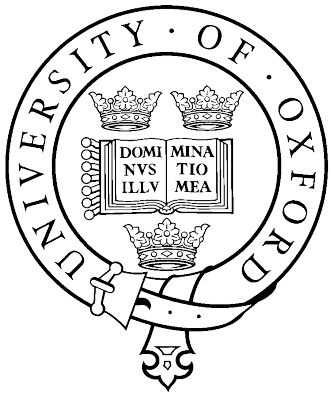
Novel approaches for embryonic competency evaluation



Columba Avila Perez
Green Templeton College
University of Oxford

Supervised by:
Prof Dagan Wells
Nuffield Department of Women's and Reproductive Health

A thesis submitted for the degree of
Doctor of Philosophy
Trinity 2023

This chapter of my life in Oxford started with a shared dream between my Dad and I. I dedicate this work to him.

Acknowledgements

I would like to first thank Dagan for opening so many doors for me. These opportunities allowed me to be part of very exciting, and many times challenging, projects that offered me a lot of academic and personal learnings. I would also like to thank Kevin, my secondary supervisor, and Celine. They have been massive supports, even before I first arrived in Oxford.

Working at Dagan's lab let me cross paths with wonderful people. To my fellow D.Phil. and postdoc students, Megan, Katharina, Kishlay, Zeynep, Ayman, Georgina and Nada, thanks for being good friends and for always enthusiastically sharing your knowledge. To my friends in Juno Genetics, Elena, Ana, Alicia, Tishy, Hadis and Saul, thanks for being amazing and fun colleagues during my time in Juno. I would like to especially acknowledge Dhruti as an integral part of my thesis, thank you for being a huge source of day-to-day guidance and advice.

Individuals outside Oxford were also fundamental. I am greatly thankful to my fiancé, Brooke. His unwavering support was the keystone supporting my D.Phil. This thesis would not have been possible either without the presence of my beloved siblings, Daniela and Jose. I would also like to thank the unconditional love during long writing sessions offered by my furry friends. Leonie, in Oxford, Terry in Guadalajara and Hachi, Toti, Glass, Lulu, Titi and Camila in Zacatecas.

Lastly, I want to acknowledge the essential role my parents played in my D.Phil. journey. Thanks, Dad, for passing on to me your eagerness for education and preparation. Without my Dad's sacrifices and my Mum's endless encouragement, I would not have been able to set foot in Oxford six years ago.

Abstract

Thesis title:	*Novel approaches for embryonic competency evaluation*
Name:	*Columba Avila Perez*
College:	*Green Templeton*
Department:	*Nuffield Department of Women's and Reproductive Health*
Degree:	*D.Phil. in Women's and Reproductive Health*
Thesis submission term:	*Trinity 2023*

One crucial factor in assisted reproductive technologies (ARTs) is the identification of embryos with the highest potential to implant and produce healthy offspring. Today, ploidy status is broadly employed as a predictive factor of embryo viability, but in vitro fertilisation (IVF) sustained implantation rates remain low.

Remarkable progress in reproductive genetics has helped to understand infertility, but much is still unknown about gametogenesis, the preimplantation stages and aneuploidy. Segmental chromosomal abnormalities result from unrepaired DNA damage, whilst numerical aneuploidy, prevalent in embryos from advanced maternal age, is the result of chromosome mis-segregation. Evidence suggests that DNA double-strand breaks (DSBs) repair capabilities may be deficient during the first mitoses in the embryo, prior to the major embryonic genome activation (EGA) wave. Several genes are involved in DNA damage repair in somatic cells, including *ATM*. Studies have found that DNA damage response (DDR) and chromosome segregation surveillance pathways share key proteins to prevent aneuploidy. Interestingly, the expression of genes involved in the ATM-mediated DSBs repair pathway decreases in germinal vesicles (GVs) from reproductively aged women. This thesis studies the expression of essential DDR genes in human preimplantation embryos and the competency of mice oocytes to produce euploid embryos when *Atm* is downregulated. The results present instances of the upregulation of DDR genes at the 5- to 6-cell stage, in line with other studies, suggesting that embryos may be capable of responding to DSBs either before or at the beginning of the major EGA wave, albeit sub optimally. This study, along with the existing literature, fails, however, to establish an association between ATM-pathway dysfunction in the mice GV and impaired meiotic chromosome segregation and DSBs repair.

Preimplantation genetic testing for aneuploidies (PGT-A) routinely requires a trophectoderm (TE) biopsy. Thus, it needs sophisticated personnel and equipment to avoid significant harm to the embryo. PGT-A, therefore, adds to the cost of IVF treatment, already considered inaccessible to many. Non-invasive prenatal testing (NIPT) is becoming a routine screening test. However, non-invasive methods for aneuploidy testing in embryos are still being investigated. So far, although spent culture media (SCM) has shown more promising results than blastocoel fluid (BF), discrepancies exist in the ploidy concordance rates with TE likely due to varying levels of embryo manipulation. This thesis assesses the accuracy of a truly non-invasive PGT (niPGT) protocol and develops PCR-based methods for the detection of the Y chromosome for NIPT and in horse BF. Results show that, when compared to other studies, embryo interventions facilitate the release of embryonic DNA into the media. They also challenge the hypothesis that a bigger blastocyst size provides better embryonic DNA amplification levels. This section of the thesis does produce promising results for a highly sensitive PCR-based Y chromosome NIPT at week-10 gestations.

Preimplantation development is characterised by a complex interplay between genetic, epigenetic and metabolic factors. The epigenome is particularly vulnerable during the preimplantation embryonic development. Studies have aimed to analyse the presence of methylation perturbations in imprinted genes in gametes and embryos with a high variation in the reported DNA methylation levels. The last part of the thesis explores the use of unique molecular identifiers (UMIs) to directly count molecules in a bisulfite sequencing (BS) approach that simultaneously detects methylation perturbations in 46 human differentially methylated regions (DMRs). One major impediment of this approach at a single-cell level is the inability to introduce the UMIs before amplification since identifiers are not able to account for bias during the first PCR of library preparation. Results also show that this is problematic because the degree of PCR amplification bias is higher in lower DNA inputs.

Table of Contents

List of abbreviations 1

Chapter 1. Introduction 3

1.1.1. Introduction 3

1.1.1.1 Thesis outline and aims 5

1.1.2. Background 7

1.1.2.1. Infertility and assisted reproductive technologies (ARTs) 7

1.1.2.2. ARTs procedures 8

1.1.2.3. Preimplantation genetic testing (PGT) for embryo selection 13

1.1.2.4. Factors affecting human embryo competency 17

Chapter 2. General materials and methods 28

2.1.1. Sample preparation 29

2.1.1.1. Buccal cell collection and isolation 29

2.1.1.2. DNA extraction from clumps of cells 29

2.1.1.3. DNA and cell-free DNA (cfDNA) isolation from blood 30

2.1.1.4. RNA extraction from embryos and reverse transcription 32

2.1.2. DNA fragmentation 33

2.1.3. DNA amplification 34

2.1.3.1. Multiplex PCR 34

2.1.3.2. Quantitative real time PCR (RT-qPCR) 35

2.1.3.3. Pre-amplification of complementary DNA (cDNA) 36

2.1.3.4. Whole genome amplification (WGA) 37

2.1.4. DNA visualisation 38

2.1.4.1. Agarose gel electrophoresis 38

2.1.4.2. TapeStation (Applied Biosystems, USA) 39

2.1.5. Measurement of DNA concentrations 39

2.1.6. DNA bisulfite conversion 39

2.1.6.1. EpiTect Bisulfite Conversion kit (Qiagen, Germany) – Protocol: 'Sodium bisulfite conversion of unmethylated cytosines in small amounts of fragmented DNA' 40

2.1.6.2. Imprint Modification kit (Sigma-Aldrich, USA) – Protocol: 'Two-Step Modification Procedure' 41
2.1.7. Library preparation and next-generation sequencing (NGS) 42
2.1.7.1. VeriSeq (Illumina, USA) 42
2.1.7.2. Oxford Nanopore MinION (Oxford Nanopore, UK) 44
2.1.7.3. PG-Seq (Perkin Elmer, USA) 47
2.1.8. Bioinformatics 49
2.1.8.1. DNA sequence retrieval 49
2.1.8.2. Primer design 50
Experimental section 1. Aneuploidy mechanisms in oocytes and preimplantation embryos 52
Chapter 3. Study of the expression of nine genes with roles in DNA repair in cleavage stage embryos and blastocysts harbouring DNA damage 52
3.1.1. Introduction 52
3.1.1.1. DNA damage and segmental aneuploidy in embryos 52
3.1.1.2. DNA repair mechanisms 53
3.1.1.3. Study aims 56
3.1.2. Materials and methods 57
3.1.2.1. Ethics statement 57
3.1.2.2. Experimental design and sample selection 57
3.1.2.3. Aneuploidy analysis of cleavage stage embryos 59
3.1.2.4. RNA extraction, cDNA reverse transcription and cDNA pre-amplification 59
3.1.2.5. Validation of the RT-qPCR method and gene expression analysis 61
3.1.2.6 Data analysis 62
3.1.2.7. Statistical analysis 63
3.1.3. Results 64
3.1.3.1. Ploidy status of cleavage stage embryos 64
3.1.3.2. RNA extraction and reverse transcription efficiency 65
3.1.3.3. Validation of the RT-qPCR method 66

3.1.3.4. Relative expression of the nine genes in embryos with signs of DNA damage 69
3.1.4. Discussion 75
3.1.4.1. Single blastomere chromosomal number analysis in arrested embryos 75
3.1.4.2. Development of the duplex RT-qPCR method for relative gene expression analysis 77
3.1.4.3. Level of expression of nine genes involved in DNA repair in 3- to 8-cell embryos and blastocysts with sings of DNA damage 81
3.1.4.4. Study limitations 88
3.1.4.5. Conclusion 89
Chapter 4. Aneuploidy analysis of *Atm* knocked-down mice parthenotes 91
4.1.1. Introduction 91
4.1.1.1. Female reproductive ageing 91
4.1.1.2. Meiosis in oogenesis 92
4.1.1.3. Ataxia telangiectasia mutated (ATM)-mediated DSBs repair 96
4.1.1.4. Study aims 98
4.1.2. Materials and methods 98
4.1.2.1. Ethics statement 98
4.1.2.2. Experimental design 99
4.1.2.3. Mouse cumulus-oocyte complexes (COCs) recovery, *Atm* knockdown, in vitro maturation (IVM), parthenogenesis and WGA 100
4.1.2.4. Aneuploidy analysis of parthenotes 101
4.1.2.5. Statistical analysis 102
4.1.3. Results 103
4.1.3.1. WGA and sequencing success rates 103
4.1.3.2. Aneuploidy analysis 103
4.1.3.3. Power analysis 109
4.1.4. Discussion 110
4.1.4.1. Aneuploidy frequency in the uninjected control group 111

4.1.4.2. Effect of GV transfection with small-interfering RNA (siRNA) in one-pronuclear stage (1PN) parthenotes ploidy status 115
4.1.4.3. Effect of *Atm* knockdown in 1PN parthenotes ploidy status 117
4.1.4.4. Study limitations 122
4.1.4.5. Conclusion 123

Experimental section 2. Non-invasive methods for embryo genetic testing 125

Chapter 5. Evaluation of a truly non-invasive preimplantation genetic testing (niPGT) for aneuploidies protocol using spent culture media (SCM) 125
5.1.1. Introduction 125
5.1.1.1. Non-invasive preimplantation embryo quality assessment 128
5.1.1.2. Study aims 131
5.1.2. Materials and methods 132
5.1.2.1. Ethics statement 132
5.1.2.2. Experimental design and sample preparation 132
5.1.2.3. Aneuploidy testing 133
5.1.2.4. Concordance between paired SCM and trophectoderm (TE) biopsy samples 135
5.1.2.5. Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) 136
5.1.2.6. Statistical analysis 136
5.1.3. Results 137
5.1.3.1 Libraries' concentration and NGS quality score 137
5.1.3.2. Chromosomal copy number analysis 145
5.1.4. Discussion 149
5.1.4.1. DNA informativity in SCM 149
5.1.4.2. DNA origin in SCM 152
5.1.4.3. Evaluation of the clinical application of the PG-Seq Rapid Non-Invasive PGT methodology 156
5.1.4.4. Conclusion 158

Chapter 6. Development of non-invasive methods for Y chromosome detection 160
6.1.1. Introduction 160

6.2. Project 1. Y chromosome detection from circulating placental cfDNA in early pregnancy 161
6.2.1. Introduction 161
6.2.1.1. Prenatal sex determination 161
6.2.1.2. Circulating placental cfDNA and non-invasive prenatal testing (NIPT) 162
6.2.1.3. Study aims 163
6.2.2. Materials and methods 164
6.2.2.1. Ethics statement 164
6.2.2.2. Sample characteristics and collection 164
6.2.2.3. Y chromosome assay 164
6.2.2.4. Experimental design for the Y chromosome NIPT method development 165
6.2.2.5. Plasma separation 167
6.2.2.6. cfDNA extraction 167
6.2.2.7. Multiplex PCR 167
6.2.2.8. Data analysis and concordance rates 168
6.2.3. Results 169
6.2.3.1. First phase: Multiplex PCR conditions optimisation 169
6.2.3.2. Second phase: Blood and plasma cfDNA isolation and volume downscaling 170
6.2.3.3. Third phase: Plasma separation from blood 176
6.3. Project 2. Y chromosome detection in horse BF 179
6.3.1. Introduction 179
6.3.1.1. Equine sex determination 179
6.3.1.2. Equine ARTs 180
6.3.1.3. Equine BF 182
6.3.1.4. Study aims 183
6.3.2. Materials and methods 183
6.3.2.1. Sample characteristics and collection 183
6.3.2.2. Experimental design for the method development 185
6.3.2.3. Development of the Y chromosome assay 185

6.3.2.4. Multiplex PCR 186
6.3.2.5. Data analysis and concordance rates 187
6.3.3. Results 188
6.3.3.1. First phase: Equine Y chromosome assay design 188
6.3.3.2. Second phase: Multiplex PCR optimisation 189
6.3.3.3. Third phase: Testing of optimised method on horse BF samples 198
6.3.4. Discussion 201
6.3.4.1. Development of a PCR-based method for the detection of uniquely foetal DNA in maternal blood during early pregnancy 201
6.3.4.2. Development of a PCR-based method for the detection of embryonic DNA in horse BF 208
6.3.4.3. Conclusion 213
Experimental section 3. Epigenetic avenues for embryo genetic testing 216
Chapter 7. Development of a targeted methylation analysis protocol for 46 imprinted genes 216
7.1.1. Introduction 216
7.1.1.1. Mechanisms of genomic imprinting 217
7.1.1.2. Technical challenges in CpGs methylation analysis 220
7.1.1.3. Study aims 223
7.1.2. Materials and methods 223
7.1.2.1. Experimental design for the method development 223
7.1.2.2. Buccal cell collection and isolation 231
7.1.2.3. Genomic DNA (gDNA) extraction 231
7.1.2.4. Bisulfite conversion 231
7.1.2.5. Library preparation 231
7.1.2.6. NGS 233
7.1.2.7. Data analysis 233
7.1.2.8. Statistical analysis 235
7.1.3. Results 236
7.1.3.1. First phase: Library preparation optimisation with an equimolar primer pool for PCR1 236

7.1.3.2. Second phase: Amplification level analysis per target 241
7.1.3.3. Third phase: Library preparation optimisation with a non-equimolar primer pool for PCR1 and index incorporation in primers for PCR2 244
7.1.3.4. Fourth phase: Optimisation of gDNA extraction of clumps of cells and bisulfite conversion 245
7.1.3.5. Fifth phase: Methylation level analysis of the targeted DMRs..... 248
7.1.4. Discussion 260
7.1.4.1. Amplification uniformity of the 46 targets after bisulfite conversion. 260
7.1.4.2. Methylation analysis of DMRs 263
7.1.4.3. Conclusions..... 269
Chapter 8. Final conclusions..... 272
8.1.1. Thesis summary..... 272
8.1.1.1. Introductory remarks 272
8.1.1.2. Aneuploidy mechanisms in oocytes and preimplantation embryos ... 273
8.1.1.3. Non-invasive methods for embryo genetic testing 274
8.1.1.4. Epigenetic avenues for embryo genetic testing 276
8.1.2. Study limitations..... 277
8.1.3. Future research 278
References..... 280
Appendices *305*
Appendix A 305
Appendix B 309
Appendix C 311
Appendix D 321

List of abbreviations

Abbreviation	Full form
1PN	One-pronuclear stage
2PN	Two-pronuclear stage
5mC	5-methylcytosine
aCGH	Array-comparative genomic hybridisation
ANOVA	Analysis of variance
APC	Anaphase-promoting complex
ARTs	Assisted reproductive technologies
ASRM	American Society for Reproductive Medicine
ATM	Ataxia telangiectasia mutated
BF	Blastocoel fluid
BS	Bisulfite sequencing
BSA	Bovine serum albumin
CCS	Comprehensive chromosomal screening
CDC	Centres of Disease Control and Prevention
cDNA	Complementary DNA
cfDNA	Cell-free DNA
CNV	Copy number variation
COCs	Cumulus-oocyte complexes
CpGs	Cytosine phosphate guanine sites
CVS	Chorionic villus sampling
DDR	DNA damage response
DMR	Differentially methylated region
DNA-PK	DNA protein kinase
DOP-PCR	Degenerate oligonucleotide prime-PCR
DPBS	Dulbecco's phosphate-buffered saline
DSBs	Double-strand breaks
dsDNA	Double-stranded DNA
EGA	Embryonic genome activation
eSET	Elective single embryo transfer
FISH	Fluorescent in situ hybridization
FSH	Follicle-stimulating hormone
gDMR	Germline-derived differentially methylated region
gDNA	Genomic DNA
GV	Germinal vesicle
HFEA	Human Fertilisation and Embryology Authority
HR	Homologous recombination
IACUC	Institutional Animal Care and Use Committee
ICM	Inner-cell mass
ICRs	Imprinting control regions
ICSI	Intracytoplasmic sperm injection
IGV	Integrated genome viewer
IVF	In vitro fertilisation
IVM	In vitro maturation
MALBAC	Multiple annealing and looping based amplification cycles
MDC1	Mediator of DNA damage checkpoint 1
MDA	Multiple displacement amplification

MFI	Mean fluorescence intensity
MII	Metaphase II
MRN	MRE11, RAD50 and NBS1
mRNA	Messenger RNA
MSY	Male-specific region of the Y chromosome
MZT	Maternal-to-zygotic transition
NCBI	National Centre for Biotechnology Information
NGS	Next-generation sequencing
NHEJ	Non-homologous end joining
NHS	National health service
NIH	National Institutes of Health
niPGT	Non-invasive preimplantation genetic testing
NIPT	Non-invasive prenatal testing
PB	Polar body
PBS	Phosphate-buffer saline
PCOS	Polycystic ovarian syndrome
PGCs	Primordial germ cells
PGT	Preimplantation genetic testing
PGT-A	Preimplantation genetic testing for aneuploidies
PGT-M	Preimplantation genetic testing for monogenic disorders
PGT-SR	Preimplantation genetic testing for structural rearrangements
PLKs	Polo kinases
PMSG	Pregnant mare's serum gonadotrophin
PSSC	Premature separation of sister chromatids
PVA	Polyvinyl alcohol
RNA-seq	RNA sequencing
RNAi	RNA interference
RQ	Relative quantification
RT-qPCR	Real-time quantitative PCR
SAC	Spindle assembly checkpoint
SCM	Spent culture media
siRNA	Small-interfering RNA
SNP	Single nucleotide polymorphism
STRs	Short tandem repeats
TE	Trophectoderm
T_m	Melting temperature
UMIs	Unique molecular identifiers
WGA	Whole genome amplification
WHO	World Health Organisation
G-H2AX	Modified histone H2AX

Chapter 1. Introduction

1.1.1. Introduction

Infertility affects approximately 9% of couples globally (Inhorn and Patrizio, 2015). The birth of the first in vitro fertilisation (IVF) baby, Louise Brown, in the late 1970s marked the beginning of a technological revolution to treat infertility. Since then, IVF has become a routine procedure, it is estimated that in the UK alone clinicians have performed 1.3 million IVF cycles since 1991 (HFEA, 2019).

There has been remarkable progress in assisted reproductive technologies (ARTs) in the years that followed the birth of Louise Brown. IVF is one of a suite of ARTs used for treating infertility. These technologies involve handling gametes to facilitate pregnancies and the delivery of healthy babies in previously untreatable cases. Nonetheless, despite the great scientific and technological developments of the past four and a half decades, IVF success rates remain low. The Human Fertilisation and Embryology Authority (HFEA) reported that in 2021, the average pregnancy rate per embryo transferred for patients aged between 18 and 34 years was 41%, while for patients aged between 43 and 50 years the rate falls to 6% (HFEA, 2023).

Given these low success rates, one of the most critical elements to IVF is selecting the embryos with the highest implantation potential. Advances in the field of reproductive genetics have focused on determining which embryos are the most competent. Today, ploidy status is used by clinics as a crucial indicator of embryo competency, due to the incompatibility of aneuploidy with normal development and life

(Sullivan et al., 2004, van den Berg et al., 2012, Fragouli et al., 2013). Transferring euploid embryos, using validated analysis methods, has proven to improve sustained pregnancy rates and live births (Tiegs et al., 2021).

However, there are two critical limitations to using ploidy status as the determining factor for embryo competency. Firstly, there are various constraints on the availability of these technologies designed to improve the success rates of ARTs. Although the World Health Organisation (WHO) recognises infertility as a health issue on a global scale, in low-income countries ARTs are only accessible to the wealthy, due to the high procedural costs (Asemota and Klatsky, 2015). Preimplantation genetic analysis of embryos also currently relies on the biopsy of a small number of cells at the blastocyst stage. This requires highly skilled personnel and specialist equipment, which significantly adds to the cost of IVF whilst also creating training and logistical bottlenecks that reduce patient access to this service.

Secondly, progress in the field of reproductive genetics has not only provided us with sensitive molecular methodologies for analysis and diagnostics but has also helped to better understand the genetics of infertility. However, much is still yet to be discovered about gametogenesis and the preimplantation stages of the embryo, including the molecular mechanisms behind aneuploidy. Increasing the existing knowledge about the biology of the egg, sperm and embryo can have positive implications for the success of ARTs.

Moreover, researchers are increasingly asking if other molecular processes should be used to assess embryo competency. For example, chromosomal status is an important predictive factor in embryo competency, yet still approximately one third of chromosomally normal embryo transfers fail to achieve a live birth (Forman et al., 2013,

Franasiak and Scott, 2017). Several other molecular processes are also defining elements in embryo quality. In fact, preimplantation development is characterised by a complex interplay between genetic, epigenetic and metabolic factors. Thus, exploring these avenues can provide new biomarkers for embryo viability to increase the predictive value of existing preimplantation testing methodologies for embryo competence, and therefore, the success of ARTs.

1.1.1.1 Thesis outline and aims

The overall aim of this thesis is to explore and develop new strategies for the evaluation of embryo competency, focused on the cytogenetic and epigenetic status. To achieve this, the thesis employs existing and novel molecular methodologies. The following is the thesis outline along with the specific aims for each chapter:

- Chapter 1. Introduction. This chapter introduces ARTs, factors affecting embryo viability and current methodologies routinely employed for preimplantation embryo genetic evaluation to provide a theoretical background to the research performed in the present thesis.
- Chapter 2. General materials and methods. This section describes in detail the general materials and methodologies used during this research. Each experimental chapter includes specific materials and methods with references to the general methodologies presented here.

Experimental section 1. Aneuploidy mechanisms in oocytes and preimplantation embryos:

- Chapter 3. Study of the expression of nine genes with roles in DNA repair in cleavage stage embryos and blastocysts harbouring DNA damage. This chapter investigates the capacity of preimplantation embryos to respond to DNA damage at different stages of development.
- Chapter 4. Aneuploidy analysis of *Atm* knocked-down mice parthenotes. This chapter studies the effect of *Atm* downregulation in oocyte capacity for DNA damage repair and faithful meiotic chromosome segregation.

Experimental section 2. Non-invasive methods for embryo genetic testing:

- Chapter 5. Evaluation of a truly non-invasive preimplantation genetic testing (niPGT) for aneuploidies protocol using spent culture media (SCM). This chapter studies the accuracy of niPGT in SCM when the manipulation of the embryo is minimised during embryological protocols.
- Chapter 6. Development of non-invasive methods for Y chromosome detection. This chapter is divided into two projects:
    - Project 1. Y chromosome detection from circulating placental cell-free DNA (cfDNA) in early pregnancy. A highly sensitive tool is developed that detects DNA from foetal origin in the mother's circulating cfDNA for sex determination.
    - Project 2. Y chromosome detection in horse blastocoel fluid (BF). This project aims to develop a highly sensitive tool that detects DNA of nuclear origin in equine BF for sex selection.

Experimental section 3. Epigenetic avenues for embryo genetic testing:

- Chapter 7. Development of a targeted methylation analysis protocol for 46 imprinted genes. A targeted bisulfite sequencing (BS) approach is developed in this chapter to detect methylation perturbations in 46 differentially methylated regions (DMRs) that control the expression of imprinted genes.
- Chapter 8. Final conclusions. This section draws out key learnings from the thesis, notes limitations and suggests avenues for further research.

1.1.2. Background

1.1.2.1. Infertility and assisted reproductive technologies (ARTs)

The American Society for Reproductive Medicine (ASRM) defines infertility as the failure to conceive after 12 months of appropriately timed unprotected sexual intercourse. In addition, the inability to carry a pregnancy to term can also be considered infertility (Olooto et al., 2012). Since conception is a complex biological process, infertility can be divided into several major categories, including ovulatory dysfunction, tubal occlusion, endometriosis, diminished ovarian reserve, uterine factors and male factors (Carson and Kallen, 2021). Successful implantation, and a healthy pregnancy, also depend on the competency of oocytes, sperm and subsequent embryos.

Although IVF was originally developed to treat women with tubal disease, the potential for ARTs is much broader and could include treatments for many other causes of infertility, such as advanced female reproductive age, mild male factor and unexplained. The use of ARTs is also now widespread. For example, 300,000 ARTs procedures in the USA were reported to the Centres of Disease Control and Prevention (CDC) in 2018 (Sunderam et al., 2022).

The growth in the use of ARTs, including IVF, is down to many factors. Due to socio-economic reasons many women are increasingly choosing to delay childbearing until later in life and, as such, have to freeze their oocytes for fertility preservation (Kamphuis et al., 2014). ARTs are also used by same-sex couples, women undergoing chemotherapeutic treatments who may freeze their eggs, and by couples at risk of having a child with a genetic condition, who may choose to undergo preimplantation genetic testing for monogenic disorders (PGT-M) (De Rycke and Berckmoes, 2020).

1.1.2.2. ARTs procedures

In essence, during natural conception, the sperm reaches and fertilises the oocyte inside the fallopian tube ('in vivo'), whilst in IVF, the oocyte is fertilised by the sperm outside the woman's reproductive system ('in vitro'). Being outside of the reproductive system creates opportunities to use technologies to achieve a healthy pregnancy and live birth, this includes ovarian stimulation, male factor infertility treatment and the employment of sophisticated molecular technologies. The following sub-sections describe some of the most commonly used procedures.

1.1.2.2.1. Ovarian hormonal stimulation and oocyte retrieval

Ovarian hormonal stimulation is a pharmacological treatment where hormones are used to promote a number of follicles to mature at once. It is often the first step for several fertility treatments. The aim is to collect multiple cumulus-oocyte complexes (COCs) at a single follicular aspiration (or IVF cycle) to increase the probability of obtaining good-quality embryos for transfer unto the uterus. Numerous hormonal

protocols exist; however, these mainly involve the administration of exogenous gonadotrophins to allow multi-follicular growth (Bosch et al., 2020, Ata et al., 2021).

1.1.2.2.2. Fertilisation

In conventional IVF, following retrieval, mature oocytes, commonly known as metaphase II (MII) oocytes, are placed along with capacitated sperm in a petri dish to allow for fertilisation (Figure 1).

Although IVF was initially in response to tubal disease, male factor infertility also impacts upon its success and new approaches have been developed to treat dysfunctions in sperm. Nowadays, intracytoplasmic sperm injection (ICSI) is also applied in epididymal aspiration and testicular biopsy cases where sperm is retrieved through microsurgery due to cryptozoospermia or azoospermia (Palermo et al., 2002).

Not only is ICSI used to treat male factor infertility, it is also applied as the fertilisation method for previously cryopreserved oocytes, because freezing causes the zona pellucida to harden (Porcu et al., 1997). Moreover, ICSI is preferred when embryos undergo preimplantation genetic testing (PGT) because it avoids paternal and maternal contamination originating from surplus sperm and cumulus cells since, prior to ICSI, the MII oocyte is denudated.

Although ICSI is commonly the procedure of choice for couples undergoing ARTs (Adamson et al., 2018), its utilisation is not associated with higher birth rates per cycle in non-male factor infertility (Bhattacharya et al., 2001, Wang et al., 2024). Some reports found a correlation between the employment of ICSI and an increased risk of embryonic *de novo* chromosome abnormalities (Bonduelle et al., 2002) and imprinting disorders

(DeBaun et al., 2003). However, follow-up studies with statistically representative sample sizes are lacking and the subject remains controversial.

1.1.2.2.3. Embryo culture

The commercial human embryo culture media and conditions used nowadays, collectively known as culture system, is the result of decades of optimisations with animal models (Figure 1). In the first decades of human ARTs, embryos were generally cultured in vitro up to the cleavage stage before transfer, due to the media's low performance in supporting blastocyst development (Chronopoulou and Harper, 2015). However, biochemical changes observed in the uterine and tubal fluid composition at different stages of the menstrual cycle led to the development of a sequential culture media system (Gardner et al., 1996). In this system, the cleavage stage embryo is transferred from one media formulation (G1) to another (G2) that better nourishes blastocyst development (Gardner, 1998). However, comparative studies have failed to show if one approach offers better IVF/ICSI sustained implantation rates than the other (Mantikou et al., 2013).

Although the culture system aims to adhere to the extent possible to the in vivo environment offered by the female reproductive tract, the embryo is inevitably subjected to cellular stress by mechanical and chemical factors (Swain, 2019). This is potentially significant since ARTs involve the manipulation of the gametes and embryos at uniquely vulnerable stages, particularly during crucial windows of epigenetic reprogramming (Mani and Mainigi, 2018).

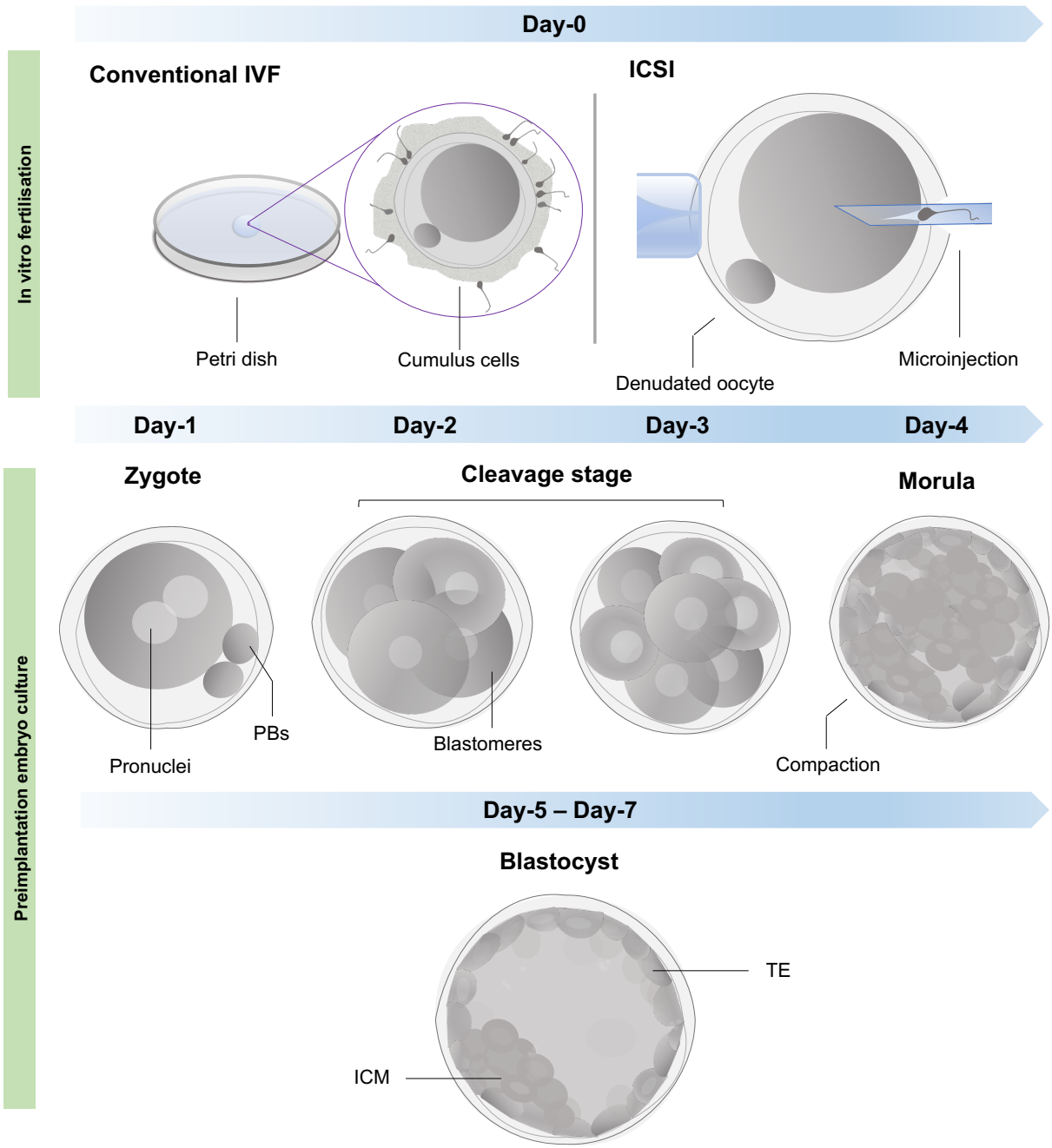


Figure 1. In vitro embryo development from fertilisation to the blastocyst stage. In conventional IVF**,** in vitro sperm capacitation evokes physiological and functional changes allowing the sperm to bind to the zona pellucida and undergo acrosome reaction and hyperactivation to penetrate and fuse with the oocyte (Bailey, 2010). With ICSI, the oocyte oolemma is breached with a micromanipulator needle to deliver the spermatozoon into the ooplasm (Palermo et al., 1996). Following fertilisation, the culture system aims to provide cells with an environment that supports their development to produce a competent embryo. Today, commercial human embryo culture media contains as many as 80 components in its formulation, including vitamins, growth factors, nutrients and amino acids (Chronopoulou and Harper, 2015).ICM, inner cell mas; ICSI, intracytoplasmic sperm injection; IVF, in vitro fertilisation; PBs, polar bodies; TE, trophectoderm.

1.1.2.2.4. Gamete and embryo cryopreservation

Oocyte and sperm cryopreservation is a key and standard procedure in ARTs. It can be used for medical and non-medical reasons, like postponing the use of gametes for programmed procedures and enabling donor sperm and oocyte banks. The cryopreservation of embryos supports the improvement of cumulative live birth rates and elective single embryo transfer (eSET) policies that aim to reduce the obstetric and neonatal complications associated with multiple gestations (Rienzi et al., 2018, Fouks and Yogev, 2022). Typically, embryos are cultured for three to six days (Figure 1) before, ideally, one selected embryo is transferred to the mother's uterus while the rest are cryopreserved for future transfers (these IVF cycles are known as frozen-thawed embryo transfer cycles). In some cases, there is no immediate transfer and all the embryos are cryopreserved for the preparation of the endometrium with natural or stimulated cycles (Glujovsky et al., 2020), after blastomere or trophectoderm (TE) biopsy for PGT or due to the patients' preference.

Currently, the most common cryopreservation technique used in oocytes and embryos is vitrification because of the superior cryosurvival rates and clinical outcomes compared to the slow-freezing technique (Rienzi et al., 2017). Vitrification is an ultra-rapid freezing method with liquid N_2 that avoids the formation of ice crystals, which can cause cell harm, with the use of high concentrations of cryoprotectants (Mukaida et al., 1998).

1.1.2.3. Preimplantation genetic testing (PGT) for embryo selection

Embryo selection refers to assessing available embryos, generally blastocysts, to select the one with the highest sustained implantation and normal development potential for transfer. Morphology evaluation remains the most common method for embryo selection because it is non-invasive and relatively easy to perform. However, good embryo morphology is not always an indicator of euploidy (Capalbo et al., 2014).

PGT, on the other hand, is considered the gold standard for the detection of embryo aneuploidy (Tiegs et al., 2021). Nonetheless, PGT can also be carried out to detect other genetic conditions. To perform PGT on embryos or oocytes requires a biopsy of blastomeres, TE or polar bodies (PBs) for the molecular analysis. The biopsy procedure is performed by penetrating the zona pellucida with laser ablation followed by aspiration with a pipette controlled with a micromanipulator (Harper et al., 2008).

