotypes and merge
# together anything coming from different families
open (IN,$file) or die "Can't read $file :$!";
$_ = <IN>;
my %genotypes;
while (<IN>) {
    chomp;
    my ($chr,$pos,$family,$genotype) = split(/\t/);
    my($genotype1,$genotype2) = split(/\//,$genotype);
    my @genotypes = ($genotype1);
    if ($genotype2 ne $genotype1) {
        push @genotypes,$genotype2;
    }
    foreach my $genotype (@genotypes) {
        my $identifier = "$chr:$pos:$genotype";
        unless (exists $genotypes{$identifier}) {
            $genotypes{$identifier} =
                {chr => $chr,
                 pos => $pos,
                 ref => undef,
                 alt => $genotype,
                 families => []};
        }
        push @{$genotypes{$identifier}->{$families}},$family;
    }
}
close IN;

# We now need to get the reference alleles for all of the variants
my @identifiers = sort keys %genotypes;
open (OUT,'>','ref_positions.txt') or die $!;
print OUT join("\t",("chr","start","end","strand")),"\n";
foreach my $id (@identifiers) {
    my $chr = $genotypes{$id}->{chr};
    $chr =~ s/chr//;
    print OUT join("\t",($chr,$genotypes{$id}->{pos},$genotypes{$id}->{pos},"+")),"\n";
}
system("./fetch_grch37.pl ref_positions.txt > ref_positions.fa") == 0 or die "Failed to fetch ref positions";

# Now we can read in the bases we pulled out of the genome and put these back into the variant data structure

open (IN,"ref_positions.fa") or die "Can't read ref_positions.fa: "$!";

my $index = 0;

while (<IN>) {
    # First line is the id, the next line is the sequence
    my $seq = <IN>;
    chomp $seq;
    $genotypes{$identifiers[$index]}->{ref} = $seq;
    ++$index;
}

close IN;

# Now we can write out the VEP files

my $outfile = $file;

$outfile =~ s/\..*/\._vep.txt/;

open (OUT,'>$outfile') or die "Can't write to $outfile: "$!";

foreach my $id (@identifiers) {
    next if ($genotypes{$id}->{ref} eq $genotypes{$id}->{alt});
    my @families = @{$genotypes{$id}->{families}};
    my $id_with_family = $id . ":" . join(":" , @families);
    warn "Families are '@families'
    warn "ID is $id_with_family"
    my $chr = $genotypes{$id}->{chr};
    $chr =~ s/chr//i;
    print OUT join("\t",
        $chr,
        $genotypes{$id}->{pos},
        $genotypes{$id}->{pos},
        $genotypes{$id}->{ref}."/" . $genotypes{$id}->{alt},
        '1',
    );
}
The same analysis was performed for the indels.

7.4.1.4 Script used for filtering valid indels in families

```perl
#!/usr/bin/perl
use warnings;
use strict;

my @files = <$INDELS*.txt>;
foreach my $file (@files) {
    process_file($file);
}

sub process_file {
    my ($file) = @_; 
    open (IN,$file) or die "Can't read $file: $!";
    open (OUT,'>',"../Valid_Family_Variant_Tables/$file") or die "Can't write outfile: $!";
    my $header = <IN>;
    print OUT $header;
    chomp $header;
    my @names = split(/\t/, $header);
    # Find the mother
    my $mother_index;
    my $father_index;
    foreach my $index (0..$#names) {
        if ($names[$index] =~ /MOTHER/) {
            $mother_index = $index;
        }
        if ($names[$index] =~ /FATHER/) {
            $father_index = $index;
        }
    }
```
unless ($mother_index and $father_index) {
    die "Couldn't find mother and father from $header";
}

my %counts;

VARIANT: while (<IN>) {
    ++$counts{total};
    chomp;
    my @variants = split(/\t/);
    my @mother_alleles;
    my @father_alleles;

    # Find which are the mother and the father
    for my $index (2..$#variants) {
        if ($variants[$index] !~ /^\[GATC\]+\/[GATC]+$/) {
            # warn
            "$variants[$index] isn't a valid variant\n"
            ++$counts{unobserved};
            next VARIANT;
        }
        if ($index == $mother_index) {
            @mother_alleles = split(/\//,$variants[$index]);
        }
        if ($index == $father_index) {
            @father_alleles = split(/\//,$variants[$index]);
        }
    }

    # Find variants which aren't compatible with the parents
    # or where all of the genotypes are the same
    my $found_multiple_genotypes = 0;
    for my $index (2..$#variants) {
        if ($variants[$index] ne $variants[2]) {
            $found_multiple_genotypes = 1;
        } else {
            next if ($index == $mother_index or $index == $father_index);
        }
    }
    if ($found_multiple_genotypes) {
        my @alleles = split(/\//,$variants[2]);
        for my $allele (0..$#alleles) {
            unless ($alleles[$allele] eq $variants[2]) {
                my $allele_valid = 0;
            }
        }
    }