1.1.2.3.1. Preimplantation genetic testing for aneuploidies (PGT-A)

Preimplantation genetic testing for aneuploidies (PGT-A) is a tool used for embryo selection based on the grounds of chromosomal status. PGT-A is mainly indicated for cases of advanced maternal reproductive age, recurrent IVF failure, recurrent miscarriage, severe male factor and eSET (Carvalho et al., 2020). The technologies behind PGT-A changed substantially since its introduction in the 1990s. Fluorescent *in situ* hybridization (FISH) was the first molecular cytogenetic technique for aneuploidy analysis of early embryos (Vidal et al., 1998). FISH allows for the interrogation of up to 12 chromosomes at the interphase nucleus from one or two blastomeres obtained from cleavage stage embryos using chromosome-specific probes. Whilst FISH can identify

numerical abnormalities (whole chromosome aneuploidies) and specific structural abnormalities (segmental aneuploidies) (Colls et al., 2009), it is unable to analyse all the chromosomes and may provide limited results because the number of embryonic cells analysed may not be representative of the cleavage stage embryo due to ploidy status mosaicism. Mosaicism is the presence of different karyotypes within the same embryo resulting from mitotic errors, most commonly during the first cleavage divisions (Fragouli et al., 2017). Furthermore, randomised control trials have concluded that there is no improvement in pregnancy and live birth rates with the use of FISH for embryo aneuploidy testing (Staessen et al., 2004, Debrock et al., 2010).

Genome-wide technologies have gradually replaced FISH over the last decades. Comprehensive chromosomal screening (CCS) captures all the chromosome pairs and different molecular techniques currently used in ARTs are based on this approach. Table 1 describes the most widely used CCS strategies. Compared to FISH, CCS provides a better representation of the whole embryo because it uses a biopsy of the TE of blastocysts, which allows for a greater number of cells (~ six) to be analysed. Nevertheless, a critical limitation of working at a single-cell level is the limited amount of material available for analysis. Therefore, generally, whole genome amplification (WGA) is required as a first step. Next-generation sequencing (NGS) has recently gained popularity in IVF due to the ability to perform parallel analyses of many samples in a single sequencing run, which reduces costs and turnaround times. Randomised control trials have found that employing CCS strategies for PGT-A can improve clinical outcomes (Forman et al., 2013, Scott et al., 2013). In addition, a recent non-selection study by Tiegs et al. (2021) demonstrated the remarkable ability of validated NGS-based PGT-A assays to predict the failure to deliver of aneuploid embryos.

Table 1. CCS strategies for PGT-A.

Technique	Description	Whole chromosome aneuploidy detection ability (Fragouli et al., 2019)	Segmental aneuploidy detection ability (Fragouli et al., 2019)
Array-comparative genomic hybridisation (aCGH)	Sample requires initial WGA. DNA is then labelled with fluorophores and hybridised to the microarray that contains spots with probes specific to different chromosomal regions. The fluorescence signals between the sample and the 'normal' reference DNA are compared to determine, with a colour signal, any losses or gains (Rubio et al., 2013).	Detects whole chromosome aneuploidies in all cells and >40-50% of mosaicism.	Detects segmental aneuploidies in all cells and >40-50% of mosaicism.
Single nucleotide polymorphism (SNP) array	Requires WGA. Similar to aCGH, DNA is then labelled with fluorophores and hybridised to an array with probes derived from specific chromosomes. Fluorescence is detected to reveal any losses or gains.	Detects whole chromosome aneuploidies in all cells but no mosaicism.	Detects segmental aneuploidies in all cells but no mosaicism.
Quantitative real time polymerase chain reaction (RT-qPCR)	Relative quantification is performed with the fluorescence emitted by assays that hybridise to specific chromosome regions through PCR (Treff et al., 2012).	Detects whole chromosome aneuploidies in all cells but no mosaicism.	Cannot detect segmental aneuploidies.
NGS	Requires WGA. Sequencing libraries are produced after fragmenting the amplified DNA followed by a low-pass sequencing in parallel. Reads are then aligned to the reference genome and counted for specific chromosomes to determine copy numbers (Fiorentino et al., 2014).	Detects whole chromosome aneuploidies in all and >20% of mosaicism.	Detects segmental aneuploidies in all cells and >20% of mosaicism.

1.1.2.3.2. Preimplantation genetic testing for monogenic disorders (PGT-M) and for structural rearrangements (PGT-SR)

In addition to PGT-A, PGT can also be carried out in a specific group of patients with a known risk of children inheriting serious genetic diseases from the parents. One

strategy is PGT-M and the other is for preimplantation genetic testing for structural rearrangements (PGT-SR). The first PGT-M case was reported by Handyside et al. (1990) who successfully determined the sex of human preimplantation embryos to avoid transmitting an X-linked genetic disorder by transferring female embryos only. As indicated by the study, PGT-M for sex determination allows for the identification of unaffected embryos before transfer. For patients, PGT strategies can offer a less stressful alternative to the more common, and widely available, approach of prenatal genetic testing (Tercyak et al., 2001). In prenatal genetic testing, through chorionic villus sampling (CVS) or amniocentesis, the genotype of the foetus is studied to determine if it is affected and the option for the patient is, therefore, termination.

PGT-M is available for couples carrying monogenic disorders, including X-linked and autosomal, recessively and dominantly inherited, as well as mitochondrial disorders involving pathogenic variants (De Rycke and Berckmoes, 2020). Couples carrying structural chromosomal rearrangements that can be prone to infertility, recurrent miscarriage and children with severe congenital disorders (Braekeleer and Dao, 1990) can be offered PGT-SR.

Multiplex PCR promptly replaced the singleplex PCR testing approach from Handyside et al. (1990) and was the gold standard for PGT-M for many years. This molecular strategy allowed for the co-amplification of short tandem repeats (STRs) used as informative polymorphic markers flanking, or linked to, the sequence containing the mutation of interest. This approach aims to identify the inherited mutant alleles and, simultaneously, allele dropout and/or contamination (Spits et al., 2006). Multiplex STRs markers near the chromosomal breakpoint are also employed as a strategy for PGT-SR (Traversa et al., 2010). Single nucleotide polymorphism (SNP) arrays, however, have

recently become popular for PGT-M. These arrays contain millions of probes distributed across the human genome to allow the genotyping of selected SNPs (which tend to be biallelic) and share the same linkage testing with multiplex PCR with STRs. Apart from the PGT-A and PGT-M application of SNP arrays, these can also be used to perform SNP haplotyping for balanced translocations or inversions (De Rycke and Berckmoes, 2020). However, NGS is currently the most widely employed PGT-SR strategy (Griffin and Ogur, 2023).

1.1.2.4. Factors affecting human embryo competency

Although ARTs mostly rely on embryo morphology and ploidy status for embryo selection, diverse elements contribute to human embryo viability and competency, such as genetic, epigenetic and transcription factors. At fertilisation, terminally differentiated cells (sperm and oocyte) fuse to trigger a highly efficient epigenetic reprogramming of the parental genomes to establish totipotency in the early embryo for subsequent normal development. RNAs and proteins of maternal origin accumulated in the egg during oogenesis are vital for the establishment of totipotency and embryonic genome activation (EGA) at the earliest stages of embryo development (Figure 2) (Zhou and Dean, 2015).

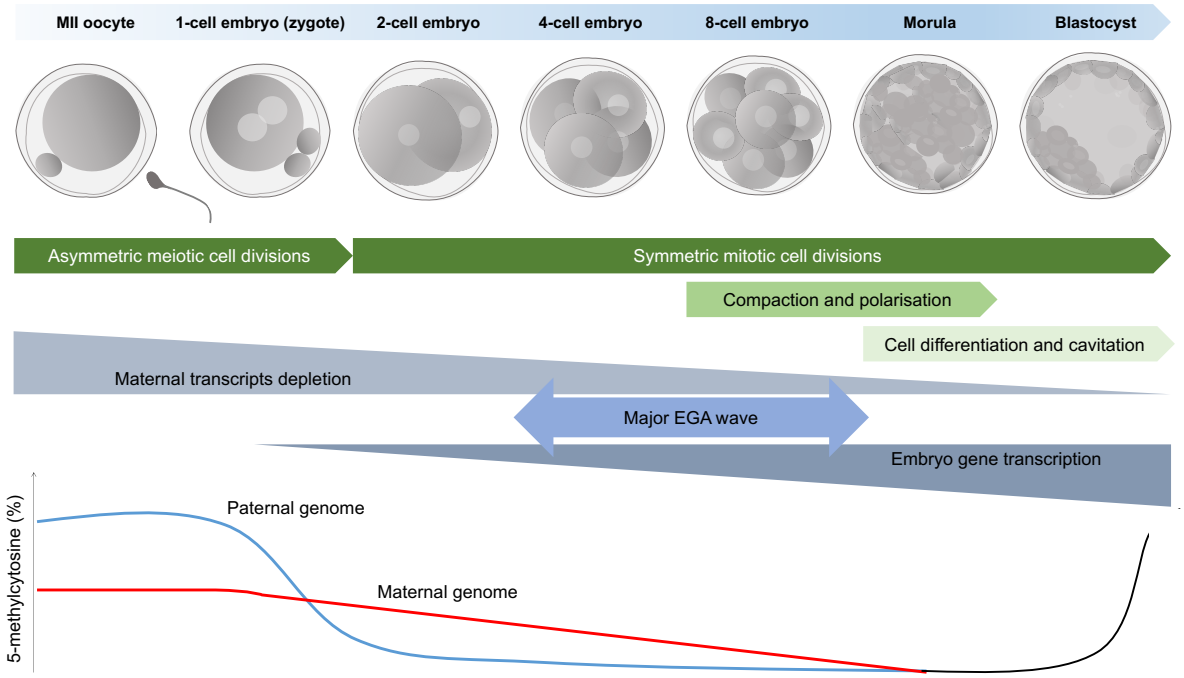


Figure 2. Human preimplantation development dynamics. Meiotic and mitotic cell divisions, oocyte-to-embryo gene expression transition dynamics and the second wave of epigenetic reprogramming. EGA, embryonic genome activation; MII, metaphase II.

1.1.2.4.1. Chromosomal ploidy status

Chromosomal abnormalities, or aneuploidy, account for approximately half of the cases of early miscarriage in humans (van den Berg et al., 2012). Aneuploidy can also lead to implantation failure and congenital birth defects (Hassold et al., 2007), such as Down syndrome (trisomy in chromosome 21), Turner syndrome (females with only one X chromosome) and Klinefelter syndrome (males with an extra X chromosome) (Tyler and Edman, 2004). ARTs have provided a helpful indication of aneuploidy rates. PGT-A findings suggest rates between 20% to 50% in in vitro-conceived embryos from young women (Baart et al., 2006, Franasiak et al., 2014). However, aneuploidy rates seem to vary at different stages of preimplantation development (Babariya et al., 2017, Pauerova et al., 2020). The earliest stages are prone to chromosomal errors that arise during meiosis or the first cell divisions (mitosis), but the preimplantation period is also

characterised by a negative selection mechanism against aneuploidy (Rubio et al., 2003). In this way, chromosomal abnormalities can be categorised based on their origin, these being meiotic and mitotic.

Aneuploidies of meiotic origin result from chromosome segregation errors during oocyte or sperm cell division (Figure 3). When gametes transmit a meiotic segregation error to the embryo all cells will carry the abnormality. Meiosis is a specialised type of cell division where the daughter cells are haploid gametes, as a result the fusion of oocyte and sperm gives rise to a zygote with a restored genome. Meiosis is, therefore, fundamental for sexual reproduction. To achieve it, cells must undergo one round of DNA replication and two rounds of asymmetric cell divisions (meiosis I and meiosis II) (Kleckner, 1996). In addition, during the prophase I stage, after DNA replication and before meiosis I is complete, paternal and maternal homologous chromosomes undergo recombination to exchange genetic material (Lam and Keeney, 2015). One major difference between the gametogenesis of oocyte and sperm is the timing of meiosis. In males, the first meiotic entry of spermatogonia occurs after puberty and throughout adult life, while oogonia enter meiosis in the foetal ovary and undergo developmental arrests to resume meiosis prior to ovulation during adult life (O'Connor, 2008). The second difference is the resulting number of haploid gametes at the end of meiosis II. Female meiosis results in a single mature oocyte (a MII oocyte) and two PBs that have no role in further development. Male meiosis, on the other hand, results in four functional sperm cells.

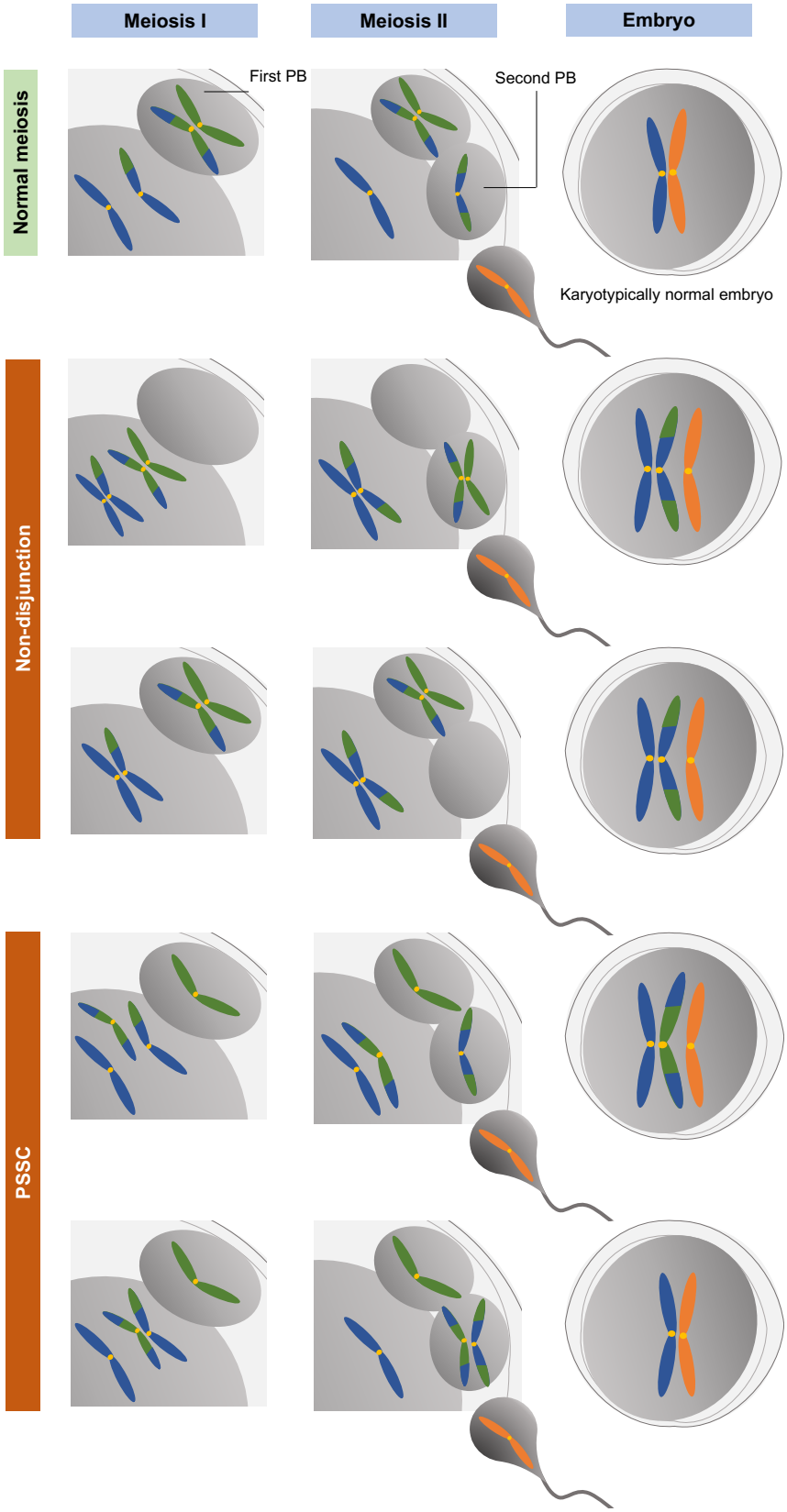


Figure 3. Mechanisms of meiotic aneuploidy in human oocytes. The type of segregation errors in female meiosis are classified into two types. One is non-disjunction, which can occur during meiosis I, where homologous chromosomes mis-segregate into the first PB, or during meiosis II, where sister chromatids segregate incorrectly into the second PB. The other segregation error type is premature separation of sister chromatids (PSSC), where, as the name indicates, individual sister chromatids mis-segregate in meiosis I due to dissociation (Charalambous et al., 2023). PB, polar body.

Although the sperm and the egg can introduce aneuploidy of meiotic origin, sperm-derived aneuploidy in embryos is relatively infrequent since aneuploidy is only observed in <5% of sperm (Petit et al., 2005). On the other hand, maternal contribution to embryo aneuploidy is highly significant. Aneuploidy is often detected in oocytes, with a strong correlation between female age and increased aneuploidy rates (Fragouli et al., 2013). Women's age-related loss of cohesion is known to be one major causing factor of premature separation of sister chromatids (PSSC) (Figure 3) (Charalambous et al., 2023). However, apart from segregation errors, chromosomal aneuploidies of meiotic origin can also originate from defective cell cycle checkpoints that safeguard the correct segregation that relies on the correct attachment of chromosomes to the meiotic spindle during cell division (Musacchio, 2015). Non-disjunction can be explained by defective microtubule dynamics and aberrant spindle assembly (Nakagawa and FitzHarris, 2017).

Aneuploidies of mitotic origin occur postzygotically (Figure 4). Fragouli et al. (2013) found chromosomal abnormalities in 74% of MII oocytes (thus, of meiotic origin) and 83% of cleavage stage embryos, an increase in aneuploidy between the two groups. Moreover, the increase in abnormality rates in cleavage stage embryos was also observed in the number of affected chromosomes per embryo (2.2 affected chromosomes in MII oocytes and 3.3 at the cleavage stage). Since spermatozoa rarely exhibit aneuploidy, it can be concluded that the increase in aneuploidy evident at the cleavage stage is of mitotic origin. Unlike meiotic aneuploidies, where all the embryonic cells present the abnormality, chromosomal abnormalities of mitotic origin introduce chromosomal mosaicism.

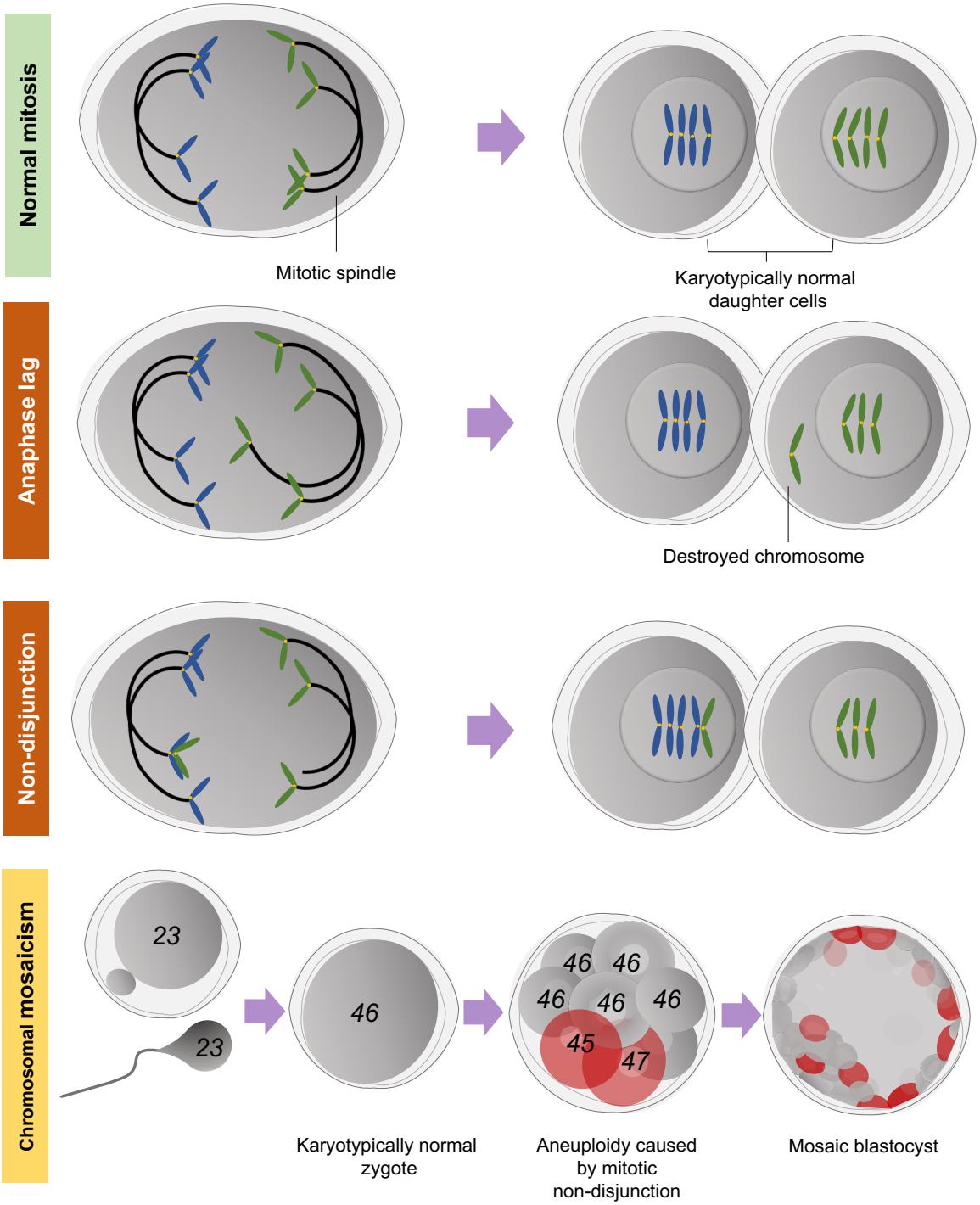


Figure 4. Mechanisms of mitotic aneuploidy in human embryos. A defective pool of maternal factors could lead to faulty cellular mechanisms and aneuploidy, including cell cycle checkpoints such as G1/S transition and inter S-phase that regulate cell division to maintain the genome integrity through safeguarding proper DNA replication. The G2/M transition checkpoint delays the progression to M-phase, if the spindle assembly checkpoint (SAC) detects a faulty kinetochore attachment to spindle microtubules that can lead to chromosome mis-segregation during mitosis (Lara-Gonzalez et al., 2021). Suboptimal cell cycle checkpoint functioning can also increase the risk of anaphase lagging where chromatids fail to attach to the mitotic spindle resulting in lost chromosomes after cell division (Mantikou et al., 2012).

Mosaicism in IVF preimplantation embryos was first reported by Delhanty et al. (1993) detected using FISH. Nowadays, NGS technologies detect even lower levels of mosaicism with high sensitivity in TE biopsies. However, there is an ongoing discussion about the management of mosaicism diagnosis for embryo selection due to the impact of technical variation during analysis, PGT-A platform-specific results and associated variable clinical outcomes when transferring mosaic embryos (Capalbo et al., 2017, Viotti et al., 2021, Babariya et al., 2023).

The origin of postzygotic aneuploidy is multifactorial. As previously mentioned, the first few mitotic divisions rely on accumulated maternal factors in the oocyte before and during EGA. It is, therefore, hypothesised that maternal proteins are vital in chromosome segregation during the first embryonic cell divisions (Figure 4) (Braude et al., 1988). Single-cell gene expression studies have shown that during the first hours of development the level of expression of a group of genes is different between euploid and aneuploid embryos (Vera-Rodriguez et al., 2015).

Another mechanism that can lead to postzygotic aneuploidy is chromosome breakages caused by unrepaired DNA double-strand breaks (DSBs). Chromosome breakages cause segmental chromosome abnormalities. While segmental aneuploidy can also be of meiotic origin, it is remarkably more common in postzygotic cell divisions (Vera-Rodríguez et al., 2016, Babariya et al., 2017). In most cell types, where these abnormalities are rare, DNA damage induces cell cycle checkpoints that pause progression to allow repair. If the damage cannot be repaired, then apoptotic pathways are triggered to eliminate the cell with the irreparable damage (Aguilera and García-Muse, 2013). However, segmental aneuploidy incidence is particularly high during the cleavage stage embryo, leading to mosaicism. It is hypothesised that the lack of

embryonic gene expression (prior to EGA) may reduce the stringency of cell cycle control that senses DNA damage (Babariya et al., 2017). Thus, maternal factors, including messenger RNAs (mRNAs) and proteins, are also defining elements in human embryo competency.

1.1.2.4.2. Transcriptional factors

Comprehensive gene expression studies with human embryos have mapped the transcription levels of thousands of genes to better understand the mechanisms of early embryo cell differentiation (Petropoulos et al., 2016). Findings show that the expression of lineage-specific genes is tightly regulated and programmed in normal preimplantation development. Recent studies using high-resolution single-cell RNA sequencing (RNA-seq) have also discovered that stringent dynamics triggered by fertilisation regulate the oocyte-to-embryo gene expression transition (Figure 2), where genes are moderately up-regulated with roles in cell-cycle progression and chromosome maintenance (Asami et al., 2022). Advances in RNA sequencing of single cells have, therefore, contributed to identifying candidate genes whose levels of expression could be associated with embryo competency (Groff et al., 2019). Thus, the analysis of transcript abundance (mRNA) in IVF preimplantation embryos has recently been proposed as a biomarker for embryo viability to be used in routine embryo selection criteria (morphology and ploidy status).

1.1.2.4.3. Epigenetic factors

Female infertility aetiologies can sometimes explain the failure of the euploid embryo to implant. For example, due to endometriosis (Simon et al., 1994), polycystic

ovarian syndrome (PCOS) (Schulte et al., 2015), the mother's immunity (Franasiak and Scott, 2017) and endometrium receptivity (Craciunas et al., 2019). However, research is increasingly focusing on epigenetic embryological factors as an explanation for implantation failure and miscarriage.

Epigenetics refers to the meiotically and mitotically heritable information that exists 'above' the encoded DNA. It helps to control gene expression without altering the nucleotide sequence, thereby, allowing for differences in genome function between different cell types (Greally, 2018). Embryo development is characterised by the tight dynamic control of gene expression patterns, or programs, that permit the full totipotency of the newly fertilised zygote at the very initial stages. Embryonic stem cells, with a lower degree of potency, then direct to new fates upon differentiation into different cell lineages. The variation in potency of embryonic cells is achieved through control exerted by histone modifications, DNA methylation and chromatin remodelling, all of which comprise epigenetic modifications (Mohn and Schübeler, 2009).

DNA methylation, which is fundamental for maintaining cell identity, is an epigenetic mark where a methyl group is transferred by a DNA methyltransferase to the 5-position of a cytosine resulting in a 5-methylcytosine (5mC) molecule. This mark must be faithfully inherited through cell division. In addition to a vital role in transcriptional regulation, DNA methylation is also fundamental for genomic imprinting (Zeng and Chen, 2019). In genomic imprinting, a subset of genes, known as imprinted genes, are expressed in a parent-of-origin dependent manner achieved by differential epigenetic marking of the maternal or paternal alleles carried out by DNA methylation.

Imprinted genes have important roles in human development, and thus, DNA methylation is critical in embryo development (Monk et al., 2019). DNA methylation

marks must be inherited for maintenance through somatic cell division. However, the epigenome, unlike the genome, is dynamic and needs to be reprogrammed after fertilisation to allow for a state of totipotency in the mammalian zygote. In fact, two waves of global DNA demethylation and remethylation exist in the mammalian life cycle (Figure 5).

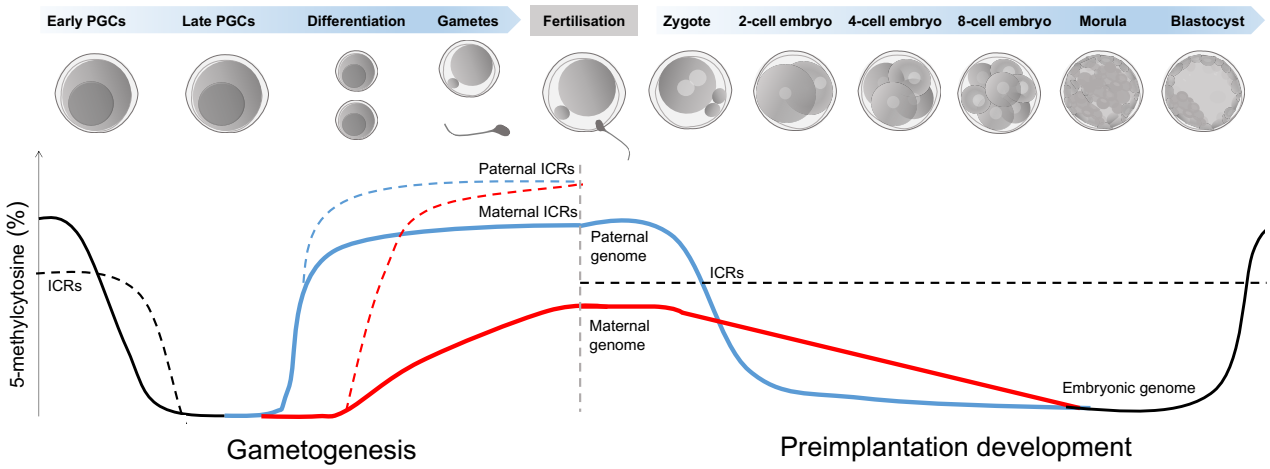


Figure 5. Two waves of DNA methylation reprogramming. The first wave occurs during the earliest stages of embryogenesis, right after fertilisation, and the other before and during gametogenesis. In early preimplantation embryogenesis, most methylation marks inherited from the highly differentiated sperm and oocyte are erased (and later re-established), excluding the marks located in imprinting control regions (ICRs). The maternal genome undergoes passive demethylation, whilst the paternal genome undergoes active demethylation, making the removal of marks from the paternal genome more abrupt (Zeng and Chen, 2019). Genomic imprints must maintain their monoallelic methylation marks throughout their generational lifespan. Thus, ICRs, also known as germline-derived differentially methylated regions (gDMRs), must be protected from global DNA demethylation during embryo somatic cell reprogramming. However, primordial germ cells (PGCs), the embryonic precursors of gametes, require epigenome reprogramming to enter gametogenesis, giving rise to the second wave of DNA demethylation and remethylation. An important distinction of PGCs epigenome reprogramming is the imprint erasure carried out through passive demethylation. The gDMRs are then remethylated following the corresponding parent-of-origin marks (Monk et al., 2019). Apart from gDMRs, differentially methylated regions (DMRs) can also be secondary. Secondary DMRs, unlike gDMRs, acquire their parent-of-origin DNA methylation post-implantation (Bhogal et al., 2004). Figure modified from Monk et al. (2019).

Aberrant DNA methylation patterns, also known as epimutations, have been associated with developmental disorders, congenital conditions and, in some instances, cancer (Peters, 2014, Hamidi et al., 2015). Imprinting disorders occur when the imprint erasure process is erroneous or when the establishment of imprints during

gametogenesis is suboptimal or fails, sometimes as a result of environmental factors (Monk et al., 2019). This is where ARTs may play a role in affecting human embryo competency.

ARTs are implemented during the developmental stages when the epigenome is particularly vulnerable. ARTs involve the manipulation of gametes and embryos with procedures such as ovarian hormonal stimulation, ICSI and embryo culture. In fact, it is through ARTs that we have learnt more about the influence of the environment on the epigenome, with reports of rare imprinting disorders found in children conceived after ARTs, including Angelman and Beckwith-Widemann syndromes (Cox et al., 2002, DeBaun et al., 2003, Cortessis et al., 2018). These are frequently caused by hypomethylation or hypermethylation of differentially methylated regions (DMRs) (Lazaraviciute et al., 2014). Moreover, abnormal expression of genes regulated by imprinting control regions (ICRs) can also affect embryo development and influence the success of implantation (Wallach et al., 1994, Tveden-Nyborg et al., 2008).

Having introduced ARTs and discussed the various factors affecting embryo viability and current methodologies employed for preimplantation embryo genetic evaluation, this thesis now turns to the materials and methods before discussing the experimental studies and data produced.

Chapter 2. General materials and methods

This chapter presents a detailed description of the general materials and methods used during the course of this thesis. Each experimental chapter describes the methodologies specific to it. The overall standard structure of the methods is shown in Figure 6. However, this was not strictly followed in every experimental chapter.

For all processes described, the manufacturer's guidelines were followed, unless stated otherwise. All biological material was handled in a clean room or in a UV-irradiated hood to avoid contamination. The same was followed for nucleic acid extractions and master mix preparations.

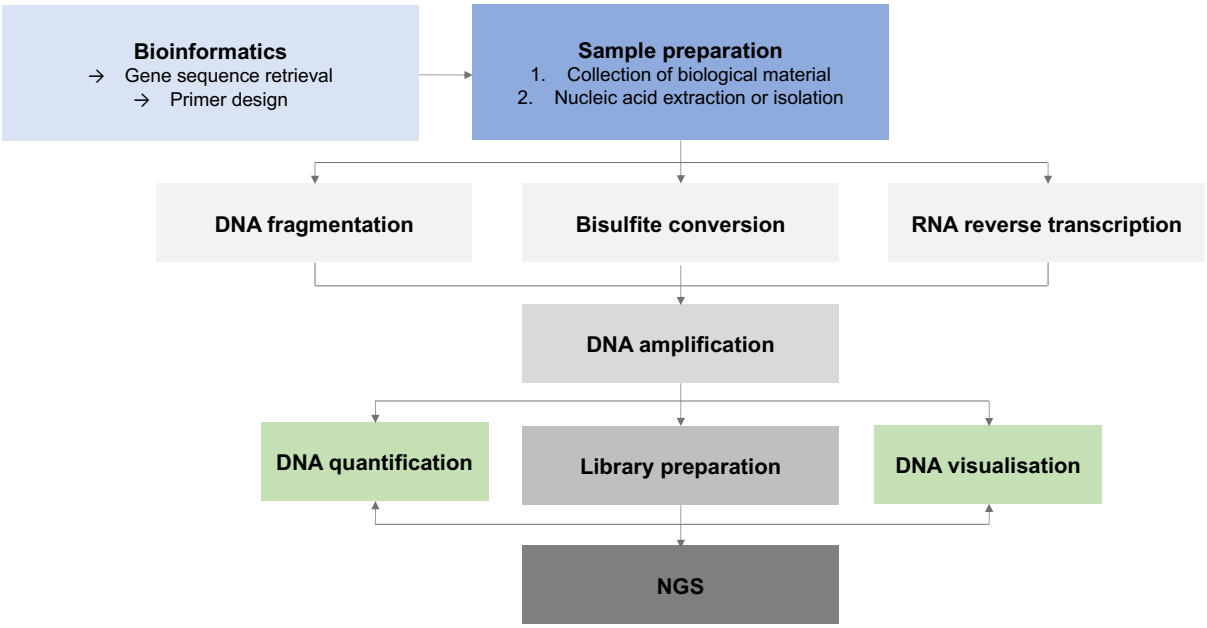


Figure 6. Standard methods structure followed in the present thesis. NGS; next generation sequencing.

2.1.1. Sample preparation

2.1.1.1. Buccal cell collection and isolation

Single-cell bisulfite sequencing (BS) methodology optimisation was conducted with clumps of buccal cells in lieu of TE biopsies. The process for obtaining samples was as follows. Individual's mouths were swabbed for one minute using the Catch All Sample Collection kit (Epicentre, USA). Until cell isolation, swabs were stored at 4°C in sealed sample collection tubes. The swab was then transferred to a sterile 5 ml Falcon conical tube (Corning, USA), containing 2 ml of Dulbecco's phosphate-buffered saline (DPBS) solution, and rotated against the wall of the tube to suspend the cells in the DPBS. Approximately 10 µl of the cell suspension were transferred to a sterile petri dish, with 5 µl and 10 µl drops of wash buffer (a mixture of phosphate-buffer saline (PBS) solution (Sigma-Aldrich, USA) and 0.1 % polyvinyl alcohol (PVA) (Sigma-Aldrich, USA)). Under an inverted microscope, the cell suspension was mixed by mouth pipetting in a 5 µl drop. It was subsequently transferred in reduced quantities to fresh 5 µl drops until clumps were observed to have from eight to nine cells. The 10 µl drops were used to wash the pipette tip between cell transfers to avoid contamination. Approximately 2.50 µl of the last drop were transferred to a 0.2 ml tube, centrifuged and stored at -80°C until DNA extraction.

2.1.1.2. DNA extraction from clumps of cells

Twenty mg/ml of Proteinase K (Qiagen, Germany) were diluted to make a 150 µg/ml concentration with nuclease-free water. Five ml of the dilution were transferred to the tube containing the clumps of cells. The cells were then incubated in a Mastercycler X50 thermocycler (Eppendorf, Germany), using the thermal protocol outlined in Table 2.

Table 2. Thermal protocol for DNA extraction with Proteinase K.

Step	Temperature	Time
First incubation	37°C	60 minutes
Second incubation	65°C	10 minutes
Inactivation	95°C	10 minutes

2.1.1.3. DNA and cell-free DNA (cfDNA) isolation from blood

Two kits were used in this thesis for the isolation of cfDNA from blood. They are both described below.

2.1.1.3.1. QIAmp DNA Micro Kit (Qiagen, Germany) – Protocol: Isolation of Genomic DNA from Small Volumes of Blood

Two hundred µl of blood, or 100 µl of plasma, were transferred to a 1.5 ml microcentrifuge tube with 10 µl of 20 mg/µl proteinase K and 100 µl of Buffer AL and vortexed. The mixture was incubated for 10 minutes at 56 °C for lysis. After incubation, 50 µl of 100% ethanol were added to the tube, which was mixed and incubated for a further three minutes at room temperature. The lysate was then transferred to a QIAmp MinElute column and centrifuged for 1 minute at 6,000 x g. After adding 500 µl of Buffer AW1 to the column, it was centrifuged for 1 minute at 6,000 x g and then for 3 minutes at 20,000 x g to dry the membrane. Then, 13 µl of Buffer AE were added to the centre of the membrane and incubated for 1 minute at room temperature. The column was centrifuged for a final time in a clean 1.5 microcentrifuge tube to elute the DNA. The DNA was immediately processed or stored at -20 °C for later use.

2.1.1.3.2. MagMAX Cell-Free DNA Isolation Kit (Applied Biosystems, USA) – Protocol: Manual Isolation of cfDNA from small volumes of plasma and serum samples

Components found in Table 3 were added in a 1.5 microcentrifuge tube, mixed and incubated for 20 minutes at 60 °C.

Table 3. MagMAX Cell Free DNA Isolation Kit table of components for the cell lysis step.

Component	Plasma volume		
	100 µl	**200 µl**	**400 µl**
Proteinase K, 20 mg/µl	2 µl	3 µl	6 µl
Plasma sample	100 µl	200 µl	400 µl
SDS, 20% Solution (Sigma-Aldrich, USA)	5 µl	10 µl	20 µl
Total volume	107 µl	213 µl	426 µl

The Binding Solution/Beads mix was prepared with the components found in Table 4, according to the plasma starting volume.

Table 4. MagMAX Cell Free DNA Isolation Kit table of components for the Binding Solution/Beads mix.

Component	Plasma volume		
	100 µl	**200 µl**	**400 µl**
MagMax Cell Free DNA Lysis/Binding Solution	150 µl	300 µl	500 µl
MagMax Cell Free DNA Magnetic Beads	5 µl		10 µl
Total volume	155 µl	305 µl	510 µl

After the incubation, the tube containing the plasma was cooled in ice for 5 minutes and the correspondent volume of Binding Solution/Beads mix was added. The mix was placed in a shaker for 10 minutes and then in a DynaMag Magnet (Thermo Fisher Scientific, USA) for 5 minutes. The resulting supernatant was discarded with a pipette. The magnetic beads were resuspended with 500 µl of MagMax Cell-Free DNA Wash Solution, placed back in the magnet for 2 minutes and the supernatant was discarded. The beads were resuspended for a second time with 500 µl of 80% ethanol and placed in

the magnet for 1 minute, the supernatant was discarded again. The ethanol wash was repeated for a second time.

The beads were then let to air dry and 13 µl of MagMAX Cell-Free DNA Elution Solution were added and the tube was placed in a shaker for 5 minutes. The mixture was placed back in the magnet for 2 minutes and the supernatant was recovered containing the purified cfDNA for immediate processing or stored at -20 °C for later use.

2.1.1.4. RNA extraction from embryos and reverse transcription

The RNA extraction from embryos, and reverse transcription, were performed using the TaqMan PreAmp Cells-to-Ct kit (Applied Biosystems, USA). Cells were first lysed. Due to the small number of cells and the small suspension volumes, the lysis solution (DNAse I diluted into Lysis Solution in a 1:100 dilution factor) added to the embryos was decreased from the 50 µl suggested by the manufacturer to 10 µl. Cells were incubated with the lysis solution for 5 minutes at room temperature. The DNAse I was then deactivated by adding 1 µl of Stop Solution and incubated for 2 minutes at room temperature. The reverse transcription master mix was assembled (Table 5) and added to the lysate. The components for the reverse transcription reaction were also modified from their original volumes and the lysate increased from 10 µl to 11 µl, as a result of the initial modifications to the protocol. The reverse transcription reaction was performed in a thermocycler. The program is described in Table 6. The products of the reverse transcription were stored at -80 °C for later use.

Table 5. Components for the reverse transcription reaction. Volumes were modified from the manufacturer's standard protocol.

Component	Volume (µl)
2X RT Buffer	11.11
20X RT Enzyme Mix	1.11
Nuclease-free water	0
Lysate	~ 11
Total volume	~ 23.22

Table 6. Thermal protocol for the reverse transcription.

Step	Temperature	Time	Number of cycles
Reverse transcription	37 °C	60 minutes	1
Inactivation	95 °C	5 minutes	1
Hold	4 °C	∞	1

2.1.2. DNA fragmentation

The NEBNext dsDNA Fragmentase (New England Biolabs, USA) was used to digest double-stranded DNA (dsDNA) into a desired size. The components in Table 7 were transferred, in the order presented, to a 0.2 ml tube and mixed.

Table 7. Components for the DNA digestion with NEBNext dsDNA Fragmentase.

Component	Volume (µl)
DNA (5 ng - 3 µg)	1 – 16
10X Fragmentase Reaction Buffer v2	2
Nuclease-free water	Variable
dsDNA Fragmentase	2
Total volume	20

The mix was then incubated at 37 °C for the recommended time found in Table 8. The reaction was stopped by adding 5 µl of 0.50 M EDTA.

Table 8. Incubation time and desired fragment sizes for the NEBNext dsDNA Fragmentase.

Desired fragment size (bp)	Incubation time (minutes)
50 – 200	23 – 35
200 – 1,000	15 – 25
1,000 – 2,000	10 – 15

2.1.3. DNA amplification

2.1.3.1. Multiplex PCR

Simultaneous amplification of multiple loci was enabled by multiplex PCR. The kit used for all multiplex reactions was Qiagen Multiplex PCR Kit (Qiagen, Germany). The manufacturer's suggested table of components is shown in Table 9. However, after conducting optimisation, the DNA input, the primer concentration and the total reaction volume used were different to the manufacturer's suggested components. If the total reaction volume was decreased, the components were decreased proportionally. The manufacturer's suggested PCR conditions are shown in Table 10. However, these were also adjusted for optimisation purposes. Reactions were carried out in a thermocycler.

Table 9. Components for the Qiagen Multiplex PCR Kit amplification reaction. The volumes shown are for single reactions. The master mixes contained enough volume for the required number of reactions, including controls and a 10% surplus.

Component	Volume (µl)
2X Qiagen Multiplex PCR Master Mix	25
10X primer mix, 2 µM each primer (final concentration: 0.2 µM)	5
Nuclease-free water	Variable
DNA	Variable
Total volume	50

Table 10. Thermal protocol for the Qiagen Multiplex PCR Kit amplification reaction.

Step	Temperature	Time	Number of cycles
Initial denaturation	95 °C	15 minutes	1
Denaturation	94 °C	30 seconds	30-45
Annealing	57-63 °C	90 seconds	
Extension	72 °C	90 seconds	
Final extension	72 °C	10 minutes	1
Hold	4 °C	∞	1

Specifications about the oligonucleotides used, modified multiplex PCR conditions and components are described in the Materials and methods section from specific chapters.

2.1.3.2. Quantitative real time PCR (RT-qPCR)

The quantitative real time PCR (RT-qPCR) amplification was carried out using the TaqMan Gene Expression Master Mix and TaqMan Gene Expression Assays (Applied Biosystems, USA). The RT-qPCR master mix was assembled (Table 11) and added to the complementary DNA (cDNA). The amplification reactions were run in a 384-well plate and in the Thermo ABI ViiA 7 Real-Time PCR (Applied Biosystems, USA). The program is described in Table 12.

Table 11. Components for the RT-qPCR. List of components and volumes obtained from the TaqMan Gene Expression Master Mix protocol. *Uracil-N glycosylase (UNG). The RT-qPCR in this thesis did not include UNG. However, this step was kept in the program.

Component	Volume (µl)	
	Singleplex	Duplex
TaqMan Gene Expression Master Mix [2X]	5	5
TaqMan Gene Expression Assay (target of interest) [20X]	0.50	0.50
TaqMan Gene Expression Assay (endogenous control) [20X]	0	0.50
Diluted cDNA	2.50	2.50
Nuclease-free water	2	1.50
Total volume	10	10

Table 12. Thermal protocol for the RT-qPCR with the TaqMan Gene Expression Master Mix.

Step	Temperature	Time	Number of cycles
UNG incubation*	50 °C	2 minutes	1
Enzyme activation	95 °C	10 minutes	1
Denaturation	95 °C	15 seconds	40
Annealing/extension	60 °C	60 seconds	
Hold	4 °C	∞	1

2.1.3.3. Pre-amplification of complementary DNA (cDNA)

The reverse-transcribed cDNA was pre-amplified targeting the genes of interest using the TaqMan PreAmp Master Mix and TaqMan Gene Expression Assays (Applied Biosystems, USA) in multiplex. The pre-amplification master mix was prepared (Table 13) and added to the cDNA. The reaction was run in a thermocycler with the program described in Table 14. Fourteen cycles were used for the pre-amplification, due to the limited amount of starting cDNA. The resulting products were stored at -80°C.

Table 13. Components for the pre-amplification reaction.

Component	Volume (µl)
TaqMan PreAmp Master Mix [2X]	25
Pooled TaqMan Gene Expression Assays [0.2X each]	12.50
cDNA	12.50
Total volume	50

Table 14. Thermal protocol for the pre-amplification reaction with the TaqMan PreAmp Master Mix. According to the manufacturer, X14 pre-amplification cycles is suitable for when the sample is extremely limited (single-cell level).

Step	Temperature	Time	Number of cycles
Enzyme activation	95 °C	10 minutes	1
Denaturation	95 °C	15 seconds	14
Annealing/extension	60 °C	4 minutes	
Enzyme inactivation	99 °C	10 minutes	1
Hold	4 °C	∞	1

2.1.3.4. Whole genome amplification (WGA)

Throughout this thesis, WGA was performed with SurePlex amplification (Illumia, USA), unless otherwise specified. The first step was cell lysis (8 µl of lysis solution were added to approximately 1 µl of cell suspension), followed by incubation and denaturation in a thermocycler. Immediately after the first thermal protocol, 5 µl of pre-amplification solution were added to the samples and pre-amplified in the thermocycler. A final amplification step was performed after adding 60 µl of amplification solution to the samples, followed by a final thermal protocol (Table 15 and Table 16). The resulting products were stored at 4 °C.

Table 15. Components for the SurePlex amplification. The volumes shown are for single reactions. The master mixes contained enough volume for the required number of reactions, including controls and a 10% surplus.

1. Cell lysis	
Reagent	**Volume (µl)**
Extraction Enzyme Dilution Buffer	4.80
Cell Extraction Buffer	3
Cell Extraction Enzyme	0.20
2. Pre-amplification	
SurePlex Pre-amp Buffer	4.80
SurePlex Pre-amp Enzyme	0.20
3. Final amplification	
SurePlex Amplification Buffer	25
SurePlex Amplification Enzyme	0.80
Nuclease-free water	34.20

Table 16. Thermal protocols for the three-step SurePlex amplification.

1. Cell lysis			
Step	**Temperature**	**Time**	**Number of cycles**
Incubation	75 °C	10 minutes	1
Denaturation	95 °C	4 minutes	1
Hold	10 °C	∞	1
2. Pre-amplification			
Enzyme activation	95 °C	2 minutes	1
Pre-amplification	95 °C	15 seconds	12
	15 °C	50 seconds	
	25 °C	40 seconds	
	35 °C	30 seconds	
	65 °C	40 seconds	
	75 °C	40 seconds	
Hold	4 °C	∞	1
3. Final amplification			
Enzyme activation	95 °C	2 minutes	1
PCR	95 °C	15 seconds	14
	65 °C	1 minute	
	75 °C	1 minute	
Hold	4 °C	∞	1

2.1.4. DNA visualisation

2.1.4.1. Agarose gel electrophoresis

Amplified products were visualised and analysed in an agarose gel electrophoresis. Five µl of DNA were mixed with 1 µl of BlueJuice Gel Loading Buffer (Invitrogen, UK) and run in an agarose gel (agarose (Invitrogen, USA), 100 ml of 1X TBE buffer (Thermo Fisher Scientific, USA) and 10 µl of GelRed (Biotium, USA)), along with 5 µl of 100 bp Low Scale Ladder (Fisher Scientific, USA). The agarose gel percentage, the voltage and the run time varied throughout this thesis depending on the purpose of the agarose gel electrophoresis and the characteristics of the nucleic acids being analysed. Once the run finished, the gel was visualised on a UV transilluminator.

2.1.4.2. TapeStation (Applied Biosystems, USA)

Nucleic acids were also visualised, and in some instances, DNA concentrations were measured using the TapeStation System (Applied Biosystems, USA). Depending on the reagents used, the ladder and the samples were prepared following the volumes in Table 17. They were mixed and loaded in the TapeStation 4200 instrument for analysis.

Table 17. Compilation of the components used for the different TapeStation reagents.

Reagents	Ladder		Sample	
	Ladder (µl)	Buffer (µl)	Sample (µl)	Buffer (µl)
Genomic DNA	1	10	1	10
D100	1	3	1	3
D100 High Sensitivity	2	2	2	2

2.1.5. Measurement of DNA concentrations

Concentrations of DNA that ranged from 10 pg/µl to 100 ng/µl were measured using the Qubit 1X dsDNA HS Assay Kit (Invitrogen, USA) and the Qubit 4 Fluorometer (Invitrogen, USA). Serial dilutions were performed in samples with concentrations above 100 ng/ul. Standard 1 and Standard 2 for calibration were prepared in 0.5 ml tubes by adding 10 µl of the standard and 190 µl of Working Solution. One to 20 µl were added to samples with Working Solution to make up a total volume of 200 µl. Standards and samples were incubated for 2 minutes and then quantified in the fluorometer.

2.1.6. DNA bisulfite conversion

Extracted genomic DNA (gDNA) from clumps of cells was subjected to bisulfite conversion of unmethylated cytosines with two commercial kits. The single-cell level amounts of gDNA required the following protocol adjustments. A single pipette tip was used to transfer the Proteinase K-digested cells to the 0.2 ml tube intended for the

bisulfite conversion. After the initial transfer of the cells, the tube was washed with reagents, as part of the first step of the bisulfite reaction, and then centrifuged to recover the material to the best extent possible. All the pipetting was performed on top of the reagents contained in the tube to avoid any binding of DNA to the walls of the pipette. Finally, the centrifugation time for the elution of DNA from the spin column was extended for each of the protocols to maximise the recovery of bisulfite-converted DNA.

2.1.6.1. EpiTect Bisulfite Conversion kit (Qiagen, Germany) – Protocol: 'Sodium bisulfite conversion of unmethylated cytosines in small amounts of fragmented DNA'

The components found in Table 18 were added to a 0.2 ml tube and incubated in a thermocycler using the thermal protocol in Table 19.

Table 18. Components for the Qiagen EpiTect Bisulfite Conversion reaction. The volumes shown are for single reactions. The master mixes contained enough volume for the required number of reactions, including controls and a 10% surplus.

Component	Volume (µl)
DNA	~8
Nuclease-free water	12
Bisulfite Mix	85
DNA Protect Buffer	35
Total volume	140

Table 19. Thermal protocol for the Qiagen EpiTect Bisulfite Conversion reaction.

Step	Temperature	Time
First denaturation	95 °C	5 minutes
First incubation	60 °C	25 minutes
Second denaturation	95 °C	5 minutes
Second incubation	60 °C	85 minutes
Second denaturation	95 °C	5 minutes
Second incubation	60 °C	175 minutes
Hold	20 °C	∞

The 0.2 ml tubes were then centrifuged and the bisulfite-converted DNA was transferred to a 1.5 ml Eppendorf tube with 310 µl of Buffer BL containing 10 µg/ml carrier RNA for DNA clean-up. For this purpose, 250 µl of 100% ethanol were added and mixed. The mix was then transferred to a spin column and centrifuged in a microcentrifuge at maximum speed for 1 minute. The flow-through was discarded and 500 µl of Buffer BW were added to the column to be centrifuged at maximum speed for 1 minute. The flow-through was discarded, 500 µl of Buffer BD were added into the column and incubated at room temperature for 15 minutes. After the incubation, the columns were centrifuged for a third time at maximum speed for 1 minute. The flow-through was discarded and 500 µl of Buffer BW were added into the column and centrifuged at maximum speed for 1 minute. The flow-through was discarded and, after adding Buffer BW, centrifugation was repeated once. The spin columns were then transferred to a 2 ml Eppendorf tube (Eppendorf, Germany) and centrifuged at maximum speed for 1 minute to remove any residual reagents. Finally, the spin column was placed in a 1.5 ml Eppendorf tube and 20 µl of Buffer EB were added. The 1.5 ml tubes were centrifuged for 2 minutes at 15,000 x g. The eluted purified bisulfite-converted DNA was recovered from the bottom of the tube, placed in a 0.2 ml tube and stored at -80°C until library preparation.

2.1.6.2. Imprint Modification kit (Sigma-Aldrich, USA) – Protocol: 'Two-Step Modification Procedure'

The extracted DNA was added to a 1.5 ml Eppendorf tube and the volume adjusted to 24 µl by adding 0.5 mg/ml of bovine serum albumin (BSA) (Sigma-Aldrich, USA). One µl of Balance Solution was also added to the tube and the mix was incubated

at 37 °C for 10 minutes. Then, 1.10 ml of DNA Modification Solution was mixed in a vial with DNA Modification Powder, and 125 µl were added to the solution and incubated at 65 °C for 90 minutes. For DNA clean-up, 300 µl of Capture Solution were added to the centre of the spin column and left for the column to absorb the reagent for 1 minute. The DNA mix was then transferred to the column placed in a Capless Collection Tube and centrifuged at 12,000 x g for 20 seconds. The flow-through was discarded and 200 µl of ethanol-diluted Cleaning Solution were added to the column to then centrifuge it a second time at 12,000 x g for 20 seconds. The flow-through was discarded. Fifty µl of Balance-ethanol Wash Solution (10 µl of Balance Solution and 1.10 ml of 90% ethanol) were added to the column and incubated at room temperature for 8 minutes and then centrifuged at 12,000 x g for 20 seconds. The flow-through was discarded again and 200 µl of 90% ethanol were added to the column and centrifuged at 12,000 x g for 20 seconds. After discarding the flow-through, the centrifugation with 90% ethanol was repeated. The column was then transferred to a 1.5 ml Collection tube and 20 µl of Elution Solution were added. The column was incubated at room temperature for 1 minute and then centrifuged at 12,000 x g for 40 seconds. The eluted purified bisulfite-converted DNA was recovered from the bottom of the tube, placed in a 0.2 ml tube and stored at -80°C until library preparation.

2.1.7. Library preparation and next-generation sequencing (NGS)

2.1.7.1. VeriSeq (Illumina, USA)

This protocol and pipeline have been validated for whole chromosomal and segmental aneuploidies detection in TE biopsies (Fragouli et al., 2017). The SurePlex method for WGA is the first step for single-cell aneuploidy screening in the VeriSeq

platform (Illumina, USA). It began with fragmenting 0.20 ng/µl (1 ng total) of diluted SurePlex products and attaching adapters to the fragments in a tagmentation reaction at 55 °C for 5 minutes in a thermocycler. Index sequences were then introduced through a PCR amplification to allow dual-index sequencing (Table 20 and Table 21). A magnetic bead-based purification was carried out on the resulting indexed sequences using AMPure XP beads (Beckman Coulter, USA) in a 1:1 dilution factor. The beads were washed away with 80% ethanol and the DNA was eluted in the Resuspension Buffer. This protocol then proceeded to normalise the purified libraries to equal quantities of DNA per sample. The normalised libraries were then pooled, denatured and finally sequenced using the MiSeq Reagent Kit V3 (300 cycles) (Illumina, USA) on the MiSeq Sequencing System (Illumina, USA). The sequencing results were analysed with BlueFuse Multi Software (Illumina, USA).

Table 20. Components for the VeriSeq tagmentation and PCR amplification for dual indexing. The volumes shown are for single reactions.

1. Tagmentation	
Reagent	**Volume (µl)**
Amplicon Tagment Mix	5
Tagment DNA Buffer	10
DNA sample ((0.2 ng/ µl)	5
2. PCR amplification for dual indexing	
Nextera PCR Master Mix	15
Index 1 (i7) primers	5
Index 3 (i5) primers	5

Table 21. Thermal program for the VeriSeq PCR amplification for dual indexing.

Step	Temperature	Time	Number of cycles
Enzyme activation	72 °C	3 minutes	1
	95 °C	30 seconds	1
PCR	95 °C	10 seconds	12
	55 °C	30 seconds	
	72 °C	30 seconds	
Final elongation	72 °C	5 minutes	1
Hold	10 °C	∞	1

2.1.7.2. Oxford Nanopore MinION (Oxford Nanopore, UK)

The flow cell used for the NGS analysis was the MinION (FLO-MIND106D). The library protocol followed was the Ligation Sequencing gDNA – PCR Barcoding (SQK-LSK109), along with the PCR Barcoding Expansion Pack 1 – 96 (EXP-PBC096) for barcoding samples.

The concentration of the SurePlex WGA products was first measured in a 1:10 dilution factor (DNA:nuclease-free water) with the Qubit 1X dsDNA HS kit to confirm optimal DNA amounts of starting template. The first step of the library preparation involved an initial end-prepping of the gDNA fragments produced after WGS. The table of components for the initial end-prep reaction and the volumes added for each sample are found in Table 22(1). Some volumes were modified from the manufacturer's protocol to economise the use of reagents. Components were mixed in 0.2 ml tubes and incubated for 5 minutes at 20 °C and 5 minutes at 65 °C in a thermocycler. The end-prepped DNA was then purified with AMPure XP beads in a 1:1 dilution factor. The beads were washed away with 70% ethanol and the DNA eluted in 12.50 µl of nuclease-free water.

Table 22. Components for the different steps of the Ligation Sequencing gDNA - PCR barcoding library preparation for MinION Oxford Nanopore sequencing. Some volumes were modified from the manufacturer's protocol. The volumes shown are for single reactions. The supplier was Oxford Nanopore, apart from those asterisked, where the supplier was New England Biolabs, USA.

1. End-prep	
Reagent	**Volume (µl)**
gDNA	8.30
Ultra II End-prep Reaction Buffer *	1.20
Ultra II End-prep Enzyme Mix *	0.50
Total volume	10
2. Ligation of barcode adapter	
End-prepped DNA	10
Barcode Adapter	6.67
NEBBlunt/TA Ligase Master Mix *	16.67
Total volume	33.34
3. Barcoding PCR	
PCR Barcode (one of BC1-BC96 from the PCR Barcoding Expansion Pack (10 µM))	0.50
Adapter-ligated DNA	2
LongAmp Taq 2X Master Mix*	5
Nuclease-free water	2.50
Total volume	10
4. DNA repair and end-prep	
NEBNext FFPE DNA Repair Buffer *	47
NEBNext FFPE DNA Repair Mix *	3.50
Ultra II End-prep Reaction Buffer *	2
Ultra II End-prep Enzyme Mix *	3.50
Total volume	59
5. Ligation of sequencing adapter	
Pooled barcoded and end-prepped DNA	60
Ligation Buffer	25
NEBNext Quick T4 DNA Ligase *	10
Adapter Mix	5
Total volume	100
6. Library loading	
Sequencing Buffer	37.50
Loading Beads	25.50
DNA libraries	12
Total volume	75

The end-prepped DNA was then transferred to a new 0.2 ml tube along with the components in Table 22(2) and mixed. The reaction was incubated for 10 minutes at room temperature. The barcode-ligated DNA was then purified with AMPure XP beads in a 1:0.90 dilution factor (33.34 µl of DNA:30 µl of beads). The beads were washed away

with 70% ethanol and the DNA eluted in 8.30 µl of nuclease-free water. A barcoding PCR reaction was subsequently set up in a new 0.2 ml tube Table 22(3) and run with the cycling conditions in Table 23 in a thermocycler.

Table 23. Thermal protocol for the Barcoding PCR of the Ligation Sequencing gDNA - PCR barcoding library preparation for MinION Oxford Nanopore sequencing. The number of cycles and time for denaturation, annealing and extension were adjusted according to the length of the target fragment (630 bp).

Step	Temperature	Time	Number of cycles
Initial denaturation	95 °C	3 minutes	1
Denaturation	95 °C	15 seconds	16
Annealing	62 °C	15 seconds	
Extension	65 °C	50 seconds	
Final extension	65 °C	2 minutes	1
Hold	4 °C	∞	

All the barcoded samples were quantified for concentration using the Qubit 1X dsDNA HS kit. The samples were then pooled and normalised in a 1.5 ml DNA LoBind Eppendorf tube (Eppendorf, Germany) containing a 47 µl dilution and a total DNA input of 1 µg with equal DNA concentrations across each sample. Subsequently, a second end-prep step and a DNA repair were carried out on the pooled barcoded samples in a 0.2 ml tube with the components in Table 22(4) after incubation for 5 minutes at 20 °C and 5 minutes at 65 °C in a thermocycler. Then, the nucleic acids were purified with AMPure XP beads in a 1:1 dilution factor. The beads were washed away with 70% ethanol and the DNA eluted in 61 µl of nuclease-free water into a new 1.5 ml Eppendorf DNA LoBind tube. The sequencing adaptors were then ligated to the fragments, and the components in Table 22(5) were mixed and incubated at room temperature for 10 minutes. Fourty µl of AMPure XP beads were added into the mix for nucleic acid purification. The beads were washed away with 250 µl of Short Fragment Buffer and the DNA eluted in 15 µl of Elution Buffer into a new 1.5 ml Eppendorf DNA LoBind tube. The flow cell was slid into the

MinION and primed by mixing 30 µl of Flush Tether into the tube of Flush Buffer, loading the flow cell first with 800 µl of the mix and then with 300 µl after 5 minutes. Seventy-five µl of DNA libraries were loaded immediately after mixing the components in Table 22(6), and the sequencing run was started. The data acquisition and base calling were performed with the MinION device and the MinKNOW software.

2.1.7.3. PG-Seq (Perkin Elmer, USA)

The PG-Seq Rapid Non-Invasive PGT (Perkin Elmer, USA) protocol and kit were used for WGA and indexing. The Illumina platform was used for denaturation and loading of the sequencing libraries (MiSeq Reagent Kit V3) and for data acquisition and base calling (MiSeq Sequencing System). A master mix was prepared for the WGA – PCR1 (Table 24(1)) and added to approximately 6 µl of SCM in a 0.2 ml tube. Positive controls (1 µl of 30 pg/µl DNA and 5 µl of nuclease-free water) and a no-template control were included. The samples were then transferred to a thermocycler and run with the thermal protocol in Table 25(1).

WGA – PCR2 was performed immediately after. A master mix was assembled (see Table 24(2)) and transferred to the WGA – PCR1 products with 8 µl of an Indexed Primer. The samples were then incubated in a thermocycler, following the thermal protocol in Table 25(2), and the products were run in an agarose gel to confirm amplification.

Table 24. Components for the two-step WGA of the PG-Seq Rapid Non-Invasive PGT protocol. The volumes shown are for single reactions. The master mixes contained enough volume for the required number of reactions, including controls and a 10% surplus.

1. WGA – PCR1	
Reagent	**Volume (µl)**
Nuclease-free water	6.20
WGA PCR Buffer	10
Round 1 Primer	3.80
dNTPs	0.50
WGA Polymerase	0.50
Total volume	21
2. WGA – PCR2	
Nuclease-free water	4.42
WGA PCR Buffer	40.50
WGA Polymerase	1.08
Total volume	46

Table 25. Thermal protocols for the two-step WGA of the PG-Seq Rapid Non-Invasive PGT.

1. WGA – PCR1			
Step	**Temperature**	**Time**	**Number of cycles**
Initial Denaturation	95 °C	3 minutes	1
Denaturation	98 °C	20 seconds	8
Annealing	25 °C (Ramp to 50 °C)	1 minute (0.20 °C/second)	
Extension	(Ramp to 72 °C) 72 °C	(1 °C/second) 30 seconds	
Denaturation	98 °C	20 seconds	16
Annealing	58 °C	1 minute	
	72 °C	30 seconds	
Hold	10 °C	∞	1
2. WGA – PCR2			
Initial denaturation	95 °C	3 minutes	1
Denaturation	98 °C	20 seconds	4
Annealing	50 °C	30 seconds	
Extension	72 °C	30 seconds	
Denaturation	98 °C	20 seconds	6
Annealing	62 °C	30 seconds	
Extension	72 °C	30 seconds	
Hold	10 °C	∞	1

Samples that amplified successfully were then purified with PG-Seq Rapid Non-invasive PGT Purification Beads in a 1:0.90 dilution factor (40 µl of DNA:36 µl of beads), washed with 80% ethanol and resuspended in 30 µl of Resuspension Buffer. Twenty-five

µl of the purified indexed library were recovered. The libraries were subsequently pooled and normalised to equal quantities of DNA per sample. Four nM of the normalised libraries were then denatured with 0.2 N of sodium hydroxide (NaOH). Finally, 15 pM of libraries were loaded and sequenced in the MiSeq Instrument. The data was analysed with the PG-Find Software V3 (Perkin Elmer, USA).

2.1.8. Bioinformatics

2.1.8.1. DNA sequence retrieval

Gene and specific DNA sequences of interest were obtained from the National Centre for Biotechnology Information (NCBI) Gene and Nucleotide databases (https://ncbi.nlm.nih.gov). If a gene was not found in the Gene database, genomic coordinates and accession numbers were retrieved from peer-reviewed academic publications. Where relevant, details about the sources of DNA sequences are described in the materials and methods section of the respective chapter.

The differentially methylated regions (DMRs) DNA sequences (from the GRCh37/hg19 genome assembly) were obtained from the UCSC Genome Browser (https://genome.ucsc.edu) (UCSC, USA) found with the published genomic coordinates of characterised regions. The UCSC Genome Browser was also used to identify SNPs. The Bisulfite Primer Seeker (https://www.zymoresearch.com/pages/bisulfite-primer-seeker/) (Zymo Research, USA) was used to generate the DMRs bisulfite-converted sequences.

2.1.8.2. Primer design

2.1.8.2.1. Primer design for conventional singleplex and multiplex PCR

The primers used in this thesis were designed using the primer design online tool Primer3 (https://primer3.ut.ee/) (Whitehead Institute for Biomedical Research, USA). Standard primer picking conditions were used, but the amplicon size ranges depended on the experiment's purpose (Table 26). If the primers designed were to be used for a multiplex reaction, then sequences with similar size, melting temperature (T_m) and GC content were selected.

Table 26. Standard Primer3 primer picking conditions.

Parameter	Minimum	Optimal	Maximum
Primer size (bp)	18	20	23
Primer T_m	57°C	59°C	62°C
Primer GC content (%)	30	50	70
Product size ranges	Variable	Variable	Variable

Primer specificity was further confirmed with Primer BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) (NCBI, USA) using the reference genome of interest and the Primer search, ePCR tool from BiSearch (http://bisearch.enzim.hu/) (Institute of Enzymology, Hungary). Primers were also checked for hairpin formations using the OligoCalc online tool (http://biotools.nubic.northwestern.edu/OligoCalc.html) (Northwestern University, USA) to avoid self-annealing.

If primers were designed to be used in a multiplex reaction, sequences were analysed in the Multiple Primer Analyser online tool (https://www.thermofisher.com/uk/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-

scientific-web-tools/multiple-primer-analyzer.html) (Thermo Fisher Scientific, USA) for complementarity to avoid cross-dimerisation.

2.1.8.2.2. Primer design for bisulfite-converted DNA

The specific primers for DMRs were designed with the online primer generator tool for methylation PCR MethPrimer (http://www.urogene.org/cgi-bin/methprimer/methprimer.cgi) (PUMCH, China). This tool allowed for the input of the entire sequence of the DMR. The software identified cytosine phosphate guanine sites (CpGs) islands and designed primers targeting loci within these regions. The primer sequences excluded CpGs to avoid allele amplification bias.

Experimental section 1. Aneuploidy mechanisms in oocytes and preimplantation embryos

Chapter 3. Study of the expression of nine genes with roles in DNA repair in cleavage stage embryos and blastocysts harbouring DNA damage

3.1.1. Introduction

3.1.1.1. DNA damage and segmental aneuploidy in embryos

Chromosomal abnormalities, including segmental and whole chromosome aneuploidies, are a major cause of embryo developmental failure. However, mechanisms behind the two forms differ. Whilst segmental aneuploidy is the result of unrepaired DNA DSBs, whole chromosome aneuploidy results from chromosome mis-segregation and is linked to advanced maternal reproductive age, unlike segmental abnormalities (Fragouli et al., 2013, Babariya et al., 2017). This chapter will focus on DNA damage, segmental aneuploidy and embryo DNA repair capacity.

The DNA damage that leads to segmental aneuploidy can be caused by endogenous and exogenous agents. It generally occurs during DNA replication, characterised by the stalling of replication forks and during DNA synthesis in the G2 phase of the cell cycle, although oxidative stress and mutagens can also provoke strand breakage. DNA damage causes DSBs and subsequent chromosome breakage,

micronucleation and aneuploidy (Babariya et al., 2017, Palmerola et al., 2022). This mechanism can be observed through CRISPR-Cas9 technology for genome editing, which operates by inducing DNA DSBs, and results in segmental aneuploidies in embryos subjected to treatment (Zuccaro et al., 2020)

Babariya et al. (2017) showed that chromosome breakages are prevalent in human preimplantation embryos, demonstrating that whilst a fraction are from meiotic origin, most breakages occur during the first few cell divisions after fertilisation. Segmental aneuploidies from meiotic origin arise during oocyte or sperm gametogenesis and are transmitted to the embryo uniformly affecting all cells. Conversely, mitotic segmental aneuploidies cause gains or losses between a subset of embryonic cells.

3.1.1.2. DNA repair mechanisms

Like all cell types, embryos are constantly exposed to stressors that have the potential to cause DSBs. Stressors, or the damage itself, elicit a robust response in somatic cells coordinated by checkpoints, known as the DNA damage response (DDR), that trigger the activity of a vast network of proteins. The DDR pauses the cell cycle with the aim of correcting the damage. If the cell fails to correct it, the cell remains in an arrested state and apoptotic pathways are activated (Figure 7(C)) (Aguilera and García-Muse, 2013). Cells employ two major pathways to resolve DSBs: non-homologous end joining (NHEJ) and homologous recombination (HR). In NHEJ, the two ends created from the break are pieced together. Although it is a quick repair that is functional in all cell cycle phases, it is error prone (Figure 7(A)). Repair through the HR pathway, on the other hand, requires an intact homologous strand, or sister chromatid, to copy information. Thus, HR is only operational in the S/G2 phases (Figure 7(B)) (Burma et al., 2006).

Consequently, these two pathways, are given different weight depending on the phase of the cell cycle, and in mammalian embryos the choice pathway is dependent on the stage of development (Khokhlova et al., 2020).

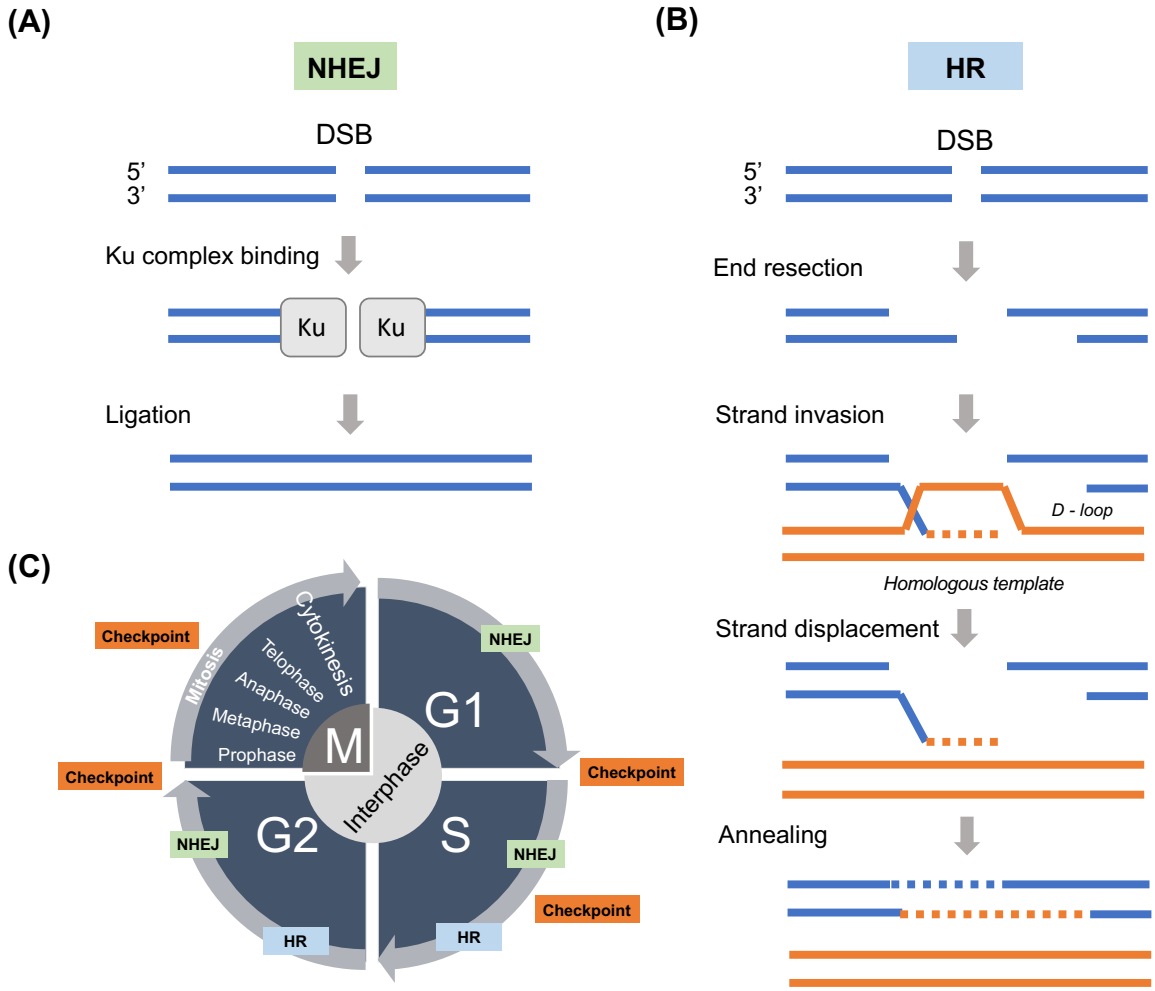


Figure 7. NHEJ and HR pathways for the repair of DSBs. (A) During NHEJ DSBs repair, damage is sensed and the Ku complex binds to the site of the break to allow the free ends to rejoin directly, sometimes leading to DNA sequence deletions. **(B)** In HR synthesis-dependent strand annealing repair, the break ends undergo resection so that a 3' single strand can invade the homologous template through a D-loop structure (also known as displacement loop) to initiate synthesis. The newly synthesised strand then anneals to the other DNA end allowing the DSB repair. **(C)** Four DNA damage checkpoints exist throughout the cell cycle: G1/S transition, inter-S phase, G2/M transition and M checkpoints. In somatic cells, the phase of the cell cycle at which DSBs are sensed impacts the repair pathway selection. HR, homologous recombination; NHEJ, non-homologous end joining.