    # Find which don't have a compatible parent genotype
    for my $index ($#variants) {
        if ($variants[$index] ne $variants[2]) {
            next if ($index == $mother_index or $index == $father_index);
        }
    }
}

VARIANT: while (<IN>) {
    ++$counts{total};
    chomp;
    my @variants = split(/\t/);
    my @mother_alleles;
    my @father_alleles;

    # Find which are the mother and the father
    for my $index (2..$#variants) {
        if ($variants[$index] !~ /^\[GATC\]+\/[GATC]+$/) {
            # warn
            "$variants[$index] isn't a valid variant\n"
            ++$counts{unobserved};
            next VARIANT;
        }
        if ($index == $mother_index) {
            @mother_alleles = split(/\//,$variants[$index]);
        }
        if ($index == $father_index) {
            @father_alleles = split(/\//,$variants[$index]);
        }
    }

    # Find variants which aren't compatible with the parents
    # or where all of the genotypes are the same
    my $found_multiple_genotypes = 0;
    for my $index (2..$#variants) {
        if ($variants[$index] ne $variants[2]) {
            $found_multiple_genotypes = 1;
        } else {
            next if ($index == $mother_index or $index == $father_index);
        }
    }
    if ($found_multiple_genotypes) {
        my @alleles = split(/\//,$variants[2]);
        for my $allele (0..$#alleles) {
            unless ($alleles[$allele] eq $variants[2]) {
                my $allele_valid = 0;
            }
        }
    }

    # Find which don't have a compatible parent genotype
    for my $index ($#variants) {
        if ($variants[$index] ne $variants[2]) {
            next if ($index == $mother_index or $index == $father_index);
        }
    }
}
7.4.1.1.5 Script used for filtering indels associated with arrested embryos and not present in blastocysts

#!/usr/bin/perl
use warnings;
use strict;

# This script goes through a variant table to find genotypes which only ever
# appear in bad embryos

my $morphologies;
read_morphologies();

my @files = <*INDELS*table.txt>;

open (OUT,'>', "only_bad_morphology_indels.txt") or die $!
print OUT join("	",("chr","pos","family","genotype","bad_with_this_genotype","bad_with_different_count","good_count")),"\n";

foreach my $file (@files) {
    process_file($file);
}

sub process_file {
    my ($file) = @_;

    my $family;
    if ($file =~ /^\w+-FAMILY/) {
        $family = $1;
        warn "Looking at $family family\n";
    } else {
        die "Couldn't get family from $file";
    }
    open (IN,$file) or die $!;
    $_ = <IN>;
    chomp;
    my @header = split(/\t/);
    my %embryo_indices;
    my %bad_morphology_indices;
    # We need to work out the indices of each of the Embryos
    # so we can later identify which of these have bad
    # morphology
    foreach my $index (0..$#header) {
        if ($header[$index] =~ /\d+\d+/) {
            $embryo_indices{$1} = $index;
            warn "Index for $1 is $index\n";
        }
    }
   }%

    if (exists $morphologies->{$family}->{morphology} eq 'BAD') {
        $bad_morphology_indices{$index} = 1;
    }
if ((scalar keys %embryo_indices) - (scalar keys %bad_morphology_indices) == 0) {
    warn "Skipping family $family as all embryos are bad\n";
    close IN;
    return;
}

warn "There are ". scalar (keys %bad_morphology_indices)." bad morphology embryos in $family\n";

# Now we can go through the variants. In each case we'll look at the genotypes found in the # bad embryos and try to find those which only occur in a bad embryo and never in a good one # or in one of the parents.
while (<IN>) {
    chomp;
    my @genotypes = split(/\t/);
    my %genotype_counts;
    for my $index (2..$#genotypes) {
        if (exists $bad_morphology_indices{$index}) {
            ++$genotype_counts{$genotypes[$index]}->{BAD};
        } else {
            ++$genotype_counts{$genotypes[$index]}->{GOOD};
        }
    }
    # Now go through the genotypes and see if there are any # which are exclusively in BAD morphology samples
    foreach my $genotype (keys %genotype_counts) {
        next unless ($genotype_counts{$genotype}->{BAD}); # We don't care about genotypes which are only good
        next if ($genotype_counts{$genotype}->{GOOD}); # We don't care about genotypes which can be both good and bad
        # We're going to report
        # Chr
        # Pos
        # Family
        # Genotype
my $chr = $genotypes[0];
my $pos = $genotypes[1];
my $bad_with_count =
$genotype_counts{$genotype}->{BAD};
my $bad_different_count =
(scalar keys %bad_morphology_indices) - $bad_with_count;
my $good_count = (scalar keys %embryo_indices) -
(scalar keys %bad_morphology_indices);
print OUT
join("\t",($chr,$pos,$family,$genotype,$bad_with_count,$bad_different_count ,
,$good_count)),"\n";

}

}

sub read_morphologies {

    open (IN, "../embryo_morphologies.txt") or die $!
    ;
    $_ = <IN>;
    while (<IN>) {
        chomp;
        next unless ($_);

        my ($family,$embryo,$haplotype,$goodbad) =
        split("\t");

        $morphologies->{$family}->{$embryo} =
        {morphology => $goodbad,haplotype=>$haplotype};
    }
    close IN;
}
Chapter 8 Publications

8.1 Published abstracts

✓ Poster presentation at the 12th Meeting of the European Human Genetics Societies 2016, Barcelona, Spain, 21-24 May 2016 (ESHG-2016).