3.1.1.2.1. DNA repair capacity in embryos

One of the most intriguing questions about preimplantation development is the capacity of embryos to respond to DNA damage. As discussed, segmental aneuploidies are particularly common during the first cell divisions in the embryo. Babariya et al. (2017) observed a significant increase in chromosome segmental abnormalities between oocytes and day-3 cleavage stage embryos (10.40% vs. 24.30%), although this percentage declined in the analysed blastocysts (15.60%). Studies using CRISPR-Cas9, where injection is performed at the two-pronuclear stage (2PN) or 2-cell stage, also provide data suggesting that embryos' DSBs repair capabilities may be deficient during the first mitoses due to the inflated rate of aneuploidy observed in embryos subjected to treatment (Fogarty et al., 2017, Zuccaro et al., 2020). Moreover, gene expression studies have demonstrated that preimplantation embryos with morphologic anomalies, such as micronucleation and fragmentation, present disturbances in the activity of genes that participate in DNA repair, cell cycle regulation and apoptosis (Wells et al., 2005b).

3.1.1.2.2. Embryonic genome activation (EGA)

The observed increased rate of segmental aneuploidy in preimplantation embryos coincides with a stage of development that has been known to be prior to EGA. Mature spermatozoa and MII oocytes are transcriptionally quiescent and the earliest events of embryonic development, including first cell divisions, are believed to be driven by maternal mRNAs and proteins accumulated in the vast cytoplasm of oocytes (Zhou and Dean, 2015). Nevertheless, the embryonic genome is eventually needed for further development (Jukam et al., 2017).

EGA is defined as the first transcription in the embryo. This event was typically understood to occur between the 4- and 8-cell stages in humans (Kakourou et al., 2013). This understanding is supported by studies in mice embryos, which show that a large cohort of genes (~2,500) becomes fully activated between the 4- and 8-cell stage (Xue et al., 2013). This rate of gene activation continues in later mitoses of early development that mark the dynamic morphological changes that give rise to the blastocyst formation. However, recent evidence suggests that the window during which embryos become fully transcriptionally active may be broader than between the 4- and 8-cell stages, even beginning as early as following fertilisation (Asami et al., 2022). This is based on an understanding that embryo gene activation may be a gradual process, albeit with a major wave of activation occurring between the 4- and 8-cell stages. This postulation, in line with the maternal-to-zygotic transition (MZT), postulates that the embryo becomes fully transcriptionally active after a series of tightly coordinated events, during which previously accumulated oocyte mRNAs become depleted (Jukam et al., 2017).

The above findings lead to the question of whether DNA damage response and DSBs repair pathways are functional prior to the major wave of EGA. To date, direct evidence associating embryo capacity to repair DNA damage at different stages of development is scarce.

3.1.1.3. Study aims

This chapter aims to evaluate the capacity of embryos to respond to DNA damage prior to, during and after the main wave of EGA. The study in this chapter analysed differences in the expression of genes with critical roles in DDR. Micronucleation and segmental aneuploidy are the models of DNA damage.

The specific aims are:

1) To validate a duplex RT-qPCR assay for the analysis of the relative expression of genes in cleavage stage embryos and blastocysts.
2) To analyse the presence of segmental aneuploidies in single blastomeres.
3) To study the relative gene expression of the genes *RAD51, RAD50, TP53, ATM, DMC1, BRCA2, BRCA1* and *XRCC4* in embryos up to the 8-cell stage and in blastocysts with DNA damage.

3.1.2. Materials and methods

3.1.2.1. Ethics statement

The experiments in this study were reviewed and approved by the Western Institutional Review Board and assigned the biomedical research protocol WIRB 20050731.

3.1.2.2. Experimental design and sample selection

A total of 79 embryos, 41 at cleavage stage and 38 blastocysts, were obtained from couples undergoing IVF treatment at RMA New Jersey, USA. Embryo manipulation techniques were performed in line with the standardised protocols for each clinic.
The 41 cleavage stage embryos comprised of three to eight cells and were identified as developmentally arrested before compaction on day-7 post-fertilisation. Thirteen of these embryos harboured micronuclei, 22 did not and for six of them, this information is unknown. A single blastomere was biopsied from the cleavage stage embryos and placed in a 0.2 ml tube. The rest of the embryo was placed in a separate tube.

The 38 blastocysts were identified as aneuploid on day-5, day-6 or day-7 post-fertilisation and, therefore, unsuitable for uterine transfer after being biopsied and undergoing routine PGT-A. Twenty-one had segmental aneuploidies as a sole defect or in addition to whole chromosome abnormalities. The rest presented whole chromosome abnormalities only (Figure 8). Embryos with three or more whole chromosome abnormalities were excluded from this study. Each blastocyst was placed in a 0.2 ml tube. All the blastomeres and embryos were sent to Juno Genetics, UK and stored at -80 °C until ready for RNA extraction. Information about the embryos obtained can be found in **Appendix A**.

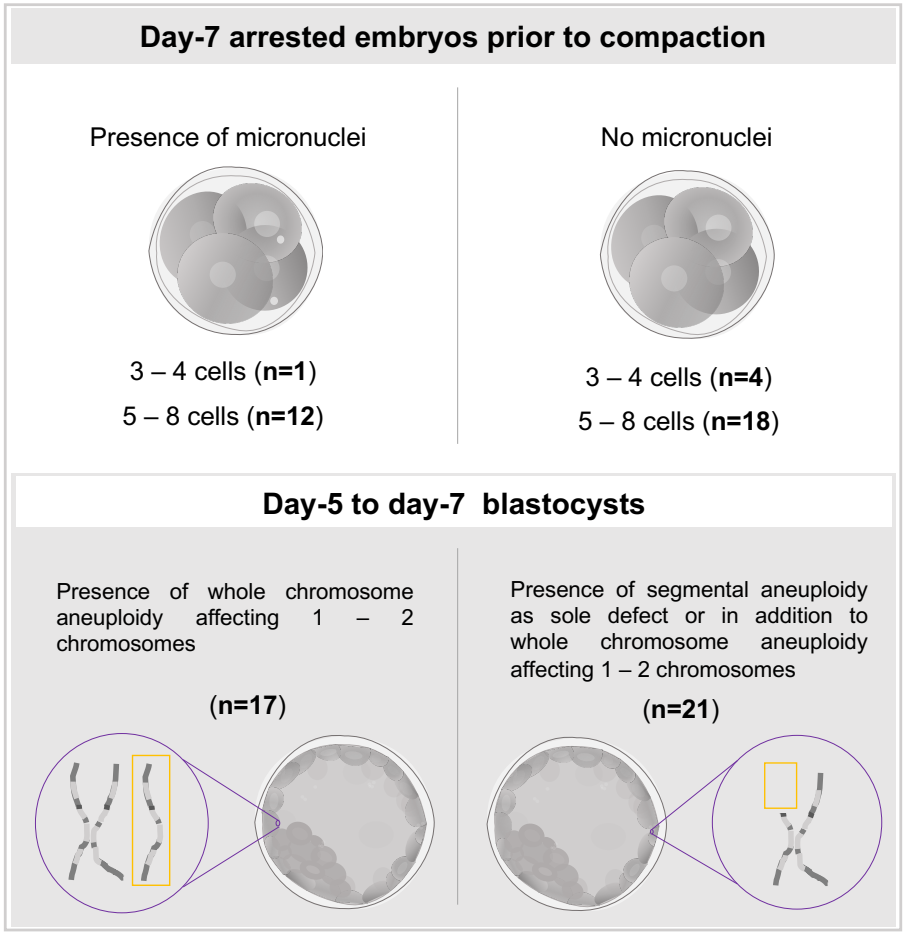


Figure 8. Experimental design for the study of the expression of nine genes with a role in DNA repair in cleavage stage embryos and blastocysts harbouring DNA damage.

3.1.2.3. Aneuploidy analysis of cleavage stage embryos

Single blastomeres were first subjected to WGA using the SurePlex method, as described in **Chapter 2**. **General materials and methods**.

Amplification products were run in a 1% agarose gel electrophoresis (1 g of agarose) for 10 minutes at 200 volts. The gel was then visualised to confirm amplification. A more detailed description of the gel electrophoresis visualisation is in **Chapter 2**.

The blastomeres that amplified successfully were subsequently subjected to NGS and screened for aneuploidies using the VeriSeq platform as described in **Chapter 2**. Whole chromosome and segmental gains or losses were determined by copy number values in the chromosome plot generated by BlueFuse. Since only one blastomere was analysed, any apparent mosaicism was disregarded. Losses were identified when reads aligned to copy number values <1.5 and gains to copy number values >2.5.

3.1.2.4. RNA extraction, cDNA reverse transcription and cDNA pre-amplification

RNA was extracted and reversed-transcribed from the 79 embryos using the TaqMan PreAmp Cells-to-Ct kit. The resulting cDNA was then pre-amplified using the TaqMan PreAmp Master Mix and 10 TaqMan Gene Expression Assays in multiplex to amplify the targets (Table 27), as per **Chapter 2**.

Table 27. Function of genes and RT-qPCR TaqMan Assays.

Gene	Function of protein	Assay ID, dye-quencher	Assay design	Amplicon length (bp)
ACTB	*ACTB* is a highly expressed housekeeping gene with a role in structural cell maintenance. This gene was included in this study as endogenous control.	Hs01060665_g1, VIC-MGB	Probe spans exons	63
RAD50	Forms part of the MRN (MRE11-RAD50-NBS1) multiprotein complex that assists with the linking of DNA damage detection to DNA repair (Petrini, 1999).	Hs00990023_m1 FAM-MGB	Probe spans exons	63
RAD51	Promotes HR in DSB repair allowing homologous strand invasion (Baumann and West, 1998).	Hs00947967_m1 FAM-MGB	Probe spans exons	58
TP53	Maintains genome stability by having a key role in DNA damage response mechanisms (Williams and Schumacher, 2016).	Hs01034249_m1 FAM-MGB	Probe spans exons	108
ATM	Upstream trigger that recognises DSBs and phosphorylates critical proteins involved in HR repair (Paull, 2015).	Hs00175892_m1 FAM-MGB	Probe spans exons	89
DMC1	RAD51 homologous that similarly promotes HR (Filippo et al., 2008).	Hs01095981_m1 FAM-MGB	Probe spans exons	76
BRCA1	Along with BRCA2, repairs DNA damage via HR (Powell and Kachnic, 2003).	Hs01556193_m1 FAM-MGB	Probe spans exons	59
BRCA2	BRCA1, BRCA2, RAD51 and DMC1 operate similarly (Chen et al., 1999).	Hs00609073_m1 FAM-MGB	Probe spans exons	110
XRCC4	XRCC4 is recruited by DNA protein kinase (DNA-PK) holoenzyme to help stabilising, along the Ku complex, the NHEJ complex (Gao et al., 1998).	Hs00243327_m1 FAM-MGB	Probe spans exons	67
LIG4	LIG4 and XRCC4 operate similarly in the NHEJ repair pathway (O'Driscoll et al., 2001).	Hs01866071_u1 FAM-MGB	Both primers and probe map within a single exon	155

3.1.2.5. Validation of the RT-qPCR method and gene expression analysis

The purpose of the RT-qPCR experimental design was to simultaneously amplify each of the nine specific targets and the endogenous control, used as a calibrator, in duplex reactions. This study initially performed a series of validation experiments to confirm the reliability of the data obtained from the RT-qPCR.

First, the amplification uniformity across the nine specific targets during the pre-amplification was analysed. Two sets of reactions were prepared per target: RT-qPCR using pre-amplified cDNA and bulk (not pre-amplified) cDNA. For the pre-amplified cDNA reactions, 4 ng of reverse-transcribed cDNA from Universal Human Reference RNA (Agilent, USA) were pre-amplified using X14 PCR cycles and the products were diluted using a 1:20 factor (7.50 µl of products in 142.50 µl of nuclease-free water). Three ng/µl (13.50 ng total) of reverse-transcribed cDNA from the Universal Human Reference RNA was used for the bulk cDNA reactions. The RT-qPCR was carried out to amplify the 10 targets in singleplex on both cDNA dilutions using the TaqMan Gene Expression Master Mix. All RT-qPCR amplifications in this study were performed in line with **Chapter 2**. A no-template control was included and all reactions were run in duplicates.

Another element of the present experimental design was to analyse all the samples, cleavage stage embryos and blastocysts, in the same 384-well plate per target. This study aimed to use the same dilution factor throughout the samples after pre-amplification for consistency. The pre-amplification products of one cleavage stage embryo and one blastocyst (both randomly selected) were diluted using 1:10, 1:20 and 1:40 dilution factors to select an optimal dilution after analysis. The 10 targets were amplified in singleplex reactions for all the dilutions. A no-template control was included and all reactions were run in triplicates.

A duplex validation was carried out following the pre-amplification uniformity analysis and the dilution factor selection. The duplex validation began with three 10-fold dilution series (1:10, 1:100 and 1:1,000) of the pre-amplification products of one randomly selected blastocyst. It was determined that this dynamic range would encompass most of the experimental samples. All 10 targets were amplified in singleplex and the nine targets of interest, along with the endogenous control, in duplex reactions for all the dilutions. In addition, the pre-amplification products of a 3-cell cleavage stage embryo were diluted using a 1:20 dilution factor and used as a representative sample. All 10 targets were amplified in singleplex and the nine targets of interest, along with the endogenous control, in duplex reactions for the representative sample. All reactions were run in triplicates and the no-template control in duplicate.

After the RT-qPCR validation, the gene expression of the nine targets of interest was analysed in duplex reactions. One 384-well plate was used per target; all 82 pre-amplified products from all the samples were run in the same plate for the same target using the same master mix (containing all the reaction components except the pre-amplified cDNA) to minimise pipetting error. For the positive control, 1.50 ng of reverse-transcribed cDNA from Universal Human Reference RNA was pre-amplified. A no-template control was included. All samples, including the controls, were run in triplicates.

3.1.2.6 Data analysis

Data was analysed using the QuantStudio Real-Time PCR Software v1.3 (Applied Biosystems, USA). A comparative C_t was used to analyse the results of the pre-amplification uniformity test to calculate the $\Delta\Delta C_t$ values between the pre-amplified and the bulk cDNA reactions. The endogenous control was used as a calibrator.

To select the dilution factor for pre-amplified products, a comparative C_t was used for analysis and the amplification plots of all targets were visualised. A target was chosen randomly and the ΔC_t between the endogenous control (calibrator) and the target C_t mean was calculated for the three dilutions.

For the duplex validation, a standard curve was generated first per target to ensure the fit of the line to the three standard points determined by the R^2 value and to calculate the reaction efficiency. Differences in the relative abundance of the target (or C_t mean) between the singleplex and the duplex reaction for the representative sample were also analysed.

For the gene expression analysis of the nine targets, a $\Delta\Delta C_t$ analysis was used to determine the relative quantification (RQ) of each target for all the samples. The RQ values obtained were then compared between the cleavage stage embryos that presented micronucleation and those that did not and between those that harboured segmental aneuploidies and those that did not. In the same way, these values were compared between the blastocyst groups with and without segmental aneuploidies. The endogenous control was used as calibrator and the positive control (the Universal Human Reference RNA) as the reference sample.

3.1.2.7. Statistical analysis

The standard deviation was calculated between replicate C_t values in the duplex validation, for both singleplex and duplex reactions. For the primary outcome data, the RQ values from the gene expression analysis were log-transformed to normalise the distribution and the differences in observations were tested using the unpaired *t*-test

with a level of significance of 5%. P-values of less than 0.05 were considered statistically significant. GraphPad Prism Software (Dotmatics, USA) was used for the analysis and graph drawing.

3.1.3. Results

3.1.3.1. Ploidy status of cleavage stage embryos

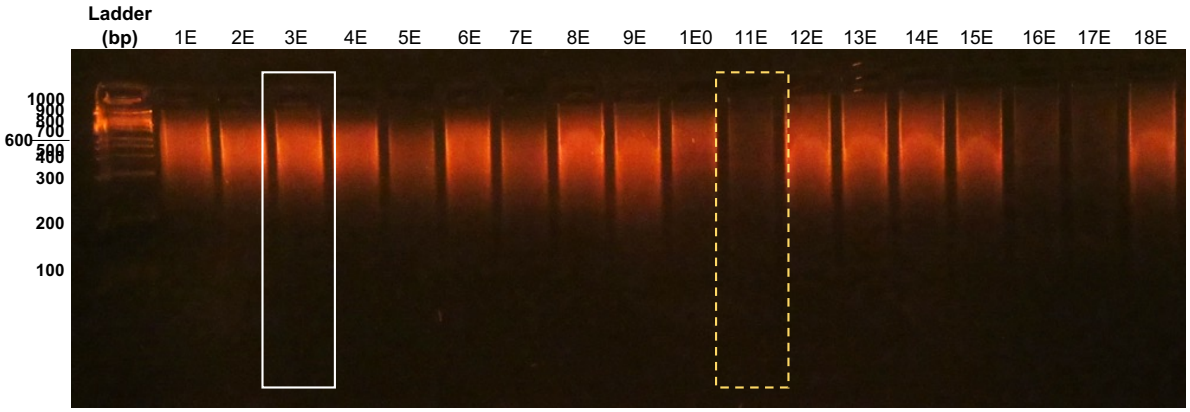
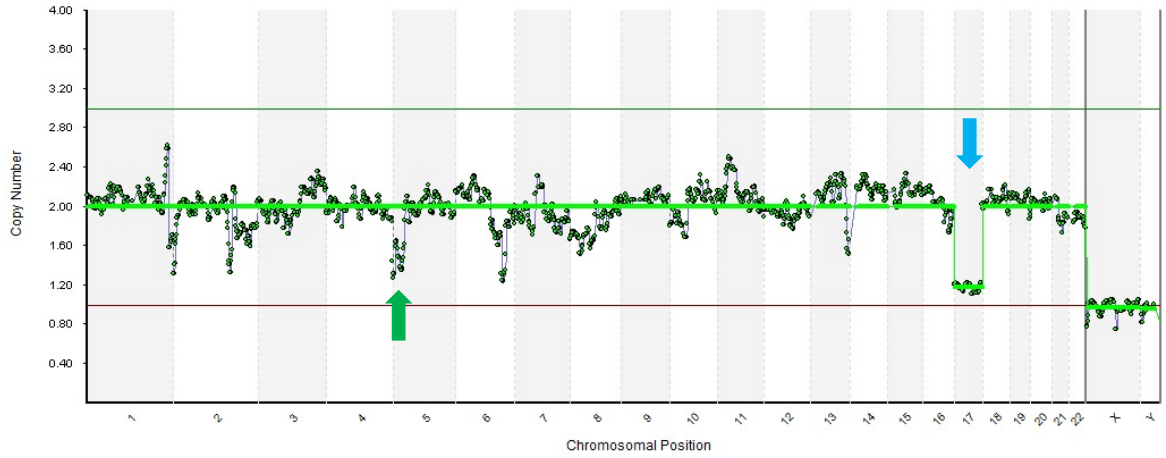


Figure 9. Agarose gel electrophoresis of WGA products from 18 representative single blastomeres following SurePlex amplification. Bands can be observed with sizes ranging 200 – 800 bp with different brightness intensities. Successful amplification is indicated with a white box, while amplification failure is indicated with a dashed yellow box. Only samples where amplification was clearly observed were analysed for aneuploidy with the VeriSeq method. Products with a size < 200 bp are primer-dimers. Samples were run in a 1% agarose gel electrophoresis for 10 minutes at 200 volts.

The presence of segmental aneuploidy in cleavage stage embryos was first studied. Due to insufficient SurePlex amplification, only 22 single blastomeres were included in the VeriSeq analysis (Figure 9). Moreover, due to poor quality sequencing data, it was only possible to determine the ploidy status of 11 of the 22 samples. Three cleavage stage embryos showed only whole chromosome aneuploidies and eight presented segmental abnormalities (Figure 10) (**Appendix A)**. This study was unable to obtain the location of chromosomal breakpoints for the segmental abnormalities.

(A)



(B)

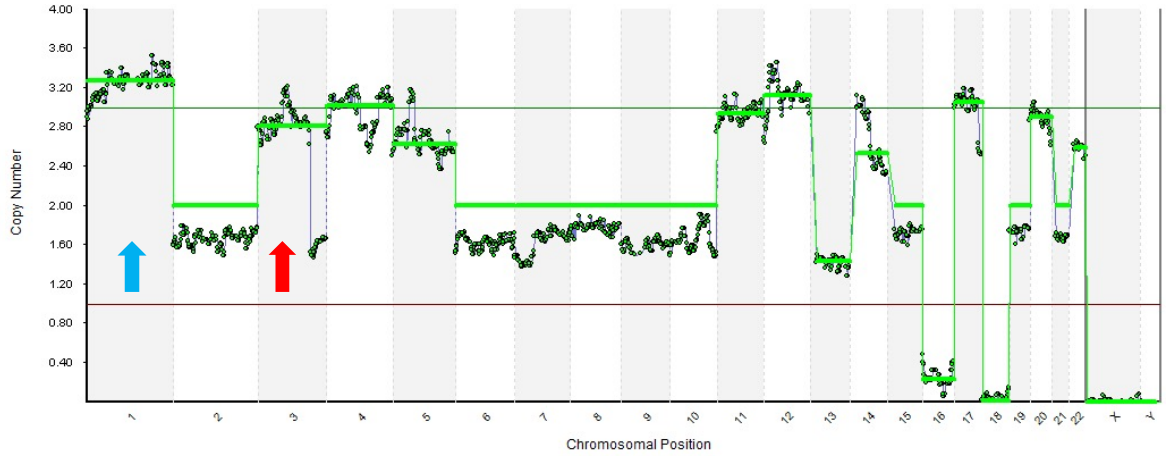


Figure 10. Chromosome copy number plots generated by the VeriSeq platform for the aneuploidy analysis of single blastomeres. (A) Representative example of a blastomere (ID: 9E, 45,XY,-17) harbouring a whole chromosome 17 loss (indicated with the blue arrow). The green arrow indicates sequencing artefacts. **(B)** A representative example of a blastomere (ID:18E, 52,00,+1,+3(seg),+4,+5,+11,+12,-13,+14,-16,+17,-17(seg),-18,+20,+22,-X,-Y) harbouring whole chromosome gains and losses (exemplified with the blue arrow) and segmental gains (exemplified with the red arrow). These chromosome plots illustrate the generalised level of noise obtained from the NGS results.

3.1.3.2. RNA extraction and reverse transcription efficiency

All the samples analysed in this study (79/79) displayed satisfactory RNA and reverse transcription efficiency, revealed by the detectable amplification levels generated during RT-qPCR of all ten targets in all the embryos.

3.1.3.3. Validation of the RT-qPCR method

3.1.3.3.1. Pre-amplification uniformity

The C_t mean values for each target were normalised to the calibrator (*ACTB*, the endogenous control) to obtain a ΔC_t per target for both the bulk and the pre-amplified cDNA groups. The ΔC_t value of the bulk cDNA group was subtracted from the ΔC_t value of the pre-amplified cDNA group to calculate the $\Delta\Delta C_t$ for the nine specific targets (Table 28). Values close to zero indicated pre-amplification uniformity.

Table 28. $\Delta\Delta C_t$ (ΔC_t(pre-amplified) – ΔC_t (bulk)) values for the nine specific targets. According to the TaqMan PreAmp Cells-to-Ct kit manufacturer, 90% of targets would typically produce values within the ± 1.5 range.

Target	$\Delta\Delta C_t$
RAD51	-1.62
RAD50	-1.68
TP53	-1.29
ATM	-1.23
DMC1	-0.43
BRCA2	-1.60
BRCA1	-1.69
XRCC4	-1.00
LIG4	0.04

3.1.3.3.2. Dilution factor selection for pre-amplified products

The amplification RT-qPCR plots for the cleavage stage embryo and blastocyst were first assessed visually to compare the C_t values between different dilutions of the pre-amplified products. The amplification plots for *RAD51* are shown in Figure 11 as a representative example. The C_t values increased with the cleavage stage embryo, as expected. The rest of the targets produced a similar pattern (not shown).

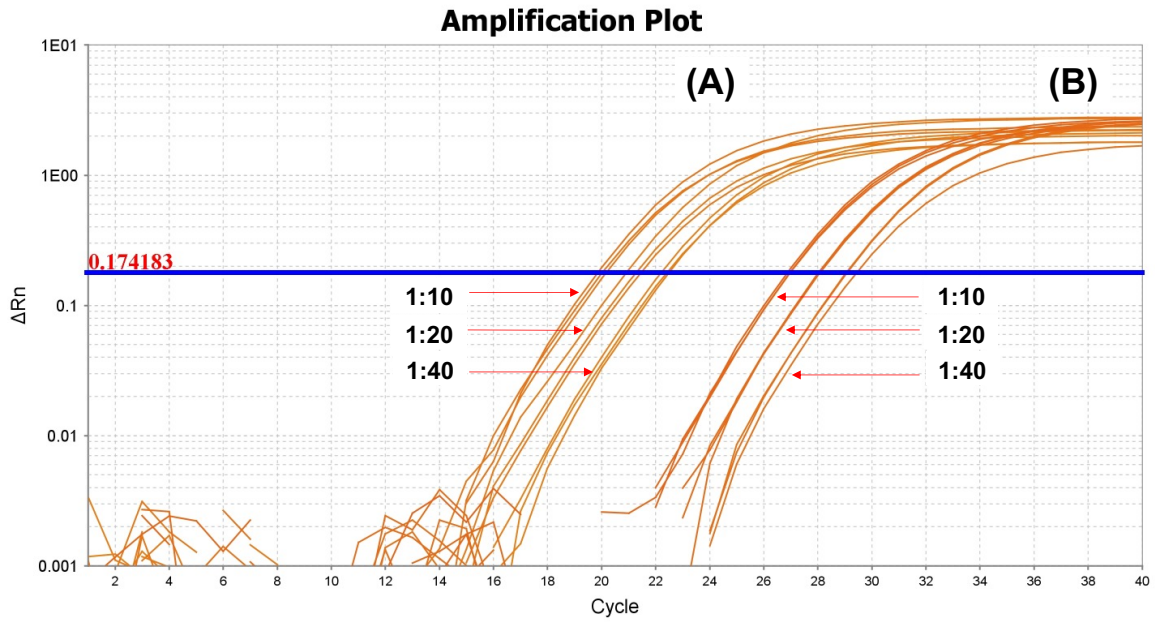


Figure 11. RT-qPCR amplification plots for *RAD51*. Blastocyst **(A)** and a cleavage stage embryo **(B)** diluted pre-amplified products.

RAD51 was selected to calculate the ΔC_t values for both the cleavage stage embryos and blastocysts dilutions by subtracting the C_t mean value of *ACTB* from the C_t mean value of *RAD51* (Table 29). Each embryo group produced consistent ΔC_t values between dilutions.

Table 29. ΔC_t (C_t mean (*RAD51*) – C_t mean (*ACTB*)) values for the three different pre-amplification dilutions. Despite minor variations, the ΔC_t values are consistent between dilutions within the cleavage stage embryo and the blastocysts groups.

ΔC_t for the cleavage stage embryo		
1:10	**1:20**	**1:40**
3.33	3.34	3.32
ΔC_t for the blastocyst		
1:10	**1:20**	**1:40**
6.23	6.13	6.22

3.1.3.3.3. Duplex validation

A standard curve per target was initially produced with the three 10-fold dilutions of the pre-amplification products of the blastocyst to study the R^2 generated. The

reaction efficiency was also calculated (Table 30). The smallest R^2 value was 0.994, demonstrating a good fit of the curve to the 10-fold dilutions for all the targets. The *BRCA1* assay showed the highest reaction efficiency (93.36%) and *LIG4* showed the lowest (76.78%).

Table 30. R^2 and reaction efficiencies for each target.

Target	R^2	Reaction efficiency
RAD51	0.998	85.48%
RAD50	0.998	87.25%
TP53	1	85.83%
ATM	0.999	81.89%
DMC1	0.997	82.09%
BRCA2	0.997	84.14%
BRCA1	0.994	93.36%
XRCC4	0.999	77.92%
LIG4	0.998	76.78%

The C_t values from the singleplex and the duplex reactions were obtained for each target in the representative sample. The C_t means and the standard deviations were compared (Figure 12). Despite some noticeable differences between the singleplex and duplex C_t mean, the highest variation is only 3.29%, corresponding to *LIG4*. The standard deviation increased in the duplex reactions for some genes, most notably *DMC1*. Nevertheless, the increment was not substantial and not all the targets followed this pattern.

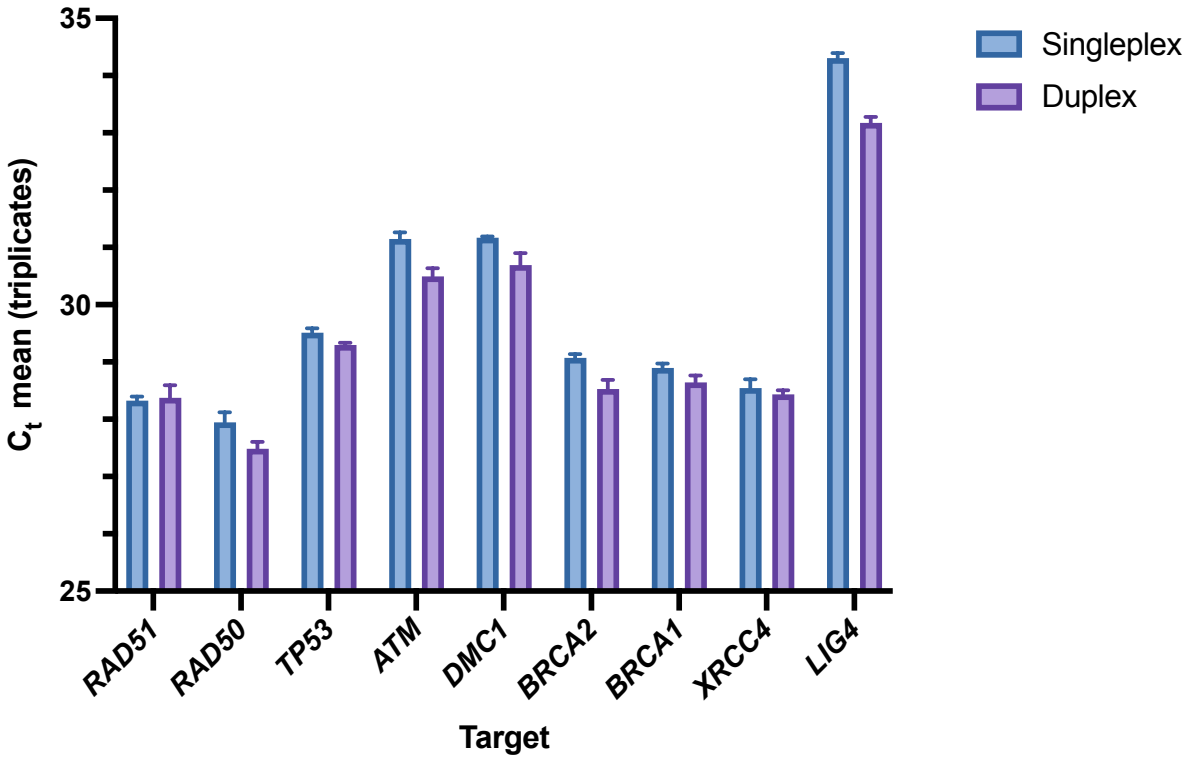


Figure 12. Relative abundance of the targets in the representative sample for the duplex validation. All targets produced similar C_t mean values between the singleplex and the duplex. *RAD51* showed the smallest difference (0.05), while *LIG4* showed the largest (1.13). For *RAD51*, *ATM*, *DMC1*, *BRCA2* and *BRCA1* the standard deviation increased for the duplex reactions. Nevertheless, the differences were not substantial.

3.1.3.4. Relative expression of the nine genes in embryos with signs of DNA damage

The $\Delta\Delta C_t$ analysis of the nine genes was subsequently carried out. As an overall observation of the RT-qPCR experiments, the C_t values for most targets of the cleavage stage embryos were higher than those of the blastocyst (Figure 13(A)). However, after normalisation with *ACTB*, the ΔC_t values of the blastocyst group are higher for most of the targets compared to the cleavage stage embryos and the standard deviation of ΔC_t values between samples increased for the cleavage stage embryos (Figure 13(B)). The final fold change, or RQ, calculations showed that the relative gene expression quantification against the reference sample (the Universal Human Reference RNA used as positive

control) is noticeably higher in the cleavage stage embryos. Only *TP53* showed an RQ value below one (0.53), while all the targets in the blastocyst group did (Figure 13(D)).

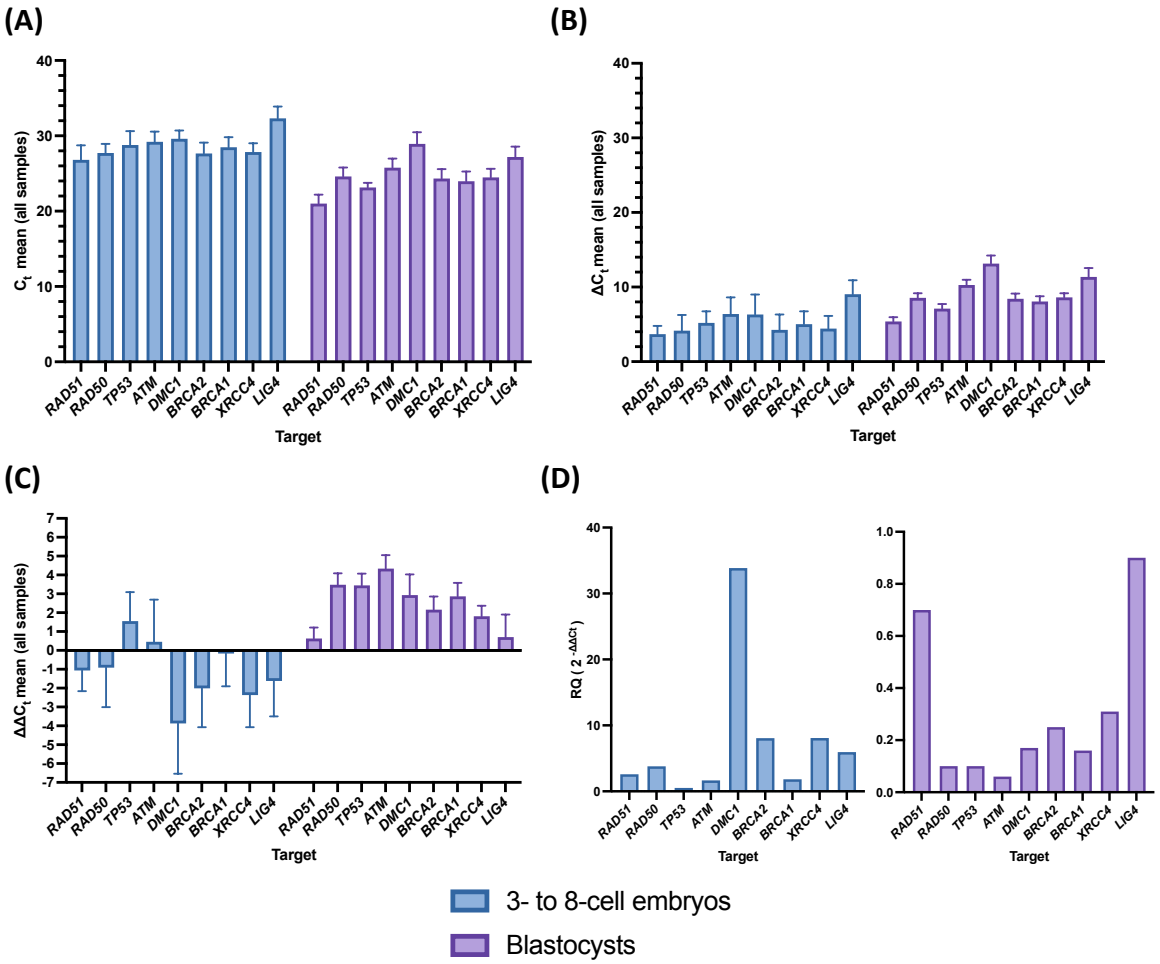
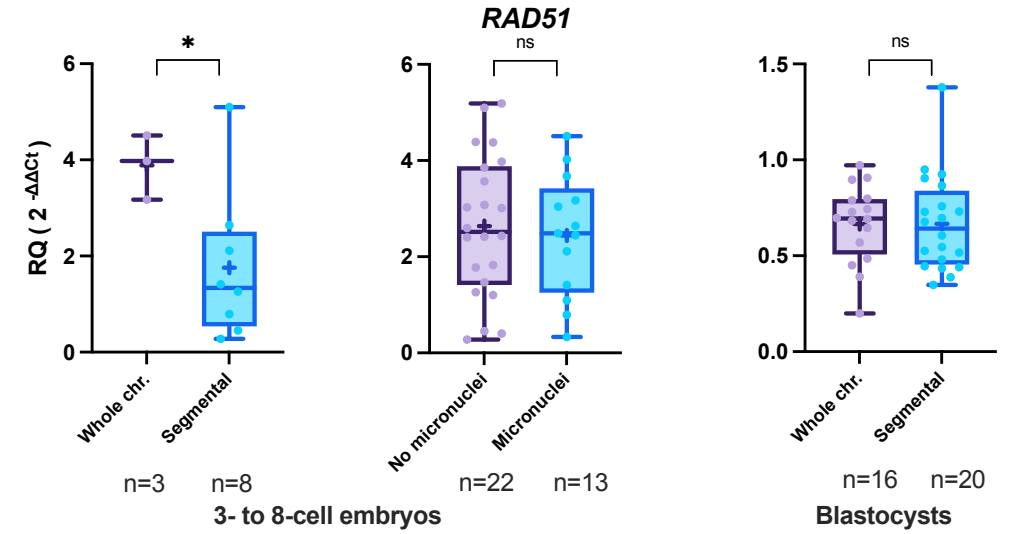
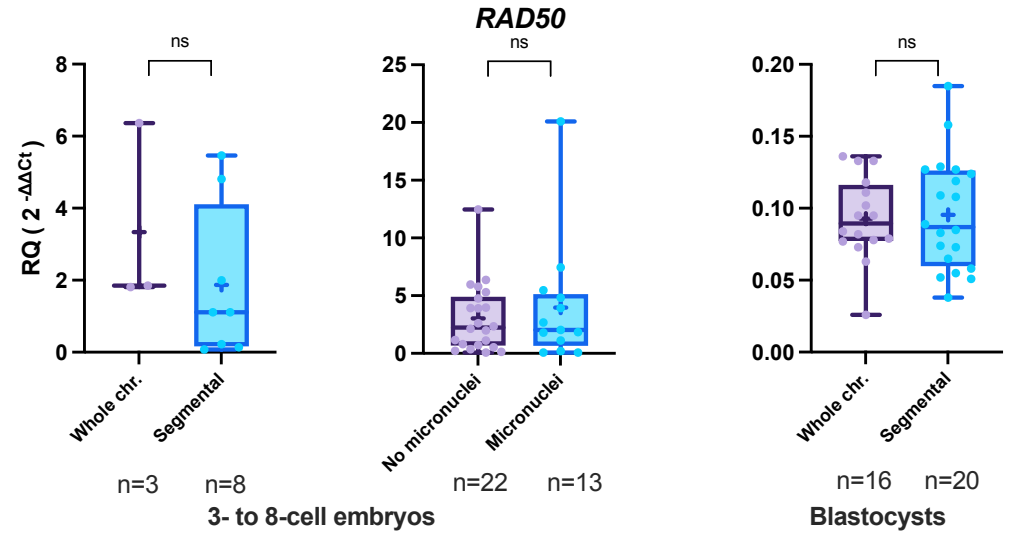
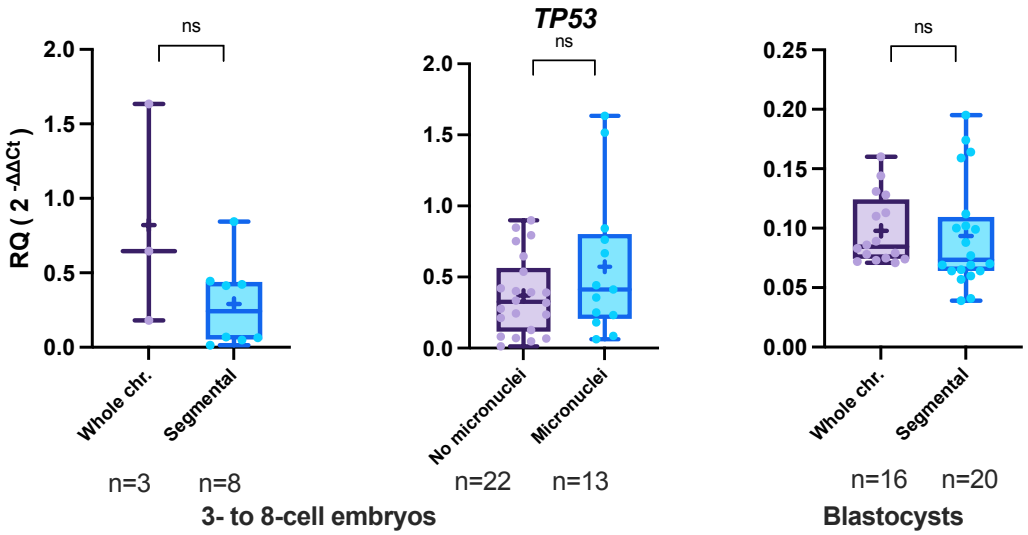


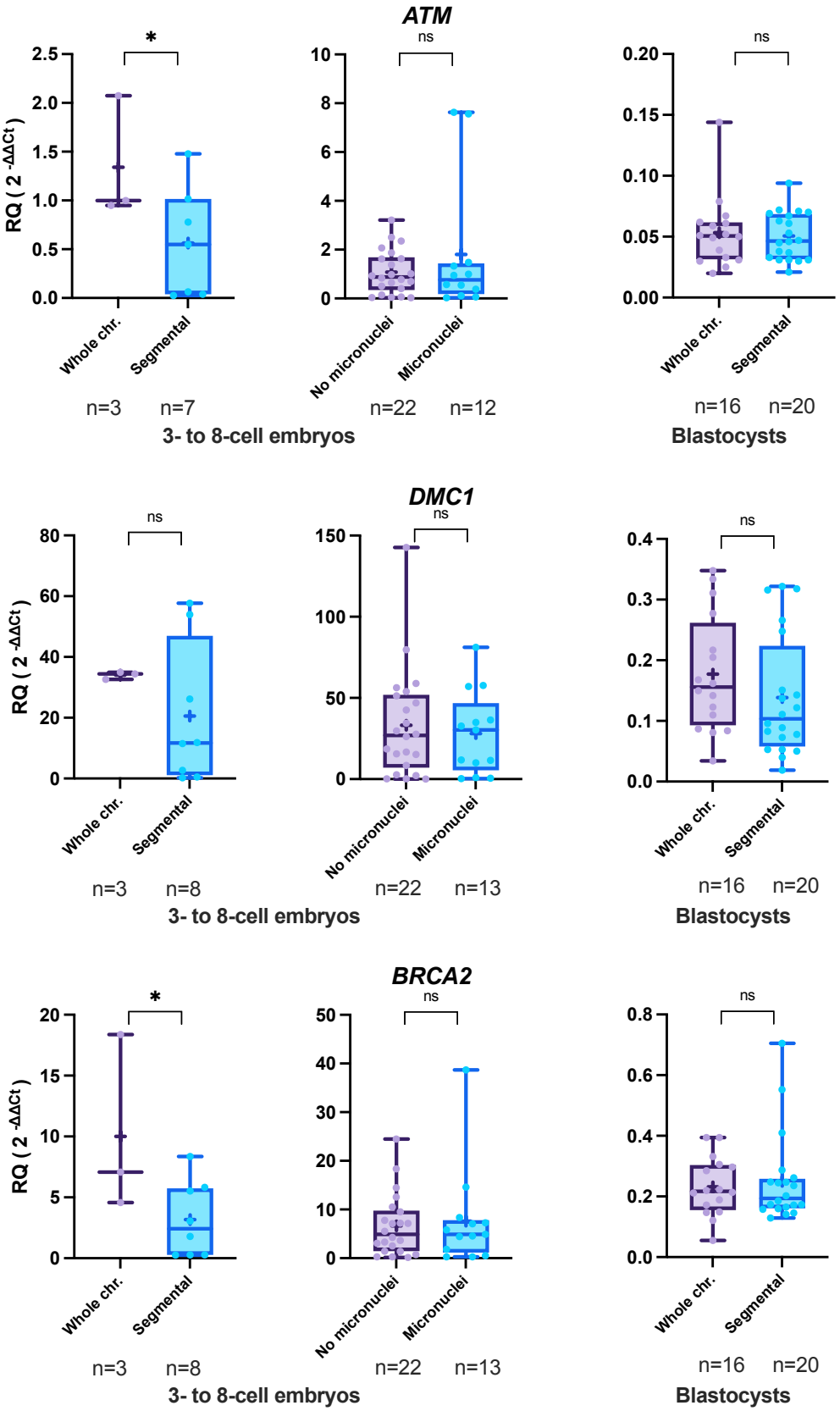
Figure 13. ΔΔC$_t$ analysis of the nine specific targets. *ACTB* was used as endogenous control and the Universal Human Reference RNA was used as the reference sample. RQ, relative quantification.

Figure 14 shows the comparisons of the gene expression relative to the Universal Human Reference RNA obtained between blastocysts with and without segmental aneuploidies and cleavage stage arrested embryos at day-7 with and without segmental aneuploidies, and with and without micronucleation. One blastocyst that consistently produced extremely high values for all targets was removed from the analysis. A generalised high variation of RQ values is observed between the targets analysed.


RAD51
*
ns
ns
RQ (2 -ΔΔCt)
Whole chr.
Segmental
No micronuclei
Micronuclei
Whole chr.
Segmental
n=3
n=8
n=22
n=13
n=16
n=20
3- to 8-cell embryos
Blastocysts


RAD50
ns
ns
ns
RQ (2 -ΔΔCt)
Whole chr.
Segmental
No micronuclei
Micronuclei
Whole chr.
Segmental
n=3
n=8
n=22
n=13
n=16
n=20
3- to 8-cell embryos
Blastocysts


TP53
ns
ns
ns
RQ (2 -ΔΔCt)
Whole chr.
Segmental
No micronuclei
Micronuclei
Whole chr.
Segmental
n=3
n=8
n=22
n=13
n=16
n=20
3- to 8-cell embryos
Blastocysts


ATM
DMC1
BRCA2
RQ (2 -ΔΔCt)
*
ns
Whole chr.
Segmental
No micronuclei
Micronuclei
n=3
n=7
n=22
n=12
n=16
n=20
n=8
n=13
3- to 8-cell embryos
Blastocysts

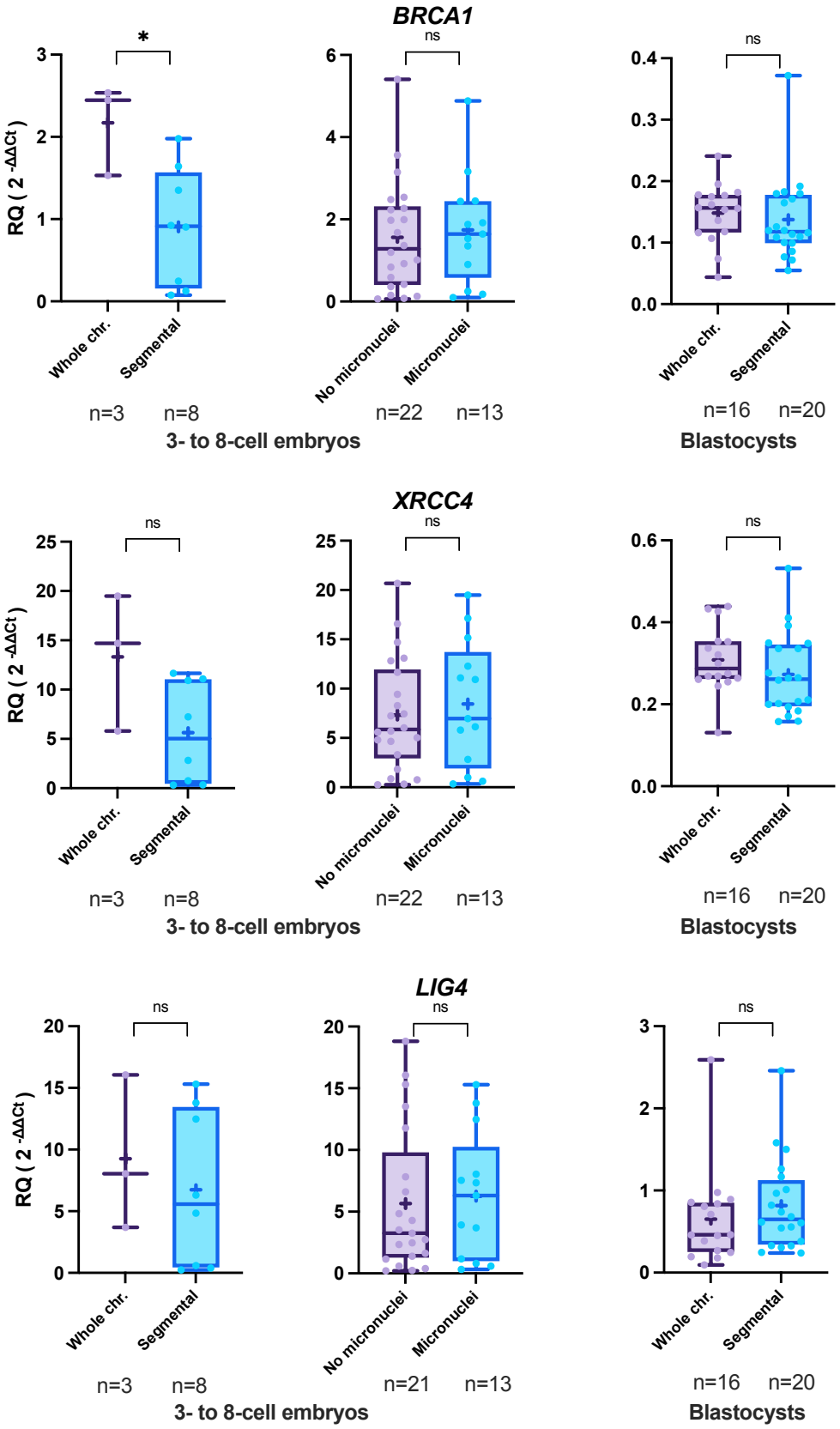


Figure 14. Relative expression of the nine genes with roles in DNA repair in embryos with signs of DNA damage. The statistical analysis (unpaired *t*-test) was performed using the $\Delta\Delta C_t$ mean value and the RQ values were plotted to visually present the levels of relative expression. '+' denotes the RQ mean in the plot. '*' denotes a level of significance P <0.05; ns, not significant.

Although, differences in relative expression when comparing the micronucleation status of blastomeres were not statistically significant for any of the nine genes analysed, there are some important observations to be made. The variation in the RQ values for *TP53* is low compared to other targets, but the mean and standard deviation are noticeably higher for cleavage stage embryos with micronuclei compared to ones without micronucleation (0.62 vs. 0.37 and 0.52 vs. 0.27, respectively). Some extremely high values (or outliers) can also be noted. A 6-cell cleavage stage embryo with micronucleation (ID: 2E) produced RQ values of 20.08 for *RAD50*, 38.68 for *BRCA2* and 4.88 for *BRCA1*. Embryo 2E and another 5-cell cleavage stage embryo with micronucleation (ID: 4E) produced RQ values of 7.63 and 7.56, respectively, for *ATM*. Interestingly, a 6-cell cleavage stage embryo without micronucleation (ID: 21E) showed an RQ value of 142.66 for *DMC1* and 5.41 for *BRCA1*.

None of the comparisons between blastocysts that harboured only whole chromosome aneuploidies and those that harboured segmental aneuploidies were statistically significant. However, like the micronucleation comparison, some outliers were observed. One blastocyst with one segmental aneuploidy and a trisomy in chromosome 21 (ID:5.2E) showed an RQ value of 1.38 for *RAD51*. Another blastocyst with one segmental aneuploidy and a trisomy in chromosome 20 (ID: 11.3E) presented an RQ value of 0.37 for *BRCA1*. Another blastocyst with two segmental aneuploidies and one trisomy in chromosome 9 presented an RQ value of 0.70 (ID:8.3E) for *BRCA2*. Interestingly, a blastocyst (ID: 3.1E) with a trisomy in chromosomes 11 and 15 presented an RQ value of 0.14 for *ATM*. Finally, a blastocyst with a trisomy in chromosomes 4 and 8 (ID: 2.3E) presented a value of 2.59 *LIG4* and a blastocyst 8.3E a value of 2.46.

The relative expression of *ATM, RAD51, BRCA2,* and *BRCA1* was significantly lower ($P<0.05$) in blastomeres that presented segmental chromosome abnormalities, according to the NGS-based aneuploidy analysis, compared to blastomeres that only harboured whole chromosome abnormalities. No statistical significance was found in the observed differences in the expression of *RAD50, TP53, DMC1, XRCC4* and *LIG4*. However, interestingly, the RQ values for *DMC1* of cleavage stage embryos with only whole chromosome aneuploidies are close, while the values vary significantly (from 0.21 to 57.69) for the embryos with segmental aneuploidies. However, as mentioned, the study determined the ploidy status of only 11 of the 22 blastomeres included for analysis and the small sample size must be considered when interpreting the results.

3.1.4. Discussion

One of the most intriguing questions about preimplantation development is the capacity of embryos to respond to DNA damage. Studies have suggested that prior to EGA, believed to occur at the 4- to 8-cell stage, embryos are more susceptible to aneuploidy (Zhou and Dean, 2015, Babariya et al., 2017). To help evaluating the DNA repair capacity at different stages of development, this chapter studied the expression of nine genes involved in DDR when cleavage stage embryos and blastocysts harbour DNA damage.

3.1.4.1. Single blastomere chromosomal number analysis in arrested embryos

This study first analysed single blastomeres from arrested cleavage stage embryos to determine if they harboured segmental aneuploidies. GDNA from only 22 of the 41 single blastomeres subjected to SurePlex successfully amplified. Furthermore, the

VeriSeq sequencing data was suboptimal and this study was only able to analyse the presence of segmental aneuploidies in 11 of the 22 blastomeres sequenced. Three cleavage stage embryos harboured only whole chromosome aneuploidies and the rest presented segmental chromosome abnormalities. None were identified as euploid.

Although low DNA input could be in part attributed to the observed suboptimal amplification levels of single blastomeres, DNA fragmentation could play a more prominent role. SurePlex is a widely used WGA method for single blastomere and TE biopsy as a step of PGT. However, there has been a transition from cleavage stage single blastomere to blastocyst TE biopsy as the preferred embryo cell sampling method, due to increased robustness. TE biopsy is also considered less harmful for the embryo (Coll et al., 2018). The cleavage stage embryos used in this study were developmentally arrested and kept in culture until day-7 before blastomere biopsy. According to Hardy (1999), extending the culture of human arrested cleavage stage embryos for up to 72 hours increases signs of apoptosis, including DNA fragmentation. These characteristics were also found in the prolonged culture of 5- to 8 – cell cleavage stage porcine embryos (Mateusen et al., 2005). VeriSeq has also been used to test chromosomal abnormalities in single blastomeres and TE biopsies. It is, therefore, likely that extended periods in culture media caused high levels of DNA fragmentation in the arrested blastomeres, resulting in suboptimal WGA, which in turn introduced artefacts during NGS.

Furthermore, slow-growing or arrested human embryos exhibit increased levels of aneuploidy. Findikli et al. (2004) found DNA fragmentation in 54.30% of these embryos and from those that did not present DNA fragmentation, 64% were identified as chromosomally abnormal by FISH. The rest produced complex aneuploidy or uninterpretable results. More recent studies where CCS was performed on arrested

embryos have confirmed an increased level of chromosome abnormalities (in 97.80%) and a high prevalence of aneuploidy involving multiple chromosomes (Qi et al., 2014). This could explain the 100% of blastomeres found to be aneuploid in this study, the aneuploidy involving multiple chromosomes and the generalised noise in the chromosome plots.

3.1.4.2. Development of the duplex RT-qPCR method for relative gene expression analysis

Experiments were then conducted to determine the validity of the $\Delta\Delta C_t$ calculations for the relative gene expression analysis. First, the pre-amplification uniformity of the nine targets was analysed. A pre-amplification reaction for the nine targets was included in the RT-qPCR experimental design, due to the expected low concentrations of total mRNA isolated from embryos. RNA pre-amplification has been widely used to overcome the limitations when minute amounts of target RNA are found in biological samples (Mestdagh et al., 2008, Vermeulen et al., 2009). For this stage of the validation, Universal Human Reference RNA was employed since human embryo samples were limited and because the required bulk cDNA (not pre-amplified) to calculate a $\Delta\Delta C_t$ was needed in a significantly higher amount (13.50 ng cDNA input) than the RNA content in cleavage stage embryos and blastocysts. The Universal Human Reference RNA bulk cDNA input value was defined as the same DNA concentration obtained in this study from pre-amplifying cDNA of a randomly selected cleavage stage embryo. The Universal Human Reference RNA cDNA input for the pre-amplification reaction (4 ng) was defined from the average theoretical maximum total RNA obtained between a 4-cell cleavage stage embryo and a blastocyst (30 pg of total RNA in a mammalian cell). The pre-amplification uniformity analysis results demonstrated that the target pre-amplification

products are representative of the biological amounts of mRNA found in the samples. Five out of the nine targets produced values within the ± 1.5 range, recommended by the reagents manufacturer. Four targets produced values within the ± 1.68 range. However, this bias was considered dismissible.

Second, a duplex RT-qPCR assay validation was performed. The study aimed to simultaneously amplify the target and the endogenous control (*ACTB*) in each reaction. Some of the advantages of duplex RT-qPCR over singleplex are cost savings, preserving limited biological samples and reducing experimental errors. The accuracy of the duplex has been widely reported for different applications in research and diagnostics (Pal et al., 2008, Cheung et al., 2021). However, a RT-qPCR assay that amplifies more than one target simultaneously needs validation to confirm that the results obtained from duplex reactions are the same results that would be produced in singleplex. This study performed a three 10-fold dilution series of the pre-amplification products of a randomly selected blastocyst. The nine targets presented R^2 values equal or greater than 0.99, showing a good fit of the dilution points to the standard curve. A blastocyst was selected, rather than a cleavage stage embryo, to be able to obtain a dynamic range that would encompass most of the experimental samples. Since RNA was expected to be in higher quantities in blastocysts, the dynamic range obtained from a cleavage stage embryo would have excluded a significant proportion of the blastocysts. Following this strategy, the pre-amplification products of a 3-cell cleavage stage embryo were used as the representative sample to study the theoretical lower end of the dynamic range when comparing the C_t mean values between duplex and singleplex. In this regard, the present study confirmed precision between C_t mean values for all targets studied in the representative sample, since the highest variation, shown by *LIG4*, was only 3.19%. Thus,

there was no significant interference between the two reactions in the duplex across the analysed genes.

Primers with high reaction efficiencies are considered those that present values ranging from 90% to 100% (Taylor et al., 2010). In this study, only the assay for *BRCA1* produced such a value. Nevertheless, studies often report employing assays that present efficiencies above 85% (Berrada et al., 2006, França et al., 2012). *RAD51*, *RAD50* and *TP53* showed reaction efficiencies above 85%. However, *ATM, DMC1, BRCA2, XRCC4* and *LIG4* produced reaction efficiencies ranging from 76.78% to 84.14%. Low reaction efficiency can be caused by suboptimal annealing temperatures, old Taq assays, amplicons with secondary structures, poor pipetting or PCR inhibitors (Taylor et al., 2010). The Taq assays employed were pre-designed and commercially available and this study also followed the standard amplification reaction conditions defined by the manufacturer. In addition, the C_t standard deviation between triplicates was low, indicating a good pipetting technique. The TaqMan assays employed in the present study were stored at -20°C for two years. This could likely have contributed to their low amplification efficiency. A primer concentration gradient could have been carried out to determine if the concentration employed was optimal. The same strategy could be applied to the annealing temperature. Nonetheless, due to financial and time constraints, the present study moved forward with the available TaqMan assays and without further PCR conditions optimisations.

The relative gene expression of the nine targets was subsequently performed and the overall elements of the RT-qPCR assay were studied. The higher C_t mean values obtained in the cleavage stage embryos suggest a lower template cDNA input. This was expected due to the small number of cells per embryo (one to eight without accounting

for the single blastomere biopsy) compared to the blastocysts (~250) and due to the accumulation of transcripts of embryonic origin following EGA in the blastocysts. The C_t mean values were then normalised against the housekeeping gene to correct the differences in quantities of input template cDNA. Interestingly, the ΔC_t values obtained indicate a higher amount of template cDNA for all nine specific targets in the cleavage stage embryos. *ACTB* is one of the most widely used housekeeping genes for RT-qPCR controls and has been employed as a reference in mice preimplantation expression studies (Willems et al., 2004). However, more recent gene expression studies show that *ACTB* can be upregulated in blastocysts compared to MII oocytes, showing RQ values of up to eight (an eight-fold change) in relative quantification studies (Kakourou et al., 2013). This could impact the accuracy of the normalisation because the embryos used in the 3- to 8-cell blastomere group are expected to be pre- or during genome activation, when the embryo relies on maternal transcripts. In addition, early studies found that the expression of *ACTB* varies from blastomere to blastomere (Krüssel et al., 1998). This could explain the higher standard deviation of ΔC_t values between cleavage stage embryos compared to blastocysts. Due to the likely difference in expression of the endogenous control between cleavage stage embryos and blastocysts, the differences in relative expression were not compared between the two groups.

It is, therefore, evident that gene expression analysis between different stages of preimplantation development requires a careful experimental design. Increasing the number of endogenous controls can provide more accuracy in the normalisations and ΔC_t calculations (Taylor et al., 2010). Nevertheless, this approach would significantly increase the complexity of RT-qPCR multiplexing, optimisation times and costs.

3.1.4.3. Level of expression of nine genes involved in DNA repair in 3- to 8-cell embryos and blastocysts with sings of DNA damage

3.1.4.3.1. Micronucleation in 3- to 8-cell embryos

This study analysed differences in relative expression of the nine targeted genes between embryos that showed signs of DNA damage and those that did not. Micronucleation, along with segmental aneuploidy, was used in this study as a model of DNA damage in the 3- to 8-cell cleavage stage embryos. The presence of micronuclei in cells has been linked to the production of encapsulated chromosome fragments caused by unrepaired DNA damage at the time of mitosis (Krupina et al., 2021). The novel technology for embryo genome editing, CRISPR-Cas9, has demonstrated that unrepaired induced DSBs after zygote treatment with Cas9 can lead to micronuclei formation in blastomeres (Papathanasiou et al., 2021). Although the differences in relative gene expression between cleavage stage embryos with and without micronucleation groups were not statistically significant in this study, some patterns and individual sample observations were noted and are discussed below.

Wells et al. (2005b) studied the expression of genes with roles in cell cycle regulation, apoptosis and DNA repair in human preimplantation embryos at different stages of development with abnormal morphology. The authors found distinct expression patterns (presumed abnormal) in 4- to 10-cell embryos displaying granular cytoplasm and micronucleation. Regarding individual genes, upregulation of *BRCA1* was identified in day-2 and day-3 cleavage stage embryos with micronucleation. *BRCA1* has a vital role in DSBs repair by facilitating HR (Powell and Kachnic, 2003). This study identified a 6-cell embryo positive for micronucleation with almost a five-fold relative *BRCA1* expression (RQ). The mean RQ values for the groups with and without micronucleation were 1.71

and 1.51, respectively. However, this study also identified an embryo without micronucleation with an extreme RQ (5.4). The ploidy status of this embryo was unknown. It is likely that other morphological and nonmorphological factors, possibly including aneuploidy, may cause the upregulation of the gene. These findings are in line with other mammalian studies, which have suggested that HR seems to be more commonly employed in preimplantation embryos compared to NHEJ, likely because NHEJ pathway is more prone to introduce deleterious mutations (Bohrer et al., 2018). Interestingly, the present study did not find noticeable differences in the expression of XRCC4, an essential protein in NHEJ, between cleavage stage embryos with and without micronucleation, between cleavage stage embryos with and without segmental aneuploidies and between blastocysts with and without segmental aneuploidies, further supporting the above hypothesis.

In addition, Wells et al. (2005b) observed an association between high levels of embryo fragmentation and an increased expression of *ATM*, and more remarkably, *TP53*. In this study, the mean *TP53* RQ is noticeably higher in micronucleated cleavage stage embryos. As already pointed out, this study did not focus on other morphological abnormalities the blastomeres may have presented. However, cells with fragmentation often display altered DNA integrity (Findikli et al., 2004). *TP53* produces the p53 transcription factor, a protein involved in DNA damage response, repair and apoptosis (Williams and Schumacher, 2016). P53 is also activated with prolonged mitotic arrest (Uetake and Sluder, 2010). Wells et al. (2005b) suggested that fragmenting blastomeres may, therefore, be subjected to DNA damage stress that activates *TP53* and this chapter's results provide evidence to support this hypothesis. Adiga et al. (2007) showed that preimplantation embryos produced with irradiated sperm are protected by mechanisms

regulated by p21 and p53. Yet, subsequent research aiming to identify the role of p53 in early development has generated contradictory conclusions and more recent data points toward the initiation of p53-dependent transcription at the blastocyst stage (Jaiswal et al., 2021).

The mean RQ for *ATM* is also noticeably higher in micronucleated embryos. Additionally, one 5- and one 6-cell embryo of this group presented extremely high *ATM* RQ values (7.56 and 7.63, respectively) while the maximum RQ in the group negative for micronucleation was 1.05. ATM kinase is an upstream and essential regulator of the DDR activated by DSBs (Paull, 2015). This data further supports the hypothesis that micronucleated and fragmented embryonic cells are subjected to stress caused by DNA damage.

Unfortunately, the observations of the expression of *TP53* and *ATM* by Wells et al. (2005b) did not stratify the analysed embryos into preimplantation stages of development and so the level of contribution of morulae and blastocysts (when the embryo's genome is active) to such observations is unknown. However, the study highlighted a subset of 4- to 8-cell embryos that exhibited extreme upregulation of *BRCA1*. The authors speculated that these embryos possessed DNA damage from gamete origin and that it was at the 4- to 8-cell stage, when the embryonic genome is activated, that DNA damage repair pathways were activated through the upregulation of genes involved in DDR. Whether the human embryo activates its genome at the 4- to 8-cell stage to stop relying on maternal transcripts has been the subject of much follow-up research and discussion. A better understanding of this dynamic is pivotal for analysing the results produced in the present study.

Consistent with the above hypothesis, a previous study by Wells et al. (2005a), indicated that the preimplantation stage that shows the highest expression of *BRCA1, BRCA2, ATM* and *TP53* is at the 5- to 10-cell, with an abrupt decrease of total transcripts per cell from the mature oocyte to the 1- to 4-cell embryo. Part of the results in the present study supports this notion since the extreme RQ values observed for the expression of *ATM, TP53, BRCA1* and *BRCA2* correspond to 5- or 6-cell embryos. However, measuring the level of micronucleation (e.g., the percentage of blastomeres harbouring micronuclei) and the inclusion of other nonmorphological factors in the study may give a more accurate picture of the impact of varying levels of DNA damage in the upregulation of genes involved in DNA repair at different stages of preimplantation development.

3.1.4.3.2. Segmental aneuploidies in 3- to 8-cell embryos

To further support the theory that the first cleavage divisions are particularly prone to DNA damage because the embryo relies on maternal transcripts, segmental aneuploidies are particularly prevalent in the first three days of human embryonic development. Babariya et al. (2017) found that most of these errors originated during the first mitoses after fertilisation. Strikingly, the study in this chapter found that the relative expression of *ATM*, *RAD51*, *BRCA1* and *BRCA2* was significantly lower ($P<0.05$) in cleavage stage embryos that harboured segmental aneuploidies than those that did not. The embryos included for this comparison ranged from the 5- to 8-cell stage.

Wells et al. (2005b) found that embryos that exhibited the highest number of *BRCA1* transcripts had minimal or no cell fragmentation, likely due to the repressive role of the accumulation of p53, caused by the stress originated from cell fragmentation, on

BRCA1. However, the RQ values of *TP53* are in the present study are not higher in blastomeres with segmental aneuploidies than those without, making this explanation for the downregulation of *BRCA1* in the first group of blastomeres unplausible. On the other hand, studies have identified an increase in aneuploidy in *Atm* $^{-/-}$ mice fibroblasts (Shen et al., 2005) and in *ATM* small interfering RNA (siRNA)-injected HeLa cells (Palazzo et al., 2014). Additionally, some authors have proposed that the increased chromosome abnormalities observed in cleavage stage embryos can be caused by inadequate resourcing of maternal factors needed for DNA repair pathways (Fragouli et al., 2011, Babariya et al., 2017). Nevertheless, it is improbable that the group of cleavage stage embryos with segmental aneuploidies exhibited a generalised downregulation of the studied genes and that this caused the structural abnormalities (albeit some comparisons lack statistical significance). Another explanation may be that segmental aneuploidies could have been erroneously identified or missed in cleavage stage embryos due to the noisy chromosome copy number plots caused by fragmented blastomere gDNA. In addition, the number of samples included in each group was very small and may not be statistically representative. Furthermore, the average number of blastomeres in the group of embryos with segmental aneuploidies was 6.83 while the average for the group with only whole chromosome aneuploidies was 7.25. The second group would probably have accumulated more transcripts from embryonic origin upon genome activation. As previously discussed, future studies could include several quantitative variables indicative of DNA damage (morphological and nonmorphological) to study the effect of DNA damage more accurately.

As mentioned, the hypothesis that human EGA occurs between the 4- and 8-cell stage has been recently brought into question after evidence that embryo genome

transcription initiates as early as the 1- or 2-cell stage (Piras et al., 2014, Asami et al., 2022). However, the major wave of EGA is during the 4- to 8-cell stage. Thus, stage-specific patterns of embryonic gene expression have been suggested instead of a 'single moment activation' where different sets of genes are activated at distinct stages of development (Jukam et al., 2017). Asami et al. (2022) found upregulation in healthy 2PN-stage human embryos of genes with roles in developmental processes, including *ATM* activation and chromosome segregation. This, according to the authors, is consistent with observations of arrest or elimination of aneuploid 2-cell embryos, since the activation of *ATM* can lead to G2/M phase arrest or apoptosis if DSBs are not repaired (Turan and Oktay, 2020). Nevertheless, the arrest could also result from the activity of gene products of maternal origin that are present in sufficient quantities in cleavage stage embryos (Betts and Madan, 2008).

3.1.4.3.3. Segmental aneuploidies in blastocysts

This chapter also aimed to study the DNA damage pathways at the blastocyst stage, which follows EGA. None of the gene RQ comparisons between blastocysts that harboured only whole chromosome aneuploidies and those that presented segmental aneuploidies were statistically significant. Nevertheless, some outliers were identified. It is currently hypothesised that all DNA repair pathways exist in the oocyte and the blastocyst. However, the beginning of embryo cell differentiation in the blastocyst is marked by a change in the expression profile of DNA repair genes to be capable of responding to DNA damage in a similar way to somatic cells (Khokhlova et al., 2020). This stage is also characterised by the acquired ability of the embryo to eliminate chromosomally abnormal cells, believed to be a selection process, via apoptosis (Babariya

et al., 2017). One key protein in somatic cells for safeguarding genome integrity is p53. However, as previously mentioned, the role of p53 in early embryo development has been the subject of much discussion. It is now believed that this transcription factor is detected only during and after the blastocyst stage and is not pivotal for embryonic development (Jaiswal et al., 2020). Consistent with this, there were no noticeable *TP53* expression level embryo outliers with segmental aneuploidies that pointed towards increased activity of p53 transcription factor caused by DNA damage. However, it is important to mention that p53 has a role in the regulation of the spindle assembly checkpoint (SAC), but studies with somatic cells have found that numerical aneuploidy, unlike structural aneuploidy, does not cause p53 activation (Soto et al., 2017).

The activity of other genes with important roles in DSBs repair has also been studied in preimplantation embryos. Studies with porcine embryos have found that the disruption of *RAD51* activity with RNA interference (RNAi) in parthenotes induces DNA damage and causes apoptosis (Jin et al., 2019). Moreover, that the expression of the gene is upregulated in 4- to 8-cell embryos, due to the induction of DSBs by vitrification earlier in development. These embryos managed to reach the blastocyst stage, suggesting that an increased activity of *RAD51* was required to repair DNA damage successfully (Somfai et al., 2023), but the expression of such a gene had not been directly studied in blastocysts. The identification of outliers in the blastocyst group with segmental aneuploidies in this study suggests that the embryo still employs a mechanism that involves *RAD51* for HR. The effects of a lack of BRCA1 and BRCA2 in embryo development, associated proteins of RAD51, have been studied with mouse knockout models and early embryo lethality is observed in both (Moynahan, 2002). The same phenotype is observed for *Rad51* $^{-/-}$ (Lim and Hasty, 1996). The disruption of *BRCA1* and

BRCA2 has not been directly studied in oocytes and preimplantation embryos, unlike with *RAD51*. Nevertheless, this study found blastocysts that harboured DNA damage with upregulation of these three genes. Previous data and the presented in this chapter support the hypothesis that *RAD51*, *BRCA1* and *BRCA2*, like in somatic cells, play a role in DSBs repair via HR in blastocysts post-EGA, possibly making the blastocyst more efficient in DDR compared to cleavage stage embryos.

3.1.4.4. Study limitations

The study design in this chapter presents some limitations that may impact the interpretation of the results. Whilst this chapter has already discussed some limitations, additional ones are presented below.

Firstly, in terms of post-zygotic DNA damage models, micronucleation is not always caused by chromosome fragments from unrepaired DNA. Lagging chromosomes that fail to incorporate into the daughter nuclei during mitosis can also cause this morphological feature (Krupina et al., 2021). Therefore, segmental aneuploidy is a more accurate model for chromosome breakage.

Secondly, this chapter did not study the presence of mosaicism in the embryos due to technical constraints and because a single blastomere was analysed from the cleavage stage embryos. Although segmental aneuploidy occurs primarily due to mitotic errors, mosaicism is an indicator of post-zygotic chromosome breakages and segregation errors, providing supplementary information about the origin of aneuploidy.

Thirdly, regarding embryo sampling, the sample size of embryos at or below the 4-cell stage was five. As already discussed, the major EGA wave starts after the 4-cell stage

and statistically representative sample sizes are required for an accurate analysis of the true biological differences between the two preimplantation stages.