8.2 Manuscripts in preparation

✓ **Jawdat; R., Lane; N., SenGupta; SB.** Mitonuclear-mismatch and preimplantation human embryo development. (Review). In preparation.

✓ **Jawdat; R., Dewhurst; E., Diekmann; Y., Haj Ali; A., Sun; X., Lane; N., Serhal; P., Philips; C., SenGupta; SB.** The effect of parental ethnicity on embryo quality. (In preparation).

✓ **Jawdat; R., Serhal; P., Lane; N., SenGupta; SB.** (2018) Carriers for mitochondrial haplogroup T show a decline in blastocyst formation through nuclear genes important in mitochondrial function. (In preparation).

✓ **Jawdat; R., Naja; R., Serhal; P., Odia; R., Lane; N., SenGupta; SB.** (In preparation).

Elevated mitochondrial DNA is not associated with aneuploidy in preimplantation embryos arrested at cleavage.
Chapter 9 References

9.1 Websites

- Ensembl Genome Browser: www.ensembl.org
- Rubicon genomics: www.rubicongenomic.com
- BLAST sequence comparison tool: www.ensembl.org/Homo_sapiens/blastview
- Primer3 primer designing software: frodo.wi.mit.edu/
- http://broadinstitute.github.io/picard
- (http://mathgene.usc.es/snipper/).
- Snipper: Binary AIM classification of multiple individuals

9.2 Books

9.3 Papers


GARDNER, R. L. & JOHNSON, M. H. 1972. An investigation of inner cell mass and
trophoblast tissues following their isolation from the mouse blastocyst. J
Embryol Exp Morphol, 28, 279-312.
other trisomic conditions. Mutat Res, 296, 69-88.
GERSHONI, M., FUCHS, A., SHANI, N., FRIDMAN, Y., CORRAL-DEBRINSKI, M.,
direct interactions between mtDNA-encoded and nDNA-encoded subunits of
GIANAROLI, L., LUISELLI, D., CRIVELLO, A. M., LANG, M., FERRARETTI, A. P.,
analysis and numerical chromosome condition in human oocytes and polar
bodies. Mol Hum Reprod, 21, 46-57.
GIGAREL, N., RAY, P. F., BURLET, P., FRYDMAN, N., ROYER, G., LEBON, S.,
Single cell quantification of the 8993T>G NARP mitochondrial DNA mutation
GORDO, A. C., RODRIGUES, P., KUROKAWA, M., JELLERETTE, T., EXLEY, G.
signal apoptosis rather than activation in in vitro aged mouse eggs. Biol
Reprod, 66, 1828-37.
a011403.
GRIMSTAD, F. & KRIEG, S. 2016. Immunogenetic contributions to recurrent
Alpha-lipoic acid as a new treatment option for Alzheimer's disease--a 48
months follow-up analysis. J Neural Transm Suppl, 189-93.
HALDER, I., SHRIVER, M., THOMAS, M., FERNANDEZ, J. R. & FRUDAKIS, T.
2008. A panel of ancestry informative markers for estimating individual
biogeographical ancestry and admixture from four continents: utility and
applications. Hum Mutat, 29, 648-58.
HALLER-KIKKATALO, K., ALTMAE, S., TAGOMA, A., UIBO, R. & SALUMETS, A.
2014. Autoimmune activation toward embryo implantation is rare in immune-
HAMAMY, H. 2012. Consanguineous marriages : Preconception consultation in
HARDE, J. & SINGLETON, A. 2009. Genomewide association studies and human
HARDY, K., HOOPER, M. A., HANDYSIDE, A. H., RUTHERFORD, A. J.,
glucose and pyruvate uptake by individual human oocytes and
embryo development. Reproduction, 123, 479-86.
more accurate time scale for the human mitochondrial DNA tree. J Mol Evol,
37, 347-54.


SMEETS, H. J. 2013. Preventing the transmission of mitochondrial DNA disorders: selecting the good guys or kicking out the bad guys. Reprod Biomed Online.


VERLAAN, D. J., GE, B., GRUNDBERG, E., HOBERMAN, R., LAM, K. C., KOKA, V., DIAS, J., GURD, S., MARTIN, N. W., MALLMIN, H., NILSSON, O.,