Fourthly, ideally, control groups would also include morphologically and chromosomally normal embryos since many genes that have roles in DDR also play a role in chromosome segregation surveillance.

The last two points, however, depend much on the availability of human embryos for research. Future studies could also employ animal models, such as the mouse, to overcome such difficulty. Finally, the present study was unable to distinguish between maternal transcripts and embryo transcripts. This last limitation is commonly reported in gene expression studies in preimplantation embryos.

3.1.4.5. Conclusion

The DDR between somatic cells and preimplantation embryos seems to differ significantly. In fact, these pathways seem to vary between different stages of the preimplantation development. Studies have suggested that EGA, which has been long believed to occur between the 4- to 8-cell stage, marks a switch between embryos relying on maternal transcripts and embryos producing their own proteins to respond to such damage, making embryos prior EGA more susceptible to aneuploidy. This study aimed to better understand the DNA damage repair mechanisms in preimplantation development by, for the first time, studying the differences in the expression of genes with roles in DNA repair at different stages of development and between embryos that harboured DNA damage and those that did not.

This study provided data that supports previous observations about the upregulation of HR pathway DNA damage repair genes in blastomeres that exhibit morphological features suggestive of DNA damage. It found a generalised, albeit statistically insignificant, upregulation of *TP53* in 3- to 8-cell embryos. It also found instances of upregulation of *ATM* in 5- to 6- cell embryos and of *BRCA1* in a 6-cell embryo that harboured micronuclei, suggesting that the embryo is capable of responding to DNA damage very early into development. Recent evidence indicates that the embryo initiates its transcription in a stage-specific pattern and that the first transcripts are detected at the 1- or 2-cell stage, including those for *ATM* activation. In addition, this chapter provided evidence to support the hypothesis that the HR pathway is more commonly employed, compared to NHEJ by early embryos.

Although this and other studies' findings suggest that embryos activate their own DNA repair pathways (albeit likely sub optimally), either before, or at the earliest stages of the major EGA wave, more transcriptomic research using high-resolution RNA-seq time courses (to distinguish maternal from embryonic transcripts) is needed as well as RNA transcription inhibition corroboration studies. This research should cover human (or animal, if unavailable) embryos at different stages of development, including 2PN and 2-cell stages, with statistically powerful sample sizes with varying degrees of DNA damage.

Chapter 4. Aneuploidy analysis of *Atm* knocked-down mice parthenotes

This chapter is based on joint work for a pilot study with Professor Kutluk Oktay, Ph.D. Katarzyna Szymanska-Vandendriessche and Ph.D. Xiu Tan from Yale University, School of Medicine, Department of Obstetrics, Gynaecology and Reproductive Sciences in Connecticut, USA.

4.1.1. Introduction

4.1.1.1. Female reproductive ageing

While several variables contribute to infertility, advanced female reproductive age, considered to be over 35, is a powerful negative predictive factor (Hart, 2016). Female reproductive ageing is distinguished by the loss of quantity and, more importantly, quality of oocytes. Bad-quality female gametes increase the risk of miscarriage and birth defects (Duncan and Gerton, 2018). A body of research based on clinical data obtained from IVF centres confirms the above by showing that aneuploidy increases predictably with women's age (Franasiak et al., 2014). The association between women's reproductive age and oocyte quality has prompted further research into the molecular mechanisms behind the effect of female ageing on gametes. These studies show that chromosome mis-segregation during oocyte meiosis plays a crucial role (Charalambous et al., 2023). This chapter aims to add to our understanding of these molecular mechanisms by exploring the role of ataxia telangiectasia mutated (ATM), an extensively studied protein that plays a key role in somatic cell DDR. Before introducing

the study aims, the introduction explains the role of oocyte meiotic DSBs repair and chromosome segregation in aneuploidy.

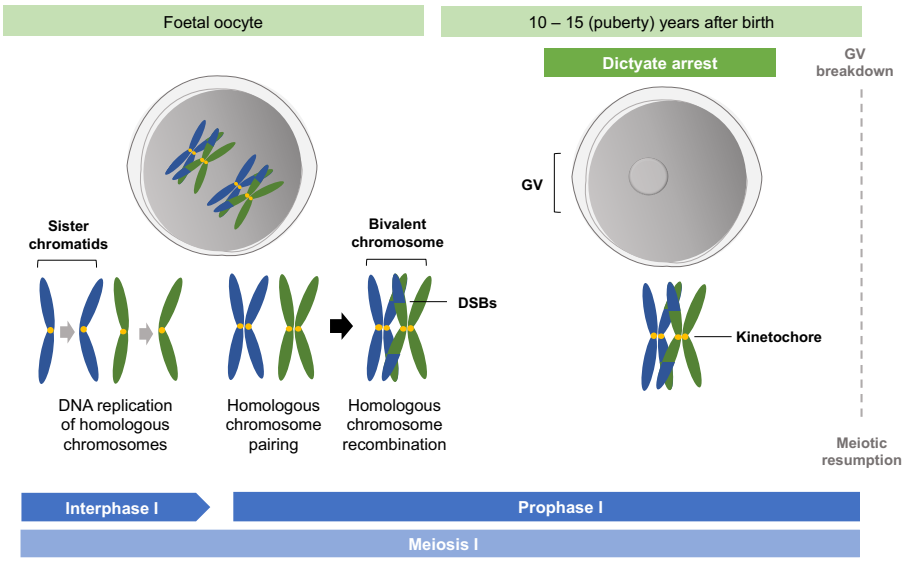
4.1.1.2. Meiosis in oogenesis

Unlike somatic cells, germ cells must reduce their chromosome content in half prior to fertilisation. To achieve this, cells undergo two rounds of asymmetric cell divisions: meiosis I and II (Figure 15).

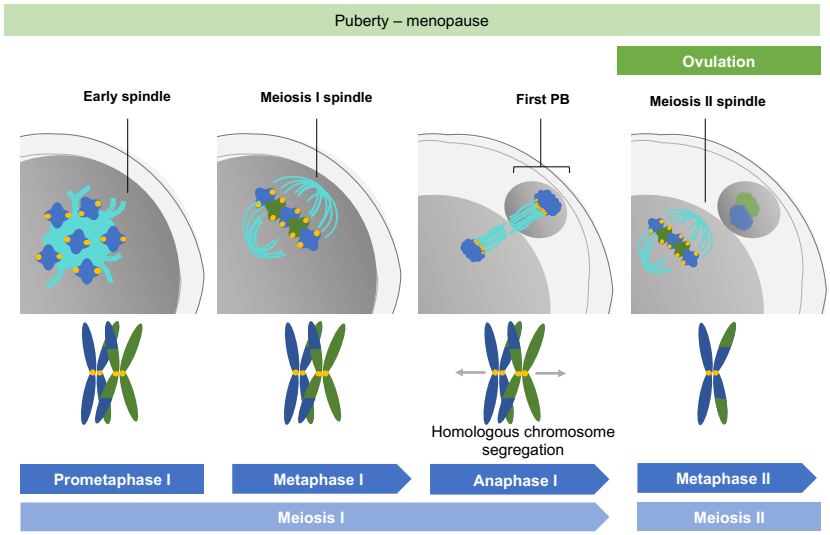
4.1.1.2.1. Chromosome recombination and meiotic arrest

During human foetal development primordial germ cells (PGCs) undergo a single round of DNA replication to produce two sister chromatids per homologous chromosome. Sister chromatids are held together by the cohesin protein complex (Haering et al., 2008). Meiosis I begins with prophase I, where the homologous chromosomes pair, a process known as synapsis, and through programmed DSBs, recombination takes place with maternal and paternal DNA strands crossing over (Lam and Keeney, 2015). Cohesins also have a role in holding the two recombined homologous chromosomes together, allowing for the constitution of a new structure known as a bivalent chromosome. Oocytes subsequently become dormant in meiosis I, a state known as dictyate arrest (Charalambous et al., 2023) (Figure 15(A)).

(A)



(B)



(C)

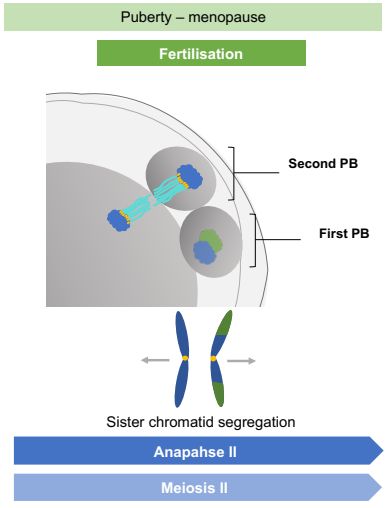


Figure 15. Human oocyte meiosis. Oocyte meiosis I begins in the foetus and it is not completed until the first cell division close to ovulation that results in the extrusion of the first PB. Meiosis II starts immediately after and culminates following fertilisation with the second PB extrusion (Charalambous et al., 2023).

4.1.1.2.2. Double-strand breaks (DSBs) and aneuploidy in oocyte meiosis

The multiple induced DSBs during recombination are mostly repaired through HR, the preferred DNA damage repair in oocytes, known in meiosis as 'crossovers repair'. This is a highly stringent process regulated by checkpoint proteins (Gray and Cohen, 2016). However, oocytes can accumulate spontaneous DSBs damage during the prolonged dictyate arrest resulting from oxidative stress and maternal ageing. The *RAD51* and *BRCA2* genes are crucial for DSBs repair through HR, and transcripts from both genes have been identified in human and mice primordial follicle oocytes. Interestingly, the expression of *Rad51* diminishes in oocytes from aged female mice, suggesting a decline in DSBs repair efficiency with age (Titus et al., 2013). Unlike programmed DSBs, prophase I-arrested oocytes have a low tolerance for DNA damage and female gametes are eliminated through DDR mechanisms if they reach a maximum threshold of unrepaired DSBs (Rinaldi et al., 2017). Failure to repair DSBs in oocytes can lead to chromosome breakages and segmental chromosomal abnormalities (MacLennan et al., 2015, Turan and Oktay, 2020). Despite this, unlike with numerical aneuploidy, studies have not found a correlation between maternal age and an increase in chromosome structural abnormalities in the embryo (Babariya et al., 2017).

4.1.1.2.3. Meiosis I and II

The germinal vesicle (GV), or prophase I-stage arrested oocyte, undergoes nuclear envelope breakdown close to ovulation. This event marks the beginning of the first cellular meiotic division. The oocyte begins prometaphase I, where bivalents slowly, and after several attempts, manage to align in the spindle's centre. Bivalents must be correctly oriented by late metaphase I (Kitajima et al., 2011). Each homologous

chromosome locates at one of the two spindle poles. The spindle and chromosomes subsequently migrate to the oocyte cortex at the start of anaphase I (Charalambous et al., 2023). Homologous chromosomes segregate (one complement is retained and the other is extruded) during anaphase I, which culminates in the first meiotic division resulting in the extrusion of the first PB. The oocyte, now arrested in metaphase II from meiosis II (MII oocyte), is ready for fertilisation (Figure 15(B)). Anaphase II is triggered by fertilisation to resume meiosis II. Sister chromatids segregate with a second meiotic spindle (one of the sister chromatids is retained and the other is extruded), resulting in the extrusion of the second PB (Figure 15(C)).

4.1.1.2.4. Chromosome mis-segregation and aneuploidy

The aetiology of aneuploidy is complex. Meiosis I and II are vulnerable stages of oocyte maturation where chromosome segregation errors may occur that result in aneuploidy. It is well established that numerical aneuploidy increases as age progresses. Nevertheless, oocytes have a natural vulnerability for chromosome mis-segregation due to the unique features of the oocyte meiotic spindle and the decreased stringency of surveillance mechanisms (checkpoints) compared to mitotic divisions (Charalambous et al., 2023).

Cohesins are of critical importance in human oocytes due to their function in maintaining bivalent chromosomes and sister chromatids close together for decades during the dictyate arrest, and until completion of meiosis I and II. Studies in mice have found that an age-dependent cohesion strength loss in oocytes (Duncan et al., 2012, Zielinska et al., 2019) is the leading cause of chromosome segregation errors associated with advanced maternal reproductive age (Mikwar et al., 2020). The deterioration of

cohesion can cause bivalent dissociation before anaphase I and PSSC that can affect meiosis I and II (Charalambous et al., 2023).

Female advanced reproductive age may also affect meiosis cell cycle checkpoints. The SAC is a surveillance mechanism that safeguards the correct kinetochore attachment to the spindle during chromosome segregation. If the SAC detects unattached kinetochores and a lack of tension, it delays the anaphase entry by inhibiting the anaphase-promoting complex (APC) activity until correcting attachments (Yang et al., 2011). Although SAC operates in mitosis and meiosis, its activity is less stringent in meiosis, contributing to oocyte predisposition to aneuploidy (Lane et al., 2017b). Human data suggest maternal ageing can lead to the loss of SAC core proteins expression, including MAD2, BUB1, BUB3 and BUBR1, exacerbating its lack of strictness (Mikwar et al., 2020).

4.1.1.3. Ataxia telangiectasia mutated (ATM)-mediated DSBs repair

DDR pathways and mechanisms for faithful chromosome segregation are two systems that work jointly during oogenesis to prevent the generation of oocytes with aneuploidy. For instance, in mice, excessive accumulation of DSBs causes SAC-mediated metaphase I arrest of primary oocytes in order to prevent this damage from influencing cellular processes, such as accurate chromosome segregation (Collins et al., 2015, Charalambous et al., 2023). Yet, findings show that this checkpoint is impaired in aged female mice (Marangos et al., 2015). Evidence also shows that the induction of DSBs in aged oocytes provokes a decline in BUB1 and BUBR1 localisation at kinetochores. But these oocytes managed to progress to meiosis I and meiosis II. The above exemplifies the link between age-related decline in oocyte competence, DNA damage repair and meiotic

chromosome segregation. The loss of essential core players in one of these systems could have downstream effects on the accurate segregation of chromosomes in meiosis I and II and chromosome integrity.

As previously mentioned, BRCA2 plays a vital role in oocyte DSBs repair. Along with BRCA1, BRCA2 is heavily involved in HR. These two proteins are an essential part of the ATM-mediated DNA DSBs repair pathway that tightly regulates HR (Turan and Oktay, 2020). ATM activates after sensing DSBs, by the MRN (MRE11, RAD50 and NBS1) complex, to trigger a wide network of responses (Titus et al., 2013). In somatic cells, about which we know more compared to gametes, the activation of ATM can then lead to three outcomes: the successful repair of DSBs, the trigger of an apoptotic pathway due to acute DNA damage or cell senescence (Turan and Oktay, 2020). In addition, it has been shown that in somatic cells, SMC1, a sub-unit of the cohesin complex, is phosphorylated by ATM kinase following ionising irradiation, suggesting that this protein is an important substrate of the ATM-mediated DSBs pathway (Kim et al., 2002), and illustrating yet again the inter-relation between DSBs repair and chromosome segregation.

Data obtained from human GVs found that the expression of genes involved in the ATM-mediated DSBs repair pathway, including *ATM, BRCA1, MRE11* and *RAD51*, significantly decreases in women who are at an advanced reproductive age (Oktay et al., 2015). Existing data suggests that ATM may have a regulatory role in aneuploidy (Yang et al., 2011, Xiao et al., 2022). However, there is limited experimental work to provide direct evidence of an association between ATM-mediated DSBs repair pathway dysfunction and impaired oocyte quality.

4.1.1.4. Study aims

This pilot study aims to analyse the effects of oocyte *ATM* downregulation in oocyte capacity for DNA damage repair and faithful meiotic chromosome segregation, to shed light on its role in oocyte competence. To achieve this, *Atm* was downregulated by gene knockdown in mice GVs and one-pronuclear (1PN) stage *Atm* knocked-down parthenotes were subjected to chromosome number and integrity analysis. The existing literature was also reviewed to study the role of ATM in mice oocyte meiotic chromosome segregation and DSBs repair.

The specific aims are:

1) To study the effect of *Atm* silencing in GVs in chromosome number and integrity in the *Atm* siRNA-injected parthenote group.
2) To study the effect of the microinjection of GVs in chromosome number and integrity in the scrambled siRNA-injected parthenote group.
3) To perform a statistical power analysis to determine how reliable the present study design is in detecting a true biological difference in the level of overall aneuploidy between the *Atm* siRNA-injected, scrambled siRNA-injected and uninjected groups.

4.1.2. Materials and methods

4.1.2.1. Ethics statement

The animals were handled at Yale University, USA, in accordance with the Guide for Care and Use of Laboratory Animals of the USA National Institutes of Health (NIH) and the Yale University Institutional Animal Care and Use Committee (IACUC).

4.1.2.2. Experimental design

Twenty-eight GVs were injected with *Atm* siRNA for the experimental *Atm* knocked-down group. For the two control groups, 24 GVs were injected with scrambled siRNA, as injection controls, and 12 were uninjected (Figure 16) producing a total of 64 mice parthenotes.

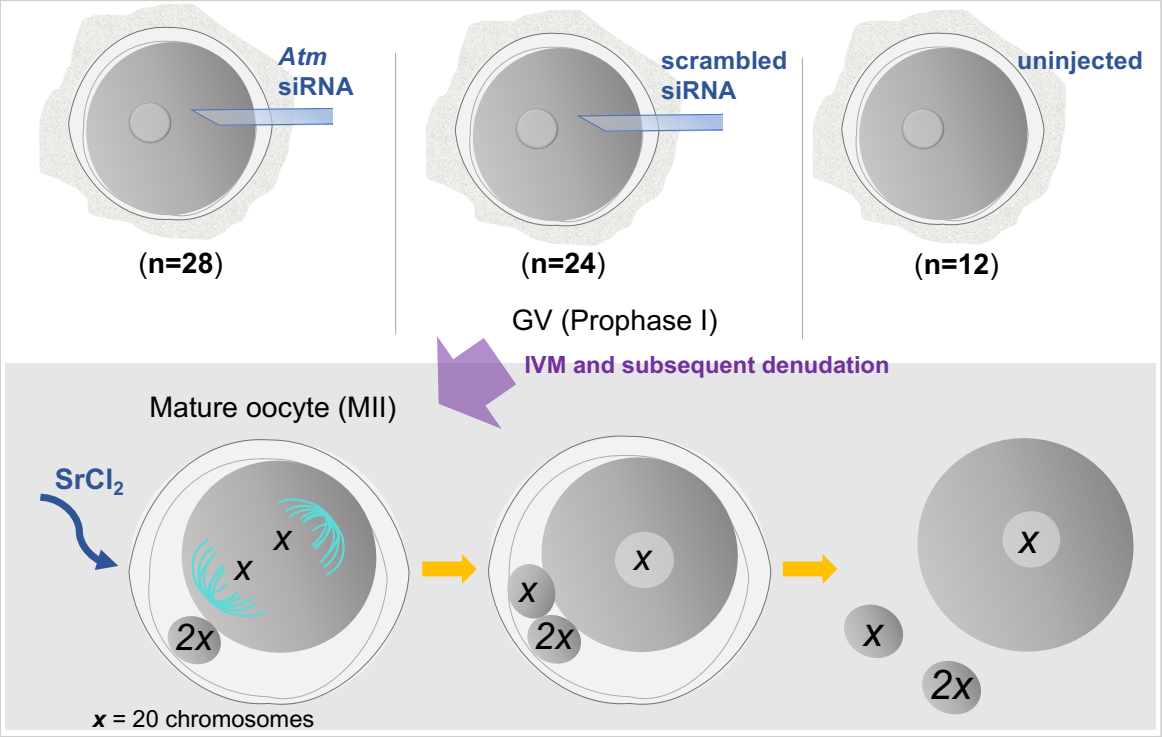


Figure 16. Experimental design for the aneuploidy analysis of *Atm* knocked-down mice parthenotes. Twenty-eight mice GVs were injected with *Atm* siRNA, 24 with scrambled siRNA and 12 were uninjected controls. COCs were subsequently subjected to in vitro maturation (IVM). Cumulus cells were removed afterward. Parthenotes were produced after subjecting the matured oocytes to chemical parthenogenic activation using strontium chloride ($SrCl_2$). $SrCl_2$ activation allows for the second PB extrusion producing an haploid 1PN embryo. The zona pellucida and the PBs were subsequently removed from the parthenotes. MII, metaphase II.

4.1.2.3. Mouse cumulus-oocyte complexes (COCs) recovery, *Atm* knockdown, in vitro maturation (IVM), parthenogenesis and WGA

The following procedures were performed by the Professor Kutluk Oktay's research group at Yale University.

Six- to 8-week-old female FVB mice underwent ovarian hormonal stimulation after injecting them with 10 IU of pregnant mare's serum gonadotrophin (PMSG). Forty-eight hours after the injection, COCs were collected by repeated follicular puncture with a fine needle. GVs were then microinjected with *Atm* siRNA and scrambled siRNA at a needle concentration of 50 µM. In a separate study, *Atm* knockdown using the present protocol was confirmed by single-cell relative RT-qPCR for *Atm* mRNA using the uninjected GVs as a calibrator. The RQ values were 0.0004 for the *Atm* siRNA-injected, 0.7330 for the scrambled siRNA-injected and 1.00 for the uninjected controls ($P<0.0005$, one-way analysis of variance (ANOVA)). The efficiency of the *Atm* gene knockdown was further confirmed with confocal microscopy to visualise the ATM protein. The protein mean fluorescence intensity (MFI) was 7.0, 13.9 and 13.2 for the *Atm* siRNA-injected, scrambled siRNA-injected and uninjected controls respectively ($P<0.05$, one-way ANOVA). COCs underwent 20 hours of in vitro maturation (IVM) and MII oocytes were removed from the cumulus cells and subjected to chemical parthenogenetic activation with strontium chloride ($SrCl_2$). 1PN stage haploid parthenotes were collected after the second PB extrusion and the zona pellucida was removed. Some second PBs were recovered (Figure 16). The parthenotes and PBs were lysed, subjected to SurePlex WGA and the products were sent to Oxford and stored at 4 °C until library preparation.

4.1.2.4. Aneuploidy analysis of parthenotes

The parthenote NGS-based aneuploidy analysis was carried out using the Oxford Nanopore MinION platform. The library preparation protocol followed was Ligation Sequencing gDNA – PCR Barcoding (SQK-LSK109) along with the PCR Barcoding Expansion Pack 1 – 96 (EXP-PBC096) according to **Chapter 2. General materials and methods**. Due to the high number of samples, these were divided into two groups and sequenced in two different runs.

The sequencing data was analysed by Ph.D. Clement Coudereau with a pipeline previously developed by Ph.D. Nada Kubikova for murine aneuploidy analysis using the MinION Oxford Nanopore platform. Briefly, the demultiplexed FASTQ files containing each sample's reads were aligned to the *Mus musculus* genome (GRCm38_68) to generate BAM files. The mapped reads were counted per chromosome using the CountsRead function from the Genome Analysis Tool Kit (Broad Institute, USA). The total number of reads mapped per chromosome was cross-referenced against multiple samples from the same cohort with a normal karyotype. The bin size used for the present bioinformatical analysis was 5 Mb. These values were subsequently plotted to visualise the predicted copy number per chromosome.

Reads that aligned to the '0' value of the Y axis of the chromosome plot meant a loss of a chromosome, reads aligned to '1' were considered normal copy number and alignment to '2' indicated a chromosome gain, since the samples included were haploid. Segmental aneuploidy was identified when a segment of the total reads from a chromosome was aligned either to '0' or '2', but this study was unable to determine location of chromosomal breakpoints.

4.1.2.5. Statistical analysis

Due to the small sample size, 95% confidence intervals for the percentage of parthenotes with aneuploidy were calculated with the Wilson/Brown method. Statistical comparisons between experimental groups were carried out using a two-tailed Fisher's exact test, with a 5% significance level. GraphPad Prism software was used for statistical analysis and graph drawing.

A post hoc power analysis with a level of significance of 5% was performed with the preliminary results obtained from this study employing the following equation (Rosner, 2011):

$$Power = \emptyset \left[\frac{\Delta}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} - z_{1-\alpha/2} \frac{\sqrt{PQ(\frac{1}{n_1} + \frac{1}{n_2})}}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} \right]$$

$$q_1 = 1 - p_1$$

$$q_2 = 1 - p_2$$

$$P = \frac{p_1 + k p_2}{1 + K}$$

$$Q = 1 - P$$

Where:

$p_1, p_2 = incidence\ of\ groups\ 1\ and\ 2$

$\Delta = |p_2 - p_1| = absolute\ difference\ between\ the\ two\ porportions$

$n_1, n_2 = sample\ size\ for\ groups\ 1\ and\ 2$

$\alpha = probability\ of\ type\ I\ error = 0.05$

$z = critical\ Z\ value\ for\ a\ given\ \alpha\ (for\ 5\%\ level\ of\ significance, z_{1-\alpha/2} = 1.96)$

$K = ratio\ of\ sample\ size\ for\ group\ 2\ to\ group\ 1$

$$\emptyset = \ function\ to\ convert\ a\ critical\ z\ value\ to\ power$$

4.1.3. Results

4.1.3.1. WGA and sequencing success rates

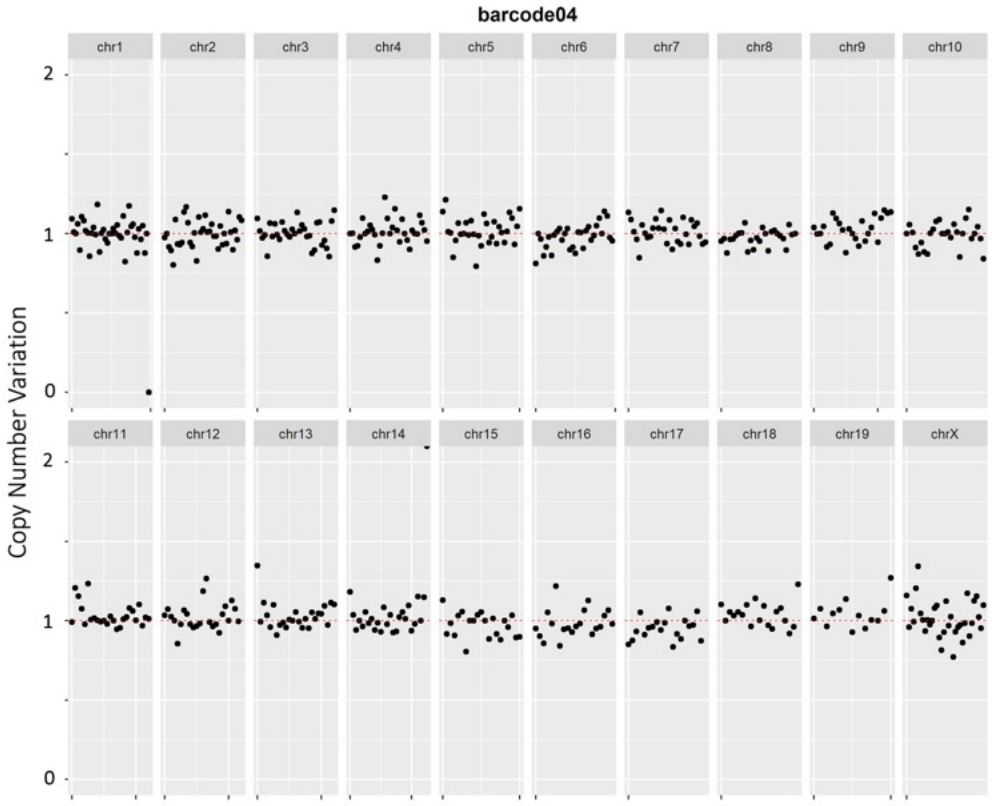
All 28 samples recovered from the *Atm* siRNA-injected GVs were parthenotes. For the scrambled siRNA-injected GVs, 24 parthenotes were recovered and two PBs. Twelve parthenotes were obtained from the uninjected GVs and one PB. After SurePlex, all parthenotes and PBs (67/67) showed above the minimum DNA concentration required to be processed for library preparation (30 ng/µl). However, two out of the 12 uninjected parthenotes failed to achieve the minimum sequencing quality score to be included for aneuploidy analysis.

4.1.3.2. Aneuploidy analysis

Two representative parthenotes are shown in Figure 17 to illustrate the predicted chromosome copy number plots generated after the Oxford Nanopore MinION sequencing. **Appendix B** shows the results for all the samples included in the analysis.

The aneuploidy analysis results obtained from the 62 parthenotes revealed close to triple the overall aneuploidy incidence in the *Atm* knocked-down group compared to the uninjected (57.14% vs. 20%, respectively). From the uninjected parthenotes, 10% showed whole chromosome aneuploidies, compared to 42.85% of the *Atm* knocked-down group. The difference was less significant with segmental aneuploidies (10% vs. 25%) (Figure 18(A)).

(A)



(B)

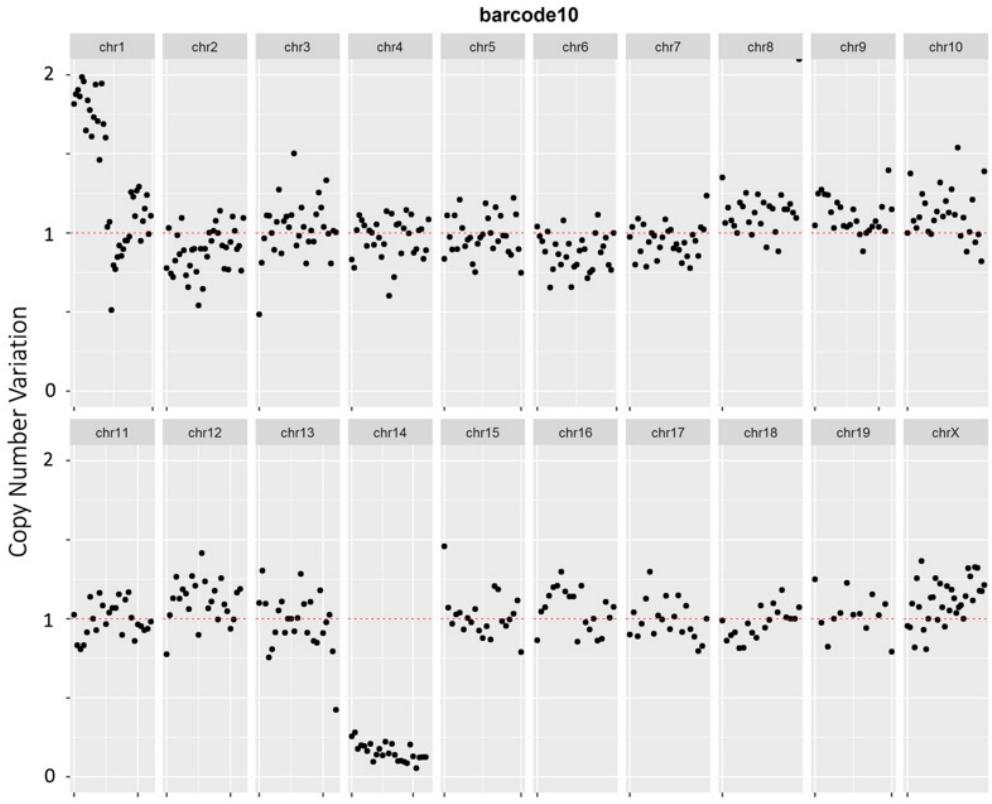
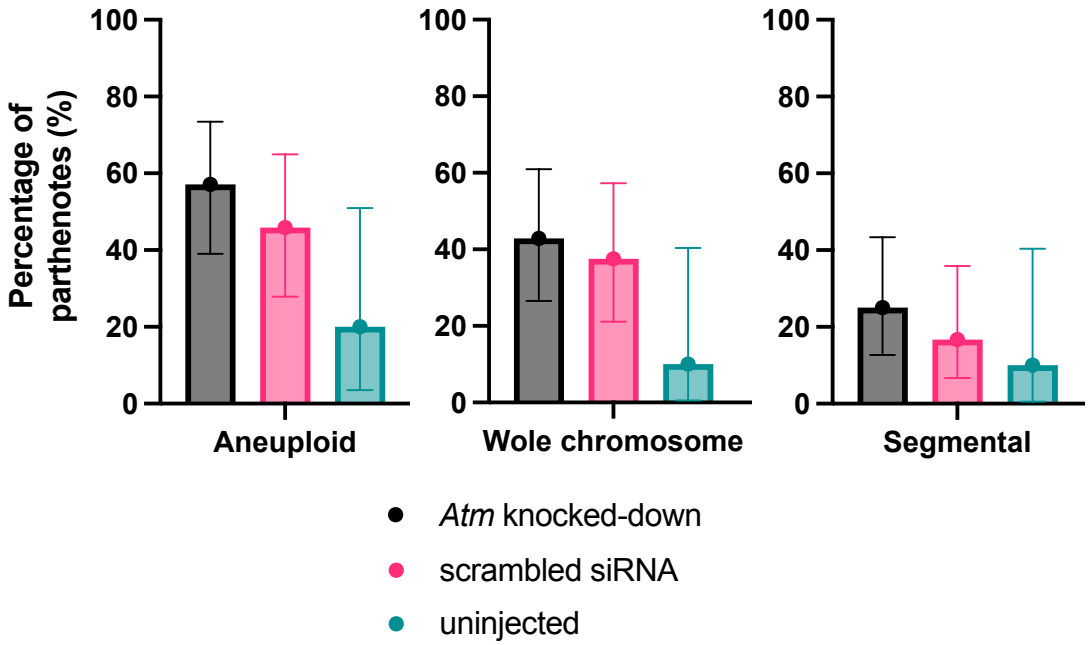


Figure 17. Plots of the predicted chromosome copy number in two representative mice parthenote samples. (A) Karyotypically normal haploid parthenote (20,X), (ID: A4). **(B)** Haploid parthenote with a segmental gain in chromosome 1 and a chromosome 14 loss (19,X,+1(seg),-14) (ID: A10). Each dot plotted represents a 5 Mb bin.

(A)



(B)

NGS results	*Atm* knocked-down (n=28)	Scrambled siRNA-injected (n=24)	Uninjected (n=10)
Aneuploid, n, %, (CI)	16, **57.14%** [a,c] (39.07 – 73.49%)	11, **45.83%** [b,c] (27.89 – 64.92%)	2, **20%** [a,b] (3.55 – 50.98%)
Aneuploid (Whole chr.), n, %, (CI)	12, **42.85%** [d,f] (26.51 – 60.93%)	9, **37.50%** [e,f] (21.16 – 57.29%)	1, **10%** [d,e] (0.51 – 40.41%)
Aneuploid (Seg.), n, %, (CI)	7, **25%** [g,i] (12.68 – 43.36%)	4, **16.67%** [h,i] (6.68 – 35.85%)	1, **10%** [g,h] (0.51 – 40.41%)

a) P=0.07 (n.s.); b) P=0.25 (n.s.); c) P=0.58 (n.s.); d) P=0.12 (n.s.); e) P=0.22 (n.s.); f) P= 0.78 (n.s.); g) P=0.65 (n.s.); h) P>0.99 (n.s.); i) P=0.52 (n.s.).

Figure 18. Comparison of aneuploidy rates in mice parthenotes from *Atm* knocked-down, scrambled siRNA-injected and uninjected control oocytes. According to Fisher's exact test, the differences in rates observed were not statistically significant for all cases. Confidence intervals were calculated with the Wilson/Brown method. CI, confidence interval.

When comparing the parthenotes from scrambled siRNA-injected oocytes to the uninjected controls, there is also a noticeable higher incidence of overall aneuploidies (45.83% vs. 20%, respectively). However, the difference is smaller than that of the *Atm* knocked-down and uninjected controls. The difference between scrambled siRNA-

injected and uninjected groups for whole chromosome aneuploidies was 37.50% vs. 10% and 16.67% vs. 10% for segmental abnormalities, respectively (Figure 18(A)). However, none of the observed rate differences were statistically significant (P>0.05 for all comparisons). The confidence interval ranges showed a high level of uncertainty due to the small number of samples analysed (Figure 18(B)).

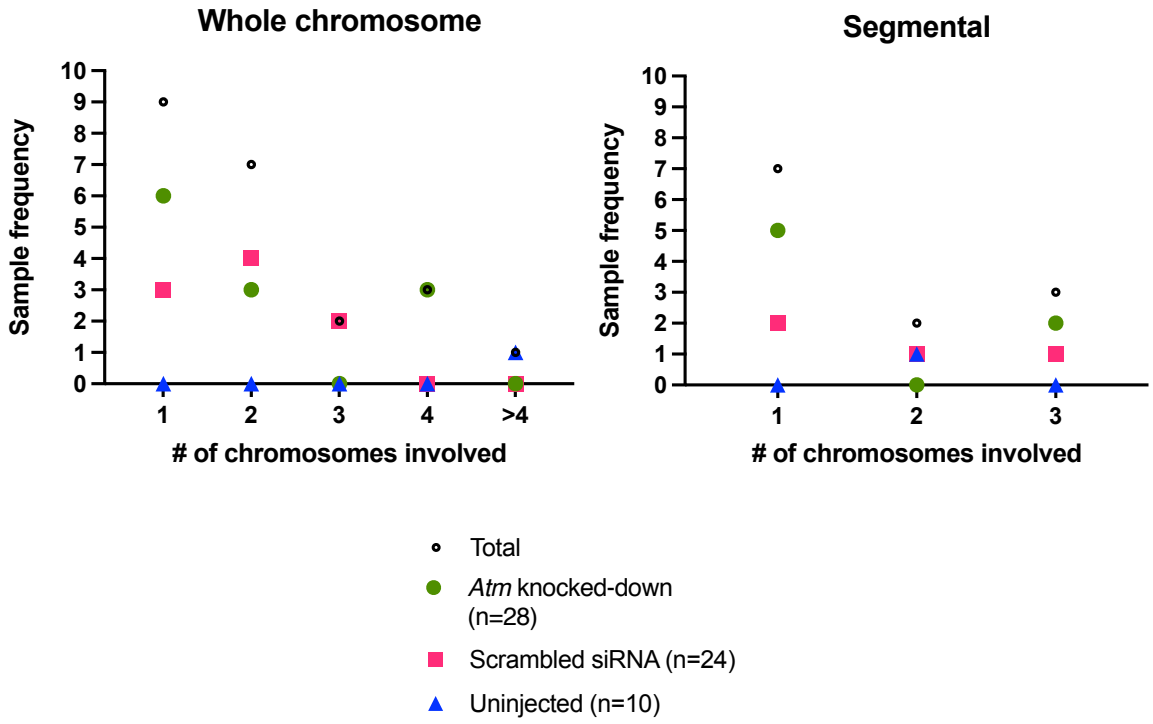


Figure 19. Number of chromosomes with numerical and segmental aneuploidies per parthenote.

Regarding the number of chromosomes with abnormalities per parthenote (Figure 19), six of the total *Atm* knocked-down samples (21.43%) and three of the total scrambled siRNA-injected control (12.50%) presented one whole chromosome gain or loss. The percentages involving numerical aneuploidies in two chromosomes and three or more chromosomes were both 10.71% for the *Atm* knocked-down parthenotes and 16.67% and 8.33% for the scrambled siRNA-injected control. The only parthenote with

whole chromosome aneuploidies in the uninjected group presented aneuploidy in three or more chromosomes (10%).

Five parthenotes from the *Atm* knocked-down group presented segmental aneuploidies in one chromosome (17.86%) compared to two samples from the scrambled siRNA-injected group (8.33%). One sample from the latter group (4.17%) and one sample from the uninjected group (10%) presented segmental aneuploidies in two chromosomes. Two samples from the *Atm* knocked-down group (7.15%) and one sample from the scrambled siRNA-injected group (4.17%) presented segmental aneuploidies in three chromosomes.

This study also analysed the frequency of chromosome-specific aneuploidies in parthenotes to determine if some chromosomes were more affected. As shown in Figure 20, including the three experimental groups, whole chromosome aneuploidies were most common in chromosome 6 (observed six times), followed by chromosome X (observed five times). However, numerical aneuploidies occurred in all the chromosomes except 4 and 20. The most common chromosomes to present numerical aneuploidies for the *Atm* knocked-down parthenotes were 5 and 10 (observed three times). For the scrambled siRNA-injected group, chromosomes 6 and X were the most common numerical aneuploidies (observed four times).

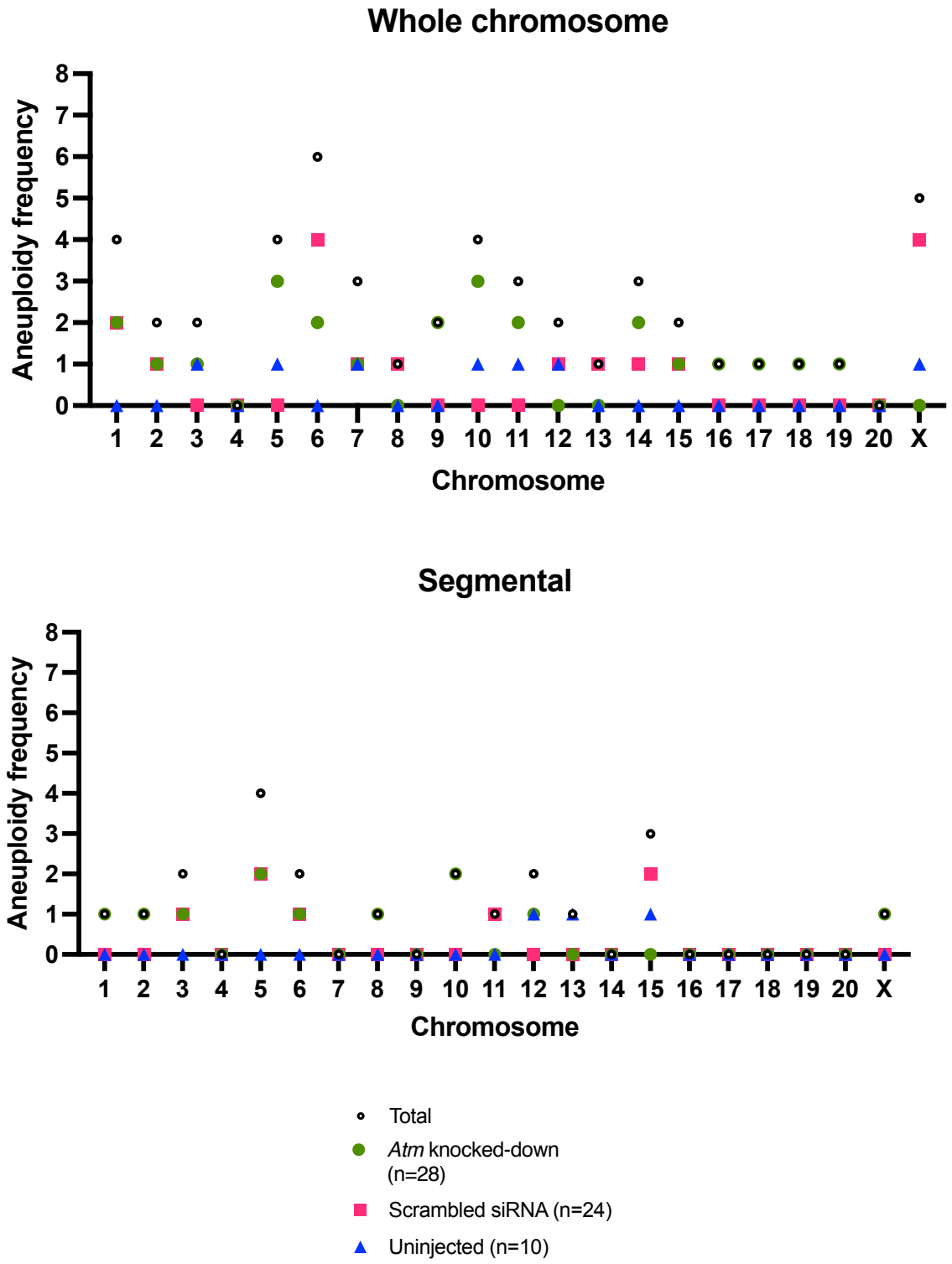


Figure 20. Chromosome-specific aneuploidy frequency in parthenotes.

Segmental aneuploidies were observed with the highest frequency overall in chromosome 5 (four times), followed by chromosome 15 (three times). The most

common segmental aneuploidy for the *Atm* knocked-down parthenotes was 5 and 10 (observed twice). Chromosomes 5 and 15 (observed twice) were the most common segmental aneuploidies for the scrambled siRNA-injected group.

Unfortunately, PBs were not paired with their corresponding 1PN parthenote. This study intended to find the pairs of PBs with their corresponding parthenote by comparing karyotypes to find any reciprocity. Nonetheless, no reciprocal karyotypes were found for two of them and one was karyotypically normal (haploid).

4.1.3.3. Power analysis

The present study was a pilot run and included small sample sizes, thus impacting the statistical significance of the findings. Therefore, to determine the ability of the present study to detect true biological differences, post hoc power analyses were calculated with the preliminary results obtained from the overall aneuploidy rate comparisons between the three study groups employing a 5% significance level.

The following formula was used for the comparison of the overall aneuploidy between the *Atm* knocked-down and uninjected groups:

$$Power = \emptyset\left[\frac{0.37}{\sqrt{\frac{0.57 \cdot 0.43}{28} + \frac{0.2 \cdot 0.8}{10}}} - 1.96\frac{\sqrt{0.47 \cdot 0.53(\frac{1}{28} + \frac{1}{10})}}{\sqrt{\frac{0.57 \cdot 0.43}{28} + \frac{0.2 \cdot 0.8}{10}}}\right]$$

$$Power = \emptyset\,[0.061] = \mathbf{52.4}\%$$

The following formula was used for the comparison of the overall aneuploidy between the scrambled siRNA-injected and uninjected groups:

$$Power = \emptyset\left[\frac{0.26}{\sqrt{\frac{0.46 \cdot 0.54}{24} + \frac{0.2 \cdot 0.8}{10}}} - 1.96\frac{\sqrt{0.38 \cdot 0.62(\frac{1}{24} + \frac{1}{10})}}{\sqrt{\frac{0.46 \cdot 0.54}{24} + \frac{0.2 \cdot 0.8}{10}}}\right]$$

$$Power = \emptyset\,[0.608] = \mathbf{27.2}\%$$

Finally, the following formula was used for the comparison of the overall aneuploidy between the *Atm* knocked-down and scrambled siRNA-injected groups:

$$Power = \emptyset\left[\frac{0.11}{\sqrt{\frac{0.57 \cdot 0.43}{28} + \frac{0.46 \cdot 0.54}{24}}} - 1.96\frac{\sqrt{0.52 \cdot 0.48(\frac{1}{28} + \frac{1}{24})}}{\sqrt{\frac{0.57 \cdot 0.43}{28} + \frac{0.46 \cdot 0.54}{24}}}\right]$$

$$Power = \emptyset\,[-1.175] = \mathbf{12}\%$$

4.1.4. Discussion

There is a strong correlation between the decline in women's fertility and advanced reproductive maternal age. Recent data found an association between oocyte ageing and a reduced expression in genes involved in the ATM-mediated DSBs repair pathway (Titus et al., 2013, Oktay et al., 2015), but the effect of a dysfunctional DNA repair ATM pathway on oocyte quality has not yet been studied. This pilot study aimed to analyse the effects of oocyte *Atm* downregulation in oocyte capacity for DNA damage repair and faithful meiotic chromosome segregation, to shed light on its role in oocyte competence.

The following discussion also includes the findings from the existing literature about the role of *Atm* in mice oocyte meiotic chromosome segregation and DSBs repair.

4.1.4.1. Aneuploidy frequency in the uninjected control group

4.1.4.1.1. Overall and whole chromosome aneuploidies

In the present mice parthenote study, 20% of samples from the uninjected control group presented a type of aneuploidy and 10% presented whole chromosome abnormalities. Aneuploidy is highly prevalent in human IVF blastocysts. Around half have a chromosomal abnormality, but the exact percentages depend on the woman's age (Franasiak et al., 2014, Rubio et al., 2019b).

Mice embryos have been investigated for their potential as a human aneuploidy model. Although mice oocyte and embryo aneuploidy has not been extensively studied and reported rates rely heavily on the method used for chromosomal number analysis, post-fertilisation mitotic chromosome abnormalities seem to occur to a lower degree in mice embryos (Lightfoot et al., 2006, Bolton et al., 2016). Data also shows that, compared to human oocytes, mice oocytes have lower rates of aneuploidy caused by meiotic errors (Duncan et al., 2009, Danylevska et al., 2014). Pauerova et al. (2020) showed that 4% of in vivo recovered mice zygotes are aneuploid, detected with kinetochore counting assay, while the aneuploidy rate observed from human young donor oocytes is around 20% (Hou et al., 2013, Gruhn et al., 2019). Nevertheless, mice oocytes, like humans, exhibit a sharp increase in aneuploidy rates as maternal age progresses (Duncan et al., 2009, Merriman et al., 2012, Shomper et al., 2014).

The GVs in the present study were collected from young mice. Thus, the resulting 1PN parthenote aneuploidy rate is inflated compared to other studies. One contributing factor for the observed discrepancy could be that previous publications have mainly employed *in situ* chromosome spreading assays, instead of CCS. However, a study by Treff et al. (2016) observed a decreased incidence of old mice IVM oocyte aneuploidy

based on NGS analysis compared to previously reported observations (15% vs. 60% at ages from 16 to 19 months) (Merriman et al., 2012).

Although the analysis performed with *in situ* chromosome spreading methodologies tends to be more subjective than quantitative methodologies, these traditional assays allow the assessment of chromatids' alignment and separation during meiosis. Chromosome mis-segregation during meiosis I and II can also be studied with NGS, but chromosomal number analysis of PBs must be included. However, as Treff et al. (2016) noted, NGS analysis performed on oocytes, unlike on zygotes, cannot identify chromosome mis-segregations of meiosis II, which would cause numerical chromosome abnormalities on the completion of meiosis and second PB extrusion after fertilisation. This study performed NGS-based aneuploidy analysis on parthenotes, thus, allowing for the detection of aneuploidy caused by chromosome segregation errors during meiosis I and II. Nevertheless, this study failed to pair the provided second PBs with their corresponding 1PN parthenote that would tell if aneuploidy originated in meiosis I and/or meiosis II for the three paired parthenotes and if it was caused by non-disjunction or by PSSC.

The parthenogenic activation of the mice oocytes could also play a role in the inflated aneuploidy rates. Initial reports showed a significant increase in mice embryo aneuploidy following ethanol activation (from 1% to 28%) (Kaufman, 1982). Evidence on the effect of $SrCl_2$ treatment on ploidy status is scarce, but early studies have shown that aneuploidies that originate at the completion of meiosis II do not significantly increase in parthenotes activated with this compound (O'neill et al., 1991). Bos-Mikich et al. (1995) reported an embryo aneuploidy incidence at the first mitotic division from $SrCl_2$-activated mice oocytes ranging from 5% to 8% and Huang et al. (2008) reported an incidence

ranging from 5.70% to 11.50% after second PB extrusion. The results from the aforementioned studies are similar to the ones presented in this thesis. Nonetheless, it is unknown whether the parthenogenic activation was a cause of the increased rate of aneuploidy observed in the 1PN uninjected controls. The sample size and lack of appropriate controls for the $SrCl_2$ treatment limit the accuracy of the conclusions that can be drawn.

CCS allows for the identification of chromosome-specific mis-segregations events. Interestingly, one of the two parthenotes with abnormal karyotypes in the uninjected control group were complex whole chromosome aneuploidies (defined as involving three or more chromosomes), as opposed to the *Atm* and scrambled siRNA-injected groups, where mis-segregations occurred, in most samples, in one or two chromosomes. This also suggests that DNA degradation, or analysis methodology artefacts, could have contributed to the inflated aneuploidy rates observed in the uninjected control group (Rubio et al., 2019b). Despite this, the observed aneuploidy in this group, considered as baseline, should be taken into account when studying the siRNA-injected and *Atm* knocked-down groups.

4.1.4.1.2. Segmental aneuploidies

Segmental aneuploidies are very infrequent in most cell types. This type of chromosomal abnormality was detected in 10% of the uninjected control parthenotes, according to the NGS-based analysis. Babariya et al. (2017) found segmental chromosomal abnormalities in 10.39% of 452 human oocytes analysed with array-comparative genomic hybridisation (aCGH). Even though the authors did not include

zygotes, they predicted that ~10% of oocytes would present segmental aneuploidies after the completion of meiosis II.

The molecular mechanisms that lead to whole chromosome and segmental aneuploidies are generally distinct. Segmental aneuploidies detected in human embryos are mostly of mitotic origin (Vera-Rodríguez et al., 2016, Girardi et al., 2020), and from the segmental abnormalities that are of meiotic origin, the majority affect paternal chromosomes (Konstantinidis et al., 2016, Kubicek et al., 2019). Unfortunately, data on segmental aneuploidy frequency in mice oocytes and embryos is lacking and it is difficult to draw conclusions from comparisons with human studies.

The mice oocytes in this study could have been predisposed to chromosome breakages that would give rise to segmental aneuploidy, due to extrinsic factors, such as, the chemical intervention during IVM and parthenogenic activation. The study used a 5 Mb resolution to detect segmental gains or losses. Compared to other studies that use above 10 Mb, this is a less conservative approach that can increase the risk of segmental aneuploidy overcalling (Girardi et al., 2020). Moreover, the 1-cell parthenotes were subjected to WGA prior to NGS which can sometimes introduce artefacts, due to preferential amplification (Voet et al., 2013, Sabina and Leamon, 2015). Another source of error is the single-cell S-phase of the cell cycle artefacts where replication domains sometimes give rise to copy number changes that are erroneously interpreted as segmental imbalances during downstream analysis (Dimitriadou et al., 2014).

More NGS studies are needed with statistically representative sample sizes to analyse the real frequency of segmental aneuploidy in mice oocytes and zygotes. They should employ appropriate controls, including in vivo matured MII oocytes and naturally

conceived 2PN embryos. Despite this, the level of segmental aneuploidies found in the uninjected control group should be considered a baseline for the other two study groups.

4.1.4.2. Effect of GV transfection with small-interfering RNA (siRNA) in one-pronuclear stage (1PN) parthenotes ploidy status

4.1.4.2.1. SiRNA sequence-dependent effects

To determine if the siRNA injection contributed to parthenote aneuploidy, GVs were injected with scrambled siRNA. Scrambled siRNA-injected parthenotes exhibited almost four times the amount of whole chromosome aneuploidies compared to the uninjected parthenotes. This points towards an off-target effect of GV transfection with siRNA in parthenote ploidy status, which the literature suggests is common. Even though Jackson and Linsley (2004) mention that the mechanisms behind off-target siRNA transfection effects are not yet fully understood, they can be caused by partial complementarity between targeted mRNA sequences and unintended mRNA, siRNA-specific off-target response, induction of non-specific interferon response, siRNAs functioning as miRNAs and cellular response to siRNA reagents toxicity and mechanisms.

Scrambled siRNA controls are designed by shuffling the effector siRNA sequence to eliminate the targeted gene down-regulation. Nevertheless, this type of negative control may exhibit its own sequence-dependent effects (Buehler et al., 2012). Jackson et al. (2003) revealed siRNA-specific expression profiles among multiple siRNAs that targeted the same mRNA transcript, but the levels of down-regulation of off-target transcripts were lower than the targeted gene and the off-target transcripts had partial complementarity to the targeted mRNA. This could suggest that the scrambled siRNA-

injected group presented a lower level of down-regulation of *Atm* compared to the *Atm* siRNA-injected group due to marginal sequence complementarity, reflected in the sample whole chromosome aneuploidy rate (37.50% and 42.85%, respectively). However, in the study by Jackson et al. (2003), the downregulated unintended transcripts shared a minimum of eight nucleotides to the 5' end of the effector siRNA sequence. Therefore, it is unlikely that any complementarity of the scrambled siRNA to the target mRNA was sufficient to cause any downregulation since the nucleotides (shared with the test siRNA) assigned to the scrambled siRNA sequence are randomised.

4.1.4.2.2. Microinjection effects

The preferred method and delivery for gene silencing in oocytes is siRNA microinjection. This is because of the ability to control the concentration of injected siRNA and because, before EGA, the oocyte and subsequent blastomeres rely on accumulated maternal transcripts prior to GV breakdown. However, the microinjection of transfection reagents exerts significant stress on the cell. Tan et al. (2020) studied the effects of siRNA microinjection by comparing the transcriptome of scrambled siRNA-injected to uninjected bovine oocytes. They identified the overexpression of genes with roles in the cytoskeleton and actin filaments, including the *ACTR2* gene, which protein is part of the ARP2/3 complex. According to mice studies, the ARP2/3 complex acts in the oocyte meiotic maturation process, including asymmetric division, spindle migration and cytokinesis (Sun et al., 2011). Tan et al. (2020) hypothesised that the disruption of the integrity of the plasma membrane could have caused the differences in gene expression between experimental groups. However, although the above effects are sequence-

independent, they also indicated that the scrambled siRNA could have interacted with the genome or oocyte mRNAs altering gene expression.

The effect of microinjection on oocyte integrity has been extensively studied after the introduction of ICSI in ARTs. A recent systematic review and meta-analysis in humans found no evidence of increased risk of overall aneuploidies in ICSI-conceived offspring, compared to conventional IVF and natural conception (Berntsen et al., 2021). Different from ICSI, where MII mature oocytes are injected with sperm, this study and Tan et al. (2020) injected GVs that subsequently matured to the MII stage after undergoing the first asymmetric division to extrude the first PB. GVs, different from MII stage oocytes, could be more susceptible to chromosome segregation errors from the stress the microinjection subjects the cell to. Such stressors can also cause chromosome breakages, in line with the 6.67% segmental aneuploidies increase in scrambled siRNA-injected parthenotes. This hypothesis is consistent with studies that reveal that most chromosome mis-segregation aneuploidies originate in the first meiotic division (Fragouli et al., 2013). Nevertheless, the correlation between microinjection of GVs, the overexpression of genes, such as *ACTR2* and aneuploidy, needs to be further investigated with statistically powerful sample sizes.

4.1.4.3. Effect of *Atm* knockdown in 1PN parthenotes ploidy status

This study hypothesised that the downregulation of *Atm* in GVs increases unrepaired DSBs and chromosome segregation errors, due to an association between ATM-mediated DSBs repair pathway deterioration and oocyte ageing. The percentage of overall aneuploid parthenotes in the *Atm* knocked-down group was 57.14% compared to 20% in the uninjected control. Whole chromosome aneuploidies were present in 42.85%

and segmental aneuploidies in 25% of the *Atm* knocked-down parthenotes. As mentioned, 10% of uninjected parthenotes presented whole chromosome aneuploidies and the same percentage showed segmental aneuploidies.

4.1.4.3.1. Wole chromosome aneuploidies

The differences in whole chromosome aneuploidy rate between the three experimental groups, the *Atm* knocked-down showing the highest, could point towards a role of ATM in meiotic chromosome segregation of the maturing oocyte. Indeed, in somatic cells, DDR proteins possess a mitotic role. ATM protein kinase is a crucial activator of several downstream molecules in response to DSBs (Paull, 2015), but around 700 proteins have been identified as substrates of ATM-mediated phosphorylation response to DNA damage (Matsuoka et al., 2007). In the absence of DNA damage, activated by Aurora-B kinase, ATM contributes to the SAC that leads to transient mitotic arrest, until spindle assembly completion, to guard against chromosome mis-segregation (Yang et al., 2011, Palazzo et al., 2014, Xiao et al., 2022).

However, the role of ATM in mitosis is complex and yet to be fully understood. Paradoxically, it has been shown that the DDR response in mitosis, to protect genome integrity from structural damage, increases the likelihood of chromosome mis-segregation. The DDR, through Aurora-A and Polo kinases (PLKs), stabilise kinetochore-microtubule attachments to chromosomes, increasing the probability of erroneous attachments that, if persistent until anaphase, increase the frequency of chromosome lagging. The inhibition of ATM eliminates this effect (Bakhoum et al., 2014). However, Bakhoum et al. (2017) argue that the type and location of the DNA damage could determine different response signalling.

Less is known about the mechanisms behind the surveillance of chromosome segregation errors in the oocyte. Although many protein players in mitosis also have a role in oocyte meiosis, their tasks can differ significantly. Chromosome segregation in meiosis I is also controlled by SAC (Stukenberg and Burke, 2015). However, unlike in mitosis, oocytes can proceed to anaphase I with a degree of bivalent misalignment and incorrect attachment (Lane et al., 2012). SAC also mediates faithful sister chromatid segregation in meiosis II (Cheng et al., 2017). It is currently unknown whether ATM plays a role in the meiotic SAC in regulating chromosome segregation, but similar to mitosis, Aurora kinases have been found to have a role in stabilising kinetochore-microtubule attachments and SAC-induced arrest (Liu and Lampson, 2009, Lane et al., 2010). As previously mentioned, in somatic cells, SMC1, a sub-unit of the cohesin complex, is phosphorylated by ATM kinase following induced DNA damage (Kim et al., 2002). Studies have found a meiosis-specific cohesin sub-unit, SMC1B. *Smc1b*$^{-/-}$ mice are sterile due to meiotic failure. Although research that directly studies the interaction between ATM and SMC1B in the oocyte is lacking, studies in somatic cells, where SMC1B activity was detected, found no physical association between SMC1B and ATM (Mannini et al., 2015), undermining the hypothesised interaction of ATM in the meiotic cohesin complex.

Something worthy of noting is chromosome-specific aneuploidy. No obvious pattern is observed in the *Atm* knocked-down samples. The study by Treff et al. (2016) with aged mice oocytes also failed to find a pattern. Importantly, the authors argued that chromosome mis-segregations associated with advanced age, and alterations in chromosome cohesion, are not reflected in gross global aneuploidy.

Although the results in this chapter could provide evidence of a yet unknown involvement of ATM in oocyte meiosis chromosome segregation mechanisms, the data in

this study must be cautiously interpreted. The scrambled siRNA-injected controls show a high degree of whole chromosome aneuploidy in the background. The small sample size and the difference in sample rate are reflected in the statistically non-significant comparisons. Nevertheless, these results raise questions that should be further studied with statistically powerful sample sizes.

4.1.4.3.2. Segmental aneuploidies

The difference in sample rate with segmental aneuploidies between the *Atm* knocked-down group and the uninjected group (25% vs. 10%, respectively) could also indicate a role of ATM in DSBs repair in oocytes from the dictyate arrest onwards. DNA repair and chromosome segregation are two inter-connected mechanisms where many of the proteins involved in DDR are also needed for faithful chromosome segregation (Bakhoum et al., 2017). When mitotic cells meet DNA damage, modified histone H2AX (G-H2AX), a mediator of DNA damage checkpoint 1 (MDC1) and members of the MRN complex are recruited. ATM is then activated along with ATR and DNA protein kinase (DNA-PK) and, subsequently, CHK1 and CHK2 kinases. This, in turn, recruits PLKs, which enable the accumulation of BUBR1/BUB3 complex at the site of DSBs for marking for later repair. BUB3 is also a SAC protein that ensures faithful chromosome segregation (Derive et al., 2015, Xiao et al., 2022). As already discussed, DNA damage can also delay the entry to anaphase via SAC proteins, demonstrating crosstalk between the DNA damage and chromosome segregation checkpoints (Kim and Burke, 2008, Bakhoum et al., 2014).

In response to DNA damage, oocytes arrest in meiosis I through the inhibition of the APC with a SAC-dependent mechanism (Lane et al., 2017b). Although there is a role of ATM in both chromosome segregation and DNA repair in mitotic cells, the inhibition of

the APC from DSBs response in oocytes depends on Aurora, MPS1 and Haspin kinases, but not on ATM kinase. However, unlike meiosis I, meiosis II does not respond to DNA damage with arrest and can complete meiosis to either repair DNA or enter apoptosis in G1, resembling a mitotic cell (Lane et al., 2017b, Hustedt and Durocher, 2017). Even though it could be speculated that ATM, therefore, plays a role in DSBs site marking in meiosis II for later repair, it is unclear whether this is the case.

Evidence shows that ATM plays a pivotal role in oocyte meiosis HR at early prophase I. Meiotic recombination is initiated with programmed DSBs executed by the SPO11 transesterase and TOPVIBL at the leptotene sub-stage (Lam and Keeney, 2015), and failure of the later repair of these breaks can compromise genome integrity and generate chromosomal rearrangements after chromosome segregation (Kim et al., 2016). In the majority of species, programmed meiotic DSBs are localised at specific regions in the genome known as 'hotspots' (Baudat et al., 2013). In most mammals and yeast, programmed DSBs, in turn, activate ATM that, with a negative feedback loop, restricts SPO11 from further catalysis of DSBs (Garcia et al., 2015). Di Giacomo et al. (2005) reported that *Atm*$^{-/-}$ female mice are infertile due to a complete loss of eggs, likely caused by defective DSBs repair at early prophase I.

In this study, GVs were transfected with *Atm* siRNA. It is believed that prophase I oocytes arrest at the diplotene sub-stage since this is the most stable chromosome conformation (Wang and Pepling, 2021). Therefore, any potential effect of *Atm* downregulation may not have impacted the programmed DSBs process since the diplotene is a prophase I sub-stage that comes after DSBs resolution. Moreover, there is limited evidence that supports a regulatory role of ATM in programmed meiotic DSBs in mice oocytes (Wang and Pepling, 2021).

ATM is highly expressed in mammalian oocytes (Menezo et al., 2007), and recent studies have shown a critical role of ATM in the HR-mediated pathway of DBSs repair in mice maturing oocytes (He et al., 2018). Nevertheless, evidence shows that only high levels of damage appear to activate ATM and the ATM-mediated checkpoint at prophase I in GVs for meiosis resumption, prompting questions about the role of ATM in DSBs repair from the GV-stage onwards (Marangos and Carroll, 2012, He et al., 2018).

Given the aforementioned limitations, the findings in this section can be used to speculate that the difference in segmental aneuploidy sample rate observed between the *Atm* knocked-down and uninjected groups could be due to *Atm* knocked-down oocytes that accumulated a higher degree of DSBs that could not be detected and repaired through ATM-dependent pathways.

4.1.4.4. Study limitations

The previous sections already raised many inherent limitations to this pilot study. These mostly likely relate to the small sample size and its impact on statistical significance. Calculations determined that the statistical power of the present study design to detect a biological true difference between overall aneuploidy rates ranged from 12 to 52.4%. Due to the inadequate power, future studies should include larger sample sizes ideally after performing a sample size calculation analysis.

Two other limitations exist. Firstly, extensive chromosome segregation studies in human and mouse oocytes have included PBs that allowed for the distinction of meiosis I and meiosis II mis-segregations. As briefly discussed, this study only included three PBs and was unable to find reciprocal aneuploidies with their parthenotes. Thus, it was

impossible to determine if the observed aneuploidies in the 1PN parthenotes were the product of meiosis I or meiosis II and the type of segregation error. This would provide crucial information about the role of ATM in mice oogenesis and could add to research in the field.

Secondly, the NGS-based aneuploidy analysis was unable to determine the locations of chromosomal breakpoints of segmental abnormalities. Such regions in the genome can be compared with the known species' DSBs recombination hotspots to determine if the chromosome breakages are random or inclined to occur within these known regions of the genome. Thousands of hotspots are known to exist in the mouse genome (Brunschwig et al., 2012). This information could shed more light on the function of ATM in DSBs repair in the oocyte.

4.1.4.5. Conclusion

There is a strong correlation between the decline in women's fertility and advanced reproductive maternal age. Recent data found an association between oocyte ageing and an impaired ATM-mediated DSBs repair system, but the effect of a dysfunctional DNA repair ATM pathway on oocyte quality has not yet been studied. This chapter aimed to analyse, for the first time, the direct impact of *Atm* downregulation on the oocyte's capacity to repair DNA and carry out faithful chromosome segregation during meiosis, by injecting GVs with *Atm* siRNA. Oocytes were subjected to parthenogenesis and 1PN parthenotes were analysed for aneuploidies with NGS.

This chapter identified a baseline aneuploidy level for whole chromosome and segmental aneuploidies in the control group of parthenotes produced from uninjected

and scrambled siRNA-injected GVs. Although the small sample size could be attributed to the observed potential bias, analysis methodology artefacts and accumulation of damage caused by in vitro manipulation of gametes can inflate the observed segmental aneuploidies. This study also identified a potential effect of the transfection through microinjection of siRNA on meiotic chromosome segregation. This is consistent with reports of the overexpression of genes with roles in spindle formation and cytokinesis. GVs, unlike MII oocytes, can be more vulnerable to mechanical and chemical stressors.

The percentage of parthenotes with whole chromosome and segmental aneuploidies in the *Atm* knocked-down group was higher compared to the uninjected and scrambled siRNA-injected controls. The exact mechanism and players in oocyte chromosome segregation and DBSs repair are yet to be defined. Nevertheless, previous research points towards little or no involvement of ATM in oocyte chromosome segregation surveillance. Moreover, evidence suggests that only high levels of DSBs appear to activate ATM in GVs. Thus, the observed elevated levels of segmental aneuploidies could result from increased accumulated damage that was not repaired with ATM-dependent pathways, due to *Atm* downregulation. The increased levels of aneuploidy in the *Atm* knocked-down parthenotes could also suggest an unknown role for ATM in faithful chromosome segregation. Nevertheless, these results should be interpreted carefully due to the study's low statistical power.

The findings in this chapter, therefore, prompt further studies with bigger sample sizes to confirm a true biological difference. Furthermore, parthenotes must be paired with their PBs to study the impact of ATM in chromosome segregation during meiosis I and II. To date, no association can be established between an ATM-pathway dysfunction and bad oocyte quality reflected in impaired chromosome segregation and DSBs repair.

Experimental section 2. Non-invasive methods for embryo genetic testing

Chapter 5. Evaluation of a truly non-invasive preimplantation genetic testing (niPGT) for aneuploidies protocol using spent culture media (SCM)

5.1.1. Introduction

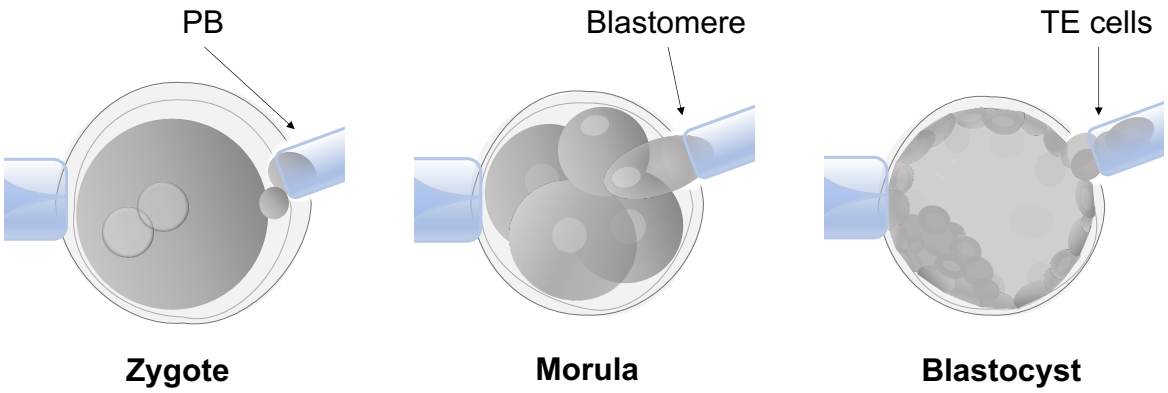
Currently, routine genetic assessment of preimplantation embryos mainly involves PGT-A for aneuploidy testing and PGT-M for testing of monogenic diseases (Figure 21(A)). Although chromosome number alterations significantly limit the reproductive competency of embryos, aneuploidy is a common occurrence that increases along with the female patients' age: 40% of embryos are aneuploid in 23-year-old women, this number doubles in 43-year-old patients and at 47, 100% of embryos are aneuploid (Franasiak et al., 2014). Data from a multicentre, blinded prospective and non-selection study on the clinical outcomes associated with PGT-A, employing a highly validated NGS-based methodology, reported that 82.10% of embryos diagnosed as euploid implanted and 64.70% achieved healthy deliveries in contrast to aneuploid embryos which 0% sustained implantation. The study obtained a 'no-call' rate in 2.80% of

the total TE biopsies. Of the total participants, 11.40% were 40 years or older (Tiegs et al., 2021).

Historically, routine PGT methods rely on the removal of blastomeres, TE cells (~6) or PBs with micromanipulation (Verlinsky et al., 1990, Munné et al., 1993) (Figure 21(A)). The biopsy is invasive and requires specialised equipment and highly skilled personnel to obtain good quality samples that provide enough high-quality DNA whilst minimising any potential damage to the embryo. PGT also considerably adds to the cost of ARTs treatment. Given the drawbacks, there is significant interest in the potential of non-invasive embryo selection and genetic testing methods.

(A) Invasive approaches				
	Analysis	**Sampling** [(1)]	**Methodology**	**Used routinely**
Preimplantation genetic testing (PGT)	Aneuploidies (PGT-A)	PBs, blastomeres or TE cells (~6)	aCGH, NGS	Yes
	Monogenic diseases (PGT-M)	PBs, blastomeres or TE cells (~6)	Sanger sequencing, PCR, RT-qPCR, and NGS	Yes

(1)



(B) Non-invasive approaches				
	Analysis	**Sampling [2]**	**Methodology**	**Used routinely**
Non-genetic	Morphology	Images	Single-time observations	Yes
	Morpho-kinetics	Images and video	Time-lapse photography	Sometimes
	Proteomics	SCM	Mass spectrometry	No
	Metabolomics	SCM	Mass spectrometry, NIR and NMR spectroscopy	No
Genetic	Small non-coding RNA	SCM	RT-qPCR	No
Non-invasive preimplantation genetic testing (niPGT)	Aneuploidies	SCM or BF	aCGH, NGS	No
	Monogenic diseases	SCM or BF	Sanger sequencing, PCR, RT-qPCR and NGS	No

(2)

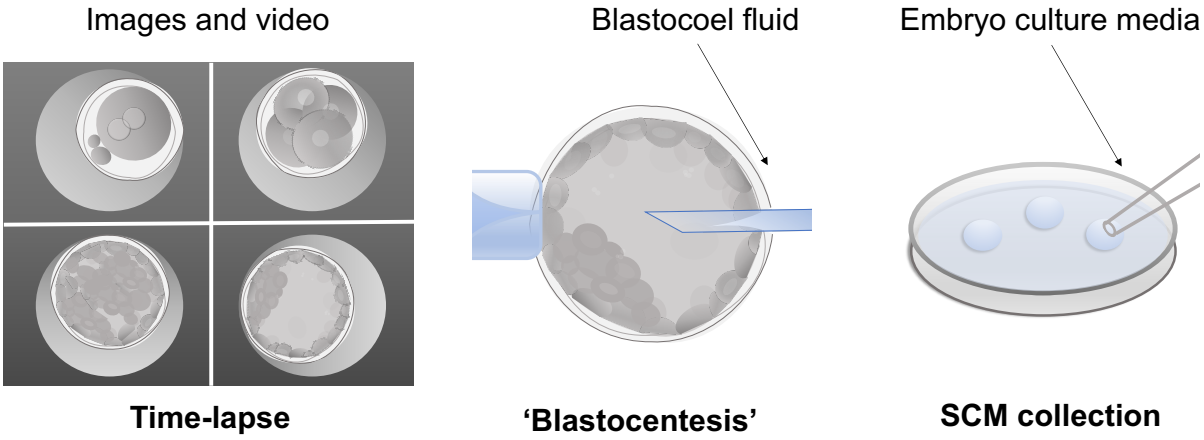


Figure 21. Preimplantation embryo quality assessment. (A) Traditional approaches used routinely include PGT-A and PGT-M where PBs, blastomeres or TE cells are obtained with micromanipulation. **(B)** Routine non-invasive approaches include morphology and morpho-kinetics assessment. The study of the proteome, metabolome and small non-coding RNA in embryo SCM has also been proposed. Nevertheless, more novel approaches include the analysis of aneuploidies and monogenic diseases in cell-free (cfDNA) from BF and SCM (Rødgaard et al., 2015, Carvalho et al., 2020, Leaver and Wells, 2020). aCGH, array comparative genomic hybridisation; BF, blastocoel fluid; NGS, next generation sequencing; NIR, Near-infrared; NMR, nuclear magnetic resonance; RT-qPCR, quantitative real-time PCR; SCM, spent culture media; TE, trophectoderm.

5.1.1.1. Non-invasive preimplantation embryo quality assessment

Various approaches were initially proposed for the non-invasive assessment of the preimplantation embryo competency. In their review, Rødgaard et al. (2015) show that researchers extensively explored morphology and morpho-kinetic evaluations, but found no evidence that their application increases live birth rates. However, they identified the analysis of embryo spent culture media (SCM) as a promising avenue to investigate the potential for analysing proteins, metabolites and small non-coding RNA as non-invasive biomarkers of embryo viability, but this approach is less practical in clinical settings and the sensitivity of the platforms used are low. More recently, the discovery of cell-free DNA (cfDNA) in the BF and in embryo SCM has prompted novel proposals for non-invasive PGT (niPGT) (Figure 21(B)).

5.1.1.1.1. Blastocoel fluid (BF) and aneuploidy testing

Mammalian embryos undergo a process of compaction that leads to the differentiation of the TE and inner cell mass (ICM) lineages. In humans, this process begins at the 8-cell stage. The preimplantation development then follows the blastocyst formation. The blastocyst encloses a fluid cavity delimited by TE monolayer epithelium, which the BF accumulates due to the presence of osmotic gradients that allow the passage of water coordinated by Na^+ K^+ pumps across the TE (Watson and Kidder, 1988, Manejwala et al., 1989).

Reports of the detection of metabolites of interest in the BF through mass spectrometry (D'Alessandro et al., 2012) prompted the study of the existence of cfDNA extruded from embryonic cells into the blastocoel cavity and its value as an indicator of

embryo viability. This novel approach, known as 'blastocentesis', attracted much enthusiasm since tight junctions that separate the blastocoel cavity from the environment would prevent the passing of extra-embryonic contaminants that would hinder its value. Palini et al. (2013) first successfully amplified the autosomal gene *TBC1D3* and the multi-copy gene *TSPY1* with RT-qPCR in 89.70% of BF samples. They estimated that the quantity of DNA found in these samples was <10 pg. However, in their review Leaver and Wells (2020) note that amplification rates vary between studies, from 34.80% to 87.50%, when attempting WGA. BF was subsequently studied for its potential as a source of embryonic gDNA for aneuploidy testing using aCGH and NGS. Nevertheless, the reported overall ploidy concordance between BF and the corresponding TE biopsy or whole embryo varies between studies, from 63% to 93.60% (Tobler et al., 2015, Capalbo et al., 2018, Tšuiko et al., 2018, Magli et al., 2019).

Consequently, there is less enthusiasm for exploring the potential of BF for preimplantation genetic testing. Differences in the platforms used for chromosomal number analysis, variable volumes of fluid recovered, embryo, sample and DNA quality and quantity may play a role (Leaver and Wells, 2020). Moreover, in the context of niPGT, blastocentesis is still considered somewhat invasive because there is a risk of damaging the cells with the micromanipulation, causing inadvertent lysis. The release of nuclear material into the blastocoel cavity may also be causing the enhanced amplification and concordance rates in some studies.

5.1.1.1.2. SCM and aneuploidy testing

SCM provides a promising and strictly non-invasive alternative to traditional PGT, where the embryo does not have to undergo any degree of manipulation to obtain

genetic material of embryonic origin. Due to the simplicity of SCM sampling, where only a small volume of embryo culture media is recovered, this proposal has generated great interest prompting the emergence of numerous studies.

The first reports of the presence of amplifiable gDNA in SCM came from Stigliani et al. (2013) and Stigliani et al. (2014). Studies followed that performed chromosomal number analysis using aCGH and, mostly, NGS on SCM from day-1 to day-6 of in vitro culture. The accuracy of this approach has generally been assessed with karyotype and overall ploidy concordance between paired SCM samples (niPGT) and their complement TE biopsy (PGT-A) or whole embryo. The successful WGA rates range from 89.93% to 100%, full karyotype concordance rates from 16% to 83.30% and overall ploidy concordance from 30.40% to 96.40% (Xu et al., 2016, Liu et al., 2017, Vera-Rodriguez et al., 2018, Kuznyetsov et al., 2018, Huang et al., 2019, Rubio et al., 2020, Yin et al., 2021).

The release mechanism of cfDNA into the BF and SCM is not yet fully understood, but its origin is a key factor in the accuracy of niPGT to determine the chromosomal status of the embryo. It is believed that a portion of the extra-embryonic cfDNA found in BF and blastocyst SCM originate from apoptotic embryonic cells which could explain the fragmented nature of the DNA found in these samples (Nagata, 2000, Hardy et al., 2003).

Given the concordance results, some studies have proved that the embryonic cfDNA found in SCM may be of better quality and in greater amounts compared to BF. Nonetheless, there are discrepancies in the SCM data. This is likely due to SCM collection at different stages of preimplantation development, cumulus cell carry-over leading to maternal contamination and difference in the degree of embryo manipulation (Vera-Rodriguez et al., 2018).

5.1.1.2. Study aims

This chapter aims to evaluate the accuracy of a truly non-invasive PGT protocol where the manipulation of the embryo is minimised prior to the collection of SCM. The embryological protocols followed in this study were based on the hypothesis that embryo manipulations such as vitrification, thawing and biopsy can cause cell lysis, facilitating the release of DNA into the media. Stringent embryological protocols were also applied to avoid contamination.

The specific aims are:

1) To define threshold values for the libraries' concentration after WGA and NGS quality score for the inclusion of informative SCM samples for aneuploidy analysis.
2) To investigate the effect of clinic of origin, ploidy status and day of SCM collection in sample informativity.
3) To analyse the overall ploidy and full karyotype concordance rates between SCM samples and their paired TE biopsy.
4) To assess the DNA contamination from maternal origin in SCM samples by comparing the presence of the X and Y chromosomes in SCM samples and their paired TE biopsies according to the niPGT and PGT-A results, respectively.
5) To evaluate the potential clinical application of the commercial kit employed for the niPGT.

5.1.2. Materials and methods

5.1.2.1. Ethics statement

The experiments in this study were reviewed and approved by the Western Institutional Review Board (Biomedical research protocol: 20031397).

5.1.2.2. Experimental design and sample preparation

This study used 128 SCM samples from 128 day-5 and day-6 post-fertilisation embryos obtained from couples undergoing IVF treatment and PGT-A. Three clinics participated. This study excluded SCM of embryos from frozen-thawed cycles. Unless otherwise specified, embryo culture and manipulation techniques were performed in line with the clinics' standardised protocols.

MII oocytes were denudated and fertilised with ICSI. Embryos were then carefully washed and separately placed in a fresh 30 µl drop of media on day-4 of in vitro culture to remove any possible remaining cumulus cells. Approximately 7 µl of media were recovered on day-5 or day-6 and transferred to an UV-irradiated 0.2 ml tube before performing the TE biopsy. One pipette per sample was used to collect the SCM to avoid cross-contamination. In parallel with the SCM sample, the TE biopsy from each blastocyst was sent to Juno Genetics, UK, to be analysed for routine PGT-A as part of the clinical service offered by the clinics (Figure 22). The SCM samples were shipped and immediately placed at -80 °C upon reception and stored until ready for WGA.

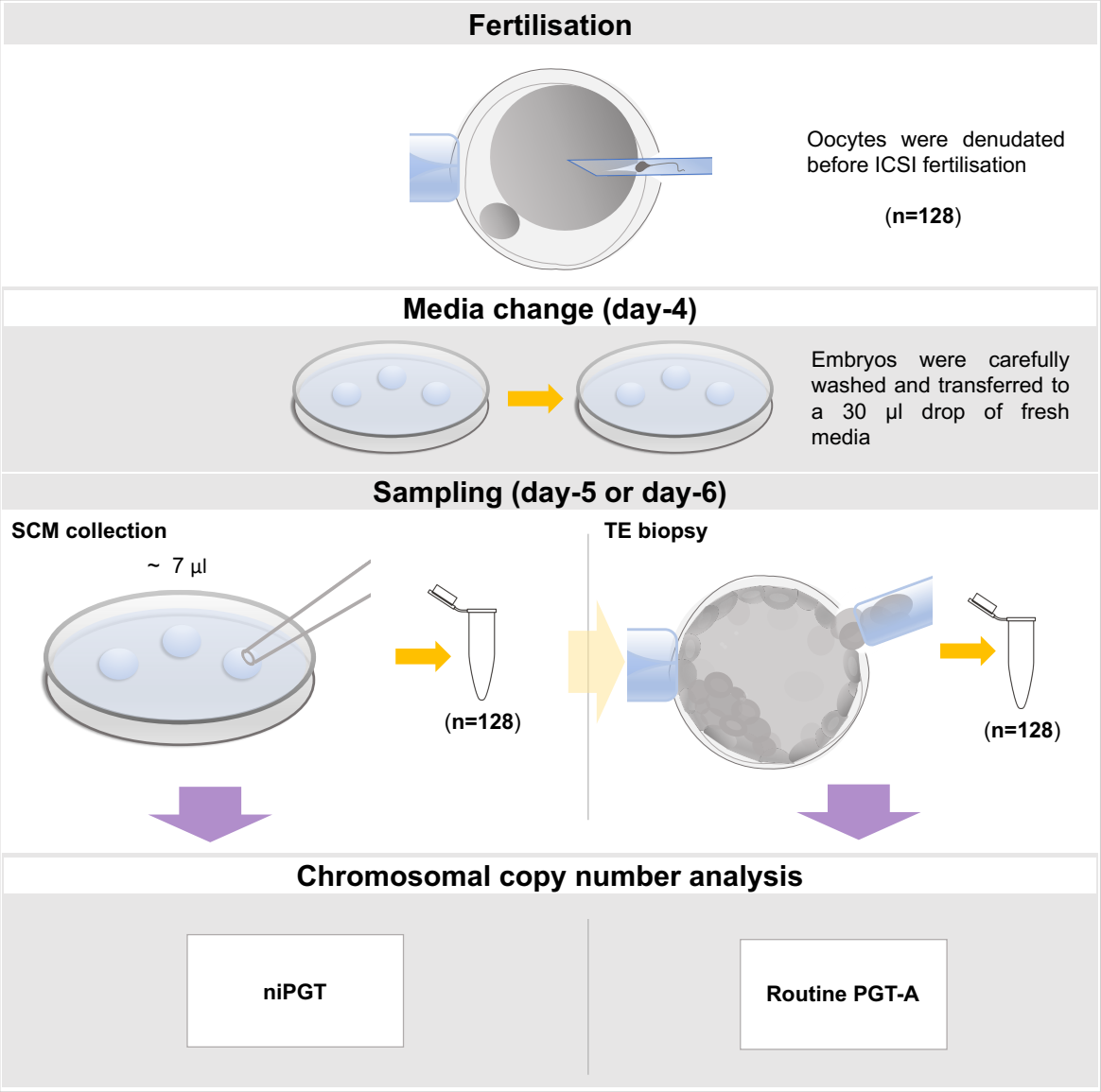


Figure 22. SCM and TE sample collection. This study excluded embryos fertilised with conventional IVF, vitrified embryos or SCM from embryos with previous biopsy. The protocol followed for the niPGT on the corresponding SCM samples was PG-Seq Rapid Non-Invasive PGT (Perkin Elmer, USA). The routine PGT-A on the TE biopsies was performed by Juno Genetics, UK with the PGTseq-A method.

5.1.2.3. Aneuploidy testing

The analysis of the ploidy status of the embryos using SCM was carried out following the protocol PG-Seq Rapid Non-Invasive PGT for WGA, library preparation and indexing as per **Chapter 2.** The libraries were normalised and denatured using the MiSeq Sequencing Kit V3 and the sequencing and data acquisition was carried out in the MiSeq

Sequencing System. The positive controls were human female DNA with a normal karyotype (46, XX) (Promega, USA), human female DNA with trisomy in chromosome 18 (47, XX, +18) (Coriell Institute, USA) and human male DNA with a segmental loss in chromosome 10 (46, XY, -10p14p12.1) (Coriell Institute, USA). A no-template control per library preparation run was also included. WGA products were run in a 1% agarose gel electrophoresis (1 g of agarose) for 15 minutes at 200 volts. The electrophoresis gel was then visualised to confirm amplification. Successful amplification was also determined by measuring the libraries' concentration with Genomic DNA TapeStation Reagents and Screen Tapes in the TapeStation System following **Chapter 2.** The niPGT data was analysed with the PG-Find Software V3. A portion of the SCM samples were processed and analysed by Miss Luisa Parnell.

The TE biopsies were processed by Juno Genetics, UK, for chromosomal copy number analysis, according to the PGTSeq-A method using the NextSeq (Illumina, USA) instrument for sequencing. The sequencing data was analysed by a custom-designed software by the developers of the PGTSeq-A method.

For the niPGT analysis, embryos were classified as euploid when the corresponding reads to all autosomal chromosomes aligned to number 2 on the copy number axis (Y-axis) of the chromosomal profile. Euploid females also displayed X chromosome reads aligned to copy number 2 and to 0 for Y chromosome. Euploid males displayed reads aligned to copy number 1 for both the X and Y chromosomes. Embryos were classified as aneuploid when reads aligned outside their default copy number. Losses were identified when reads aligned below 1.3 copy number axis and gains above 2.7. The niPGT detection limit for segmental aneuploidies was ≥5 Mb.

Due to the low-quality and low-abundance of the gDNA in SCM, the chromosome profiles tend to be noisier than the profiles from TE biopsies. The mosaicism criteria were not applied to the niPGT analysis and any mosaicism detected by PGT-A was disregarded for concordance analysis for this reason. The PGT-Seq-A detection limit for segmental aneuploidies was 18 Mb. However, this limit varies between chromosomes. To assess concordance between niPGT and PGT-A results, due to the differences in DNA quality between TE and SCM and the difference in copy number variation (CNV) size detection limit between analyses, the location of chromosomal breakpoints of segmental aneuploidy did not have to match precisely.

5.1.2.4. Concordance between paired SCM and trophectoderm (TE) biopsy samples

Concordance between SCM samples and their complement TE biopsy aneuploidy analysis was classified as follows:

a) 'Overall ploidy concordance' was defined when both SCM and TE biopsy were euploid or aneuploid.
b) 'Full karyotype concordance' was defined when both SCM and TE biopsy showed the same karyotype (disregarding mosaicism detected by PGTSeq-A or any apparent by niPGT and location of chromosomal breakpoints did not have to match precisely).
c) 'Aneuploidy reciprocals' were defined when a gain (or a loss) of a whole chromosome or segment was detected in a TE biopsy sample and a loss (or a gain) of the same chromosome or segment was observed in the complement SCM sample (location of chromosomal breakpoints did not have to match precisely).

d) ‘Male SCM maternal contamination’ was defined when TE biopsies were confirmed male, but SCM analysis showed a female result.

5.1.2.5. Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV)

Sensitivity, specificity, positive predictive values (PPV) and negative predictive values (NPV) were calculated using the following equations:

$$Sensitivity\ (\%) = \frac{True\ positives}{True\ positives + False\ negatives}$$

$$Specificity\ (\%) = \frac{True\ negatives}{True\ negatives + False\ positives}$$

$$PPV\ (\%) = \frac{True\ positives}{True\ positives + False\ positives}$$

$$NPV(\%) = \frac{True\ negatives}{True\ negatives + False\ negatives}$$

5.1.2.6. Statistical analysis

A three-way ANOVA determined if the clinic of origin, the day of SCM collection, or the ploidy status of the embryo had an impact on the libraries’ concentration and NGS quality score. The ANOVA was carried out with a 95% confidence interval. P-values of less than 0.05 were considered statistically significant. The analysis was performed using the SPSS Software (IBM, USA). A two-tailed with a 95% confidence interval Pearson r test was

performed to analyse the statistical correlation between libraries' concentrations and NGS quality scores. The differences in the observed female DNA contamination rates between clinics were tested using the unpaired *t*-test with a 95% confidence interval. GraphPad Prism was the software used for the Pearson r test, the *t*-test, the lognormal distribution analysis of the datasets and graph drawing.

5.1.3. Results

5.1.3.1 Libraries' concentration and NGS quality score

The 128 SCM samples were obtained from three different clinics (75 from Clinic 1, 46 from Clinic 2 and seven from Clinic 3). After WGA, and before nucleic acid purification, libraries from 86 samples were run in an electrophoresis gel (Figure 23) and in the TapeStation. In the TapeStation Software, a region was selected between 200 bp and 800 bp to measure concentrations, since, according to the manufacturer, the libraries' fragment size ranges between these sizes (Figure 24). The resulting concentrations ranged from 0.44 ng/µl to 36.20 ng/µl. After juxtaposing band intensity and concentration results from the 86 samples only libraries that showed a concentration above 2 ng/µl were included in the sequencing run (70.93% (61/86)). All samples for which this information was unknown were also sequenced. Therefore, 77% of the total of SCM samples were sequenced (99/128). Information of the 99 SCM samples and controls included for NGS can be found in **Appendix C**.

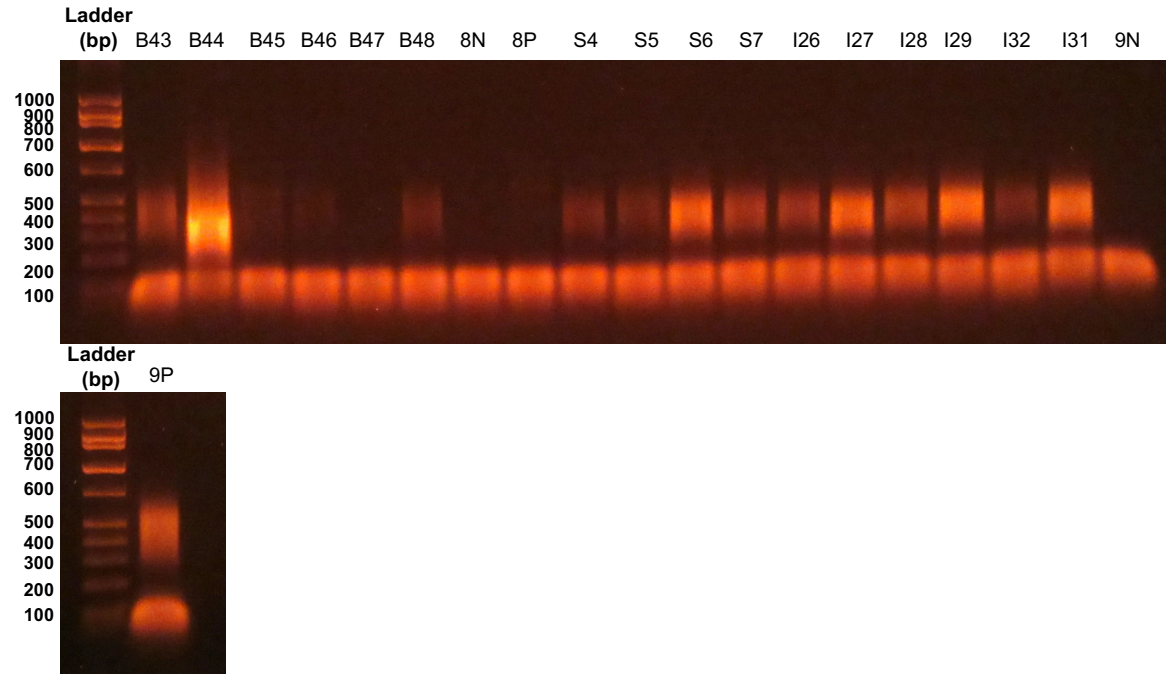


Figure 23. Agarose gel electrophoresis of WGA products from 20 representative samples following the PG-Seq Rapid Non-Invasive PGT protocol for library preparation and indexing. 8N and 9N are no-template controls, while 8P and 9P are positive controls. Different 200 – 800 bp band intensities are observed. Products with a size <200 bp are primer-dimers. Samples were run in a 1% agarose gel electrophoresis for 15 minutes at 200 volts.

However important the concentration of libraries, an optimal downstream copy number analysis will depend on the quality of the data obtained from the NGS. The PG-Find Software calculates a quality score based on the number of 'usable' reads per sample and bin. The greater the quality score value, the lower the quality of the libraries. This information was known for 61 of the 99 SCM samples sequenced. The libraries' concentrations and NGS quality scores for the 61 samples and one control are found in **Appendix C**.

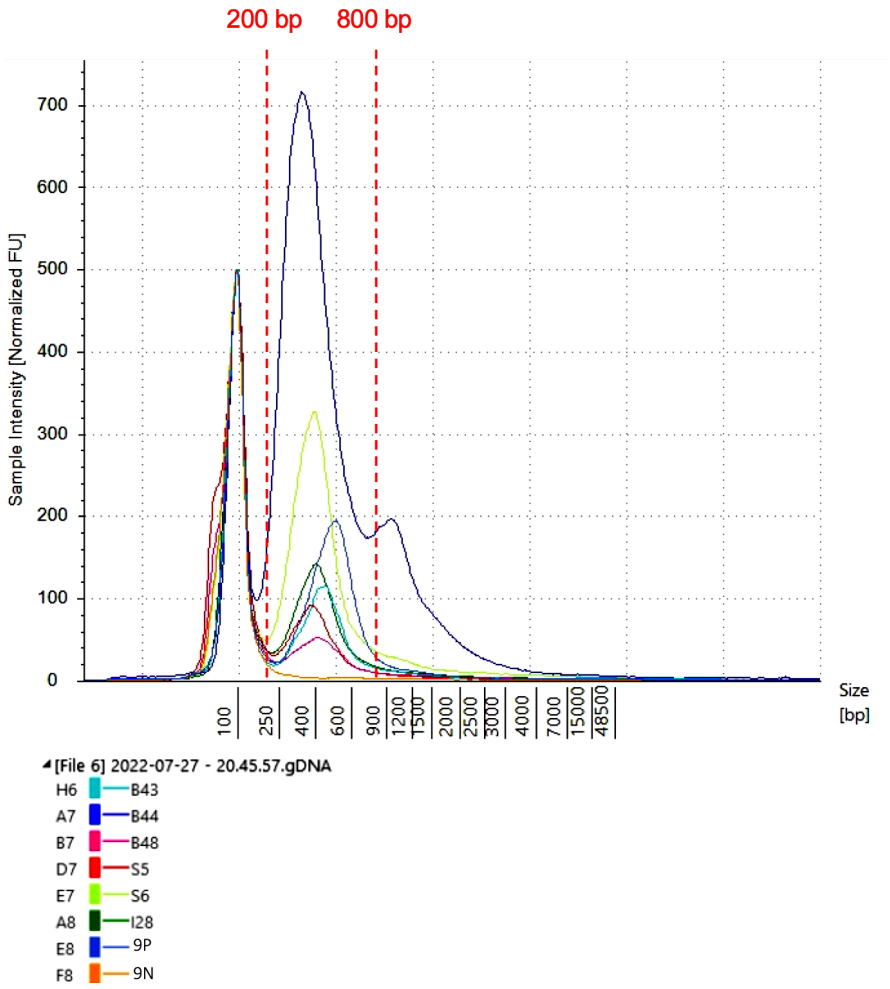


Figure 24. Genomic DNA TapeStation sample intensity comparison of eight representative samples after WGA. A region was selected between 200 bp and 800 bp size products to calculate library concentrations per sample. Products with a size <200 bp are primer-dimers.

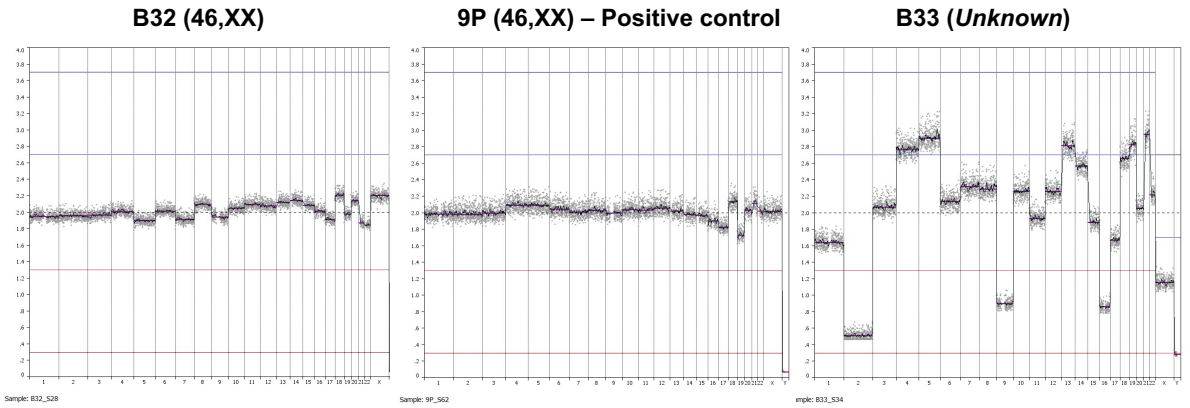
The NGS quality scores were first compared between samples using the karyotypically normal female commercial gDNA positive control (9P) as a reference to determine the maximum acceptable value (Figure 25(A)). The NGS quality score produced by 9P was 0.43. Some of the SCM samples achieved lower scores. A SCM sample where the karyotype matched the results of PGT-A, except for a detected segmental gain in chromosome 8, was selected from the male samples. The gender was confirmed by PGT-A and it was not contaminated with female DNA (the female contamination assessment of this study is discussed in the following section) (Figure 25 (B)). The value of this

sample's NGS quality score was 0.68, and therefore, the maximum acceptable value selected was 0.69, this number being less than double the quality score produced by 9P (0.43).

Fifty-two of the 61 sequenced samples (85.24%) from which NGS quality score values were known produced a value lower than 0.68 and were used for further analysis. For the remaining 38 sequenced samples, a 'pass' or 'fail' value was assigned depending on the observed noise in the chromosome plots. From these 38 samples, 30 (78.95%) passed for further analysis. Overall, 82.83% of all the sequenced samples (82/99) were included for analysis.

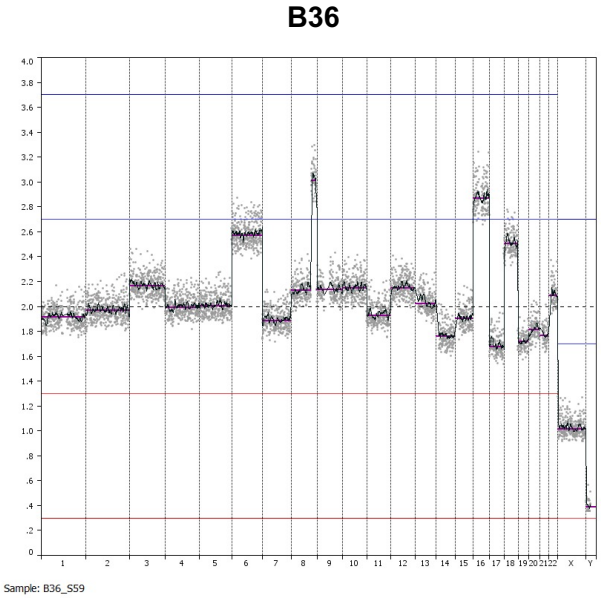
A three-way ANOVA was performed to analyse whether the clinic, the SCM collection day and the ploidy status of the embryo had a significant effect on the libraries' concentration and NGS quality score (Table 31.). Only data from Clinic 1 and Clinic 2 were included, since the sample size from Clinic 3 was not statistically representative. A lognormality distribution was first confirmed for both the libraries concentration and NGS quality scores and the data was subsequently log-transformed for analysis. A second analysis was carried out after removing male samples where female contamination was confirmed after comparing PGT-A and niPGT results.

(A)



NGS quality scores		
0.17	0.43	0.72

(B)



PGT-A karyotype result:

47, XY, +16

niPGT karyotype result:

47, XY, 8(seg), +16

NGS quality score
0.68

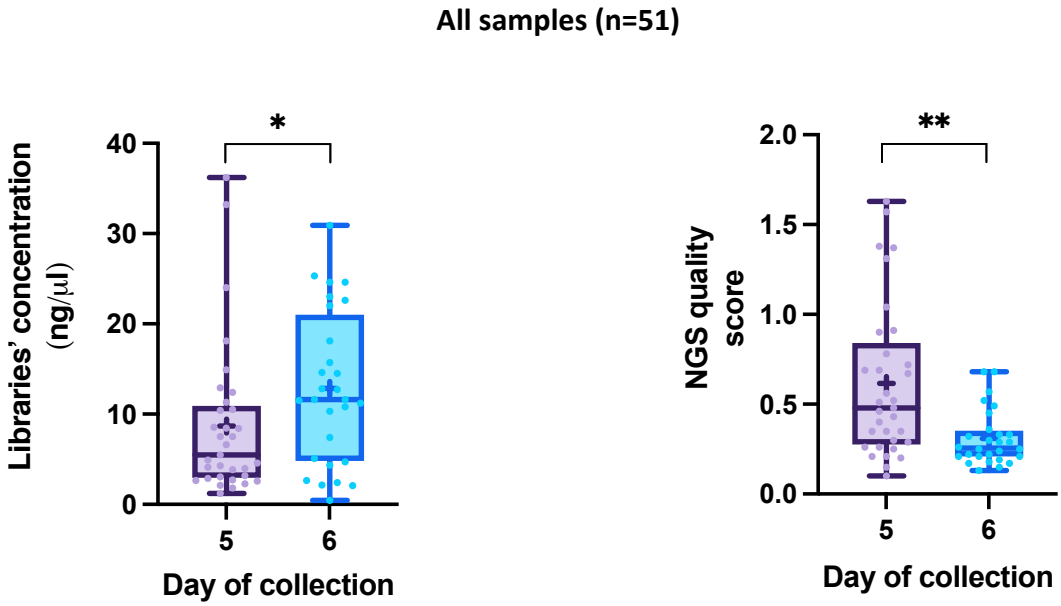
Figure 25. Representative chromosome plots produced by the PG-Find software. (A) The mapped reads on copy number 2 (Y-axis) are mostly gathered for B32, whereas the reads for B33 are very scattered, producing a noisy profile. Sample B33 was discarded for further analysis**. (B)** Male sample which niPGT karyotype matched almost exactly the PGT-A karyotype except for a segmental gain in chromosome 8. The threshold for copy gain is indicated with a blue line (Y=2.7) on the Y-axis and copy loss with a red line (Y=1.3). Any apparent mosaicism that occurred when chromosome number alignments fell between 2 and 2.7 or 2 and 1.3 was disregarded for concordance analysis. NGS, next-generation sequencing.

Table 31. P-values from the three-way ANOVA to determine statistical significance of the effect of the clinic of origin, collection day and ploidy status of the embryo in libraries' concentration and NGS quality score.

All samples (n=51)	Factor(s)	P-value (significance <0.05)
Libraries' concentration	Clinic (1 and 2)	0.578
	SCM collection day (5 and 6)	**0.040**
	Ploidy status (euploid or aneuploid)	0.330
	Clinic * Collection day	0.145
	Clinic * Ploidy status	0.532
	Collection day * Ploidy status	0.636
	Clinic * Collection day * Ploidy status	0.466
NGS quality score	Clinic (1 and 2)	0.520
	SCM collection day (5 and 6)	**0.005**
	Ploidy status (euploid or aneuploid)	0.397
	Clinic * Collection day	0.579
	Clinic * Ploidy status	0.260
	Collection day * Ploidy status	0.639
	Clinic * Collection day * Ploidy status	0.164
After removal of known contaminated samples with female DNA (n=41)	**Factor(s)**	**P-value (significance <0.05)**
Libraries' concentration	Clinic (1 and 2)	0.914
	SCM collection day (5 and 6)	**0.011**
	Ploidy status (euploid or aneuploid)	0.156
	Clinic * Collection day	0.077
	Clinic * Ploidy status	0.665
	Collection day * Ploidy status	0.509
	Clinic * Collection day * Ploidy status	0.211
NGS quality score	Clinic (1 and 2)	0.906
	SCM collection day (5 and 6)	**0.002**
	Ploidy status (euploid or aneuploid)	0.328
	Clinic * Collection day	0.279
	Clinic * Ploidy status	0.541
	Collection day * Ploidy status	0.634
	Clinic * Collection day * Ploidy status	**0.033**

In the first ANOVA, where all the informative samples were included, the SCM collection day had a significant effect on both libraries' concentration and NGS quality score. Figure 26(A) shows that collecting the SCM on day-6 significantly increased libraries' concentration after WGA and significantly decreased NGS quality scores compared to day-5.

(A)



(B)

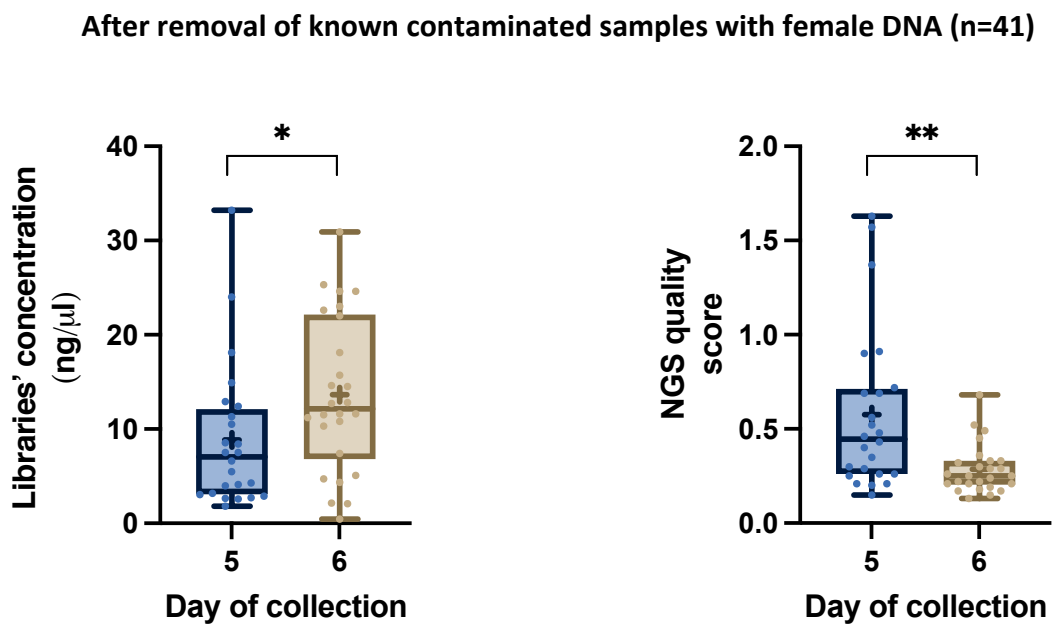


Figure 26. Libraries' concentrations and NGS quality scores comparison between day-5 and day-6 of SCM collection. All data points are represented with dots. '+' denotes the mean. '*' denotes a level of significance P <0.05; '**' level of significance P <0.005.

In the second ANOVA, where known male SCM samples contaminated with female DNA were removed, the P-values for the day of SCM collection for both libraries' concentration and NGS quality score demonstrated more significance, compared to the first group of samples (Table 31 and Figure 26(B)). Interestingly, the results of this second statistical analysis showed the interaction of clinic, collection day and ploidy status to also be significant.

The mean of the libraries' concentrations and NGS quality scores were compiled in Table 32 for comparison between SCM collection day, ploidy status and clinic. Although not significant, the results of the experiments showed higher library concentrations and lower NGS quality scores for SCM from euploid embryos compared to aneuploid. Despite being the comparison with the smallest differences, Clinic 1 showed higher library concentrations and lower NGS quality scores compared to Clinic 2.

Table 32. Libraries concentration and NGS quality score mean comparison. Excluding data from Clinic 3. SCM, spent culture media.

All samples				
	Libraries concentration (ng/µl)		**NGS quality score**	
SCM collection day	**5 (n=31)**	**6 (n=23)**	**5 (n=31)**	**6 (n=23)**
	9.08	13.45	0.59	0.31
Ploidy status	**Euploid (n=24)**	**Aneuploid (n=27)**	**Euploid (n=24)**	**Aneuploid (n=27)**
	11.62	11	0.42	0.51
Clinic	**1 (n=23)**	**2 (n=31)**	**1 (n=23)**	**2 (n=31)**
	11.70	10.38	0.40	0.52
After removal of known contaminated samples with female DNA				
	Libraries concentration (ng/µl)		**NGS quality score**	
SCM collection day	**5 (n=23)**	**6 (n=21)**	**5 (n=23)**	**6 (n=21)**
	9.13	14.49	0.53	0.28
Ploidy status	**Euploid (n=20)**	**Aneuploid (n=21)**	**Euploid (n=20)**	**Aneuploid (n=21)**
	13.09	11.30	0.36	0.45
Clinic	**1 (n=21)**	**2 (n=23)**	**1 (n=21)**	**2 (n=23)**
	12.29	11.14	0.39	0.43

It was then analysed if there was a correlation between libraries' concentration and NGS quality score (Figure 27). According to the Pearson r test r-value, a moderate negative correlation exists between libraries' concentration and NGS quality score (P<0.0001), as expected.

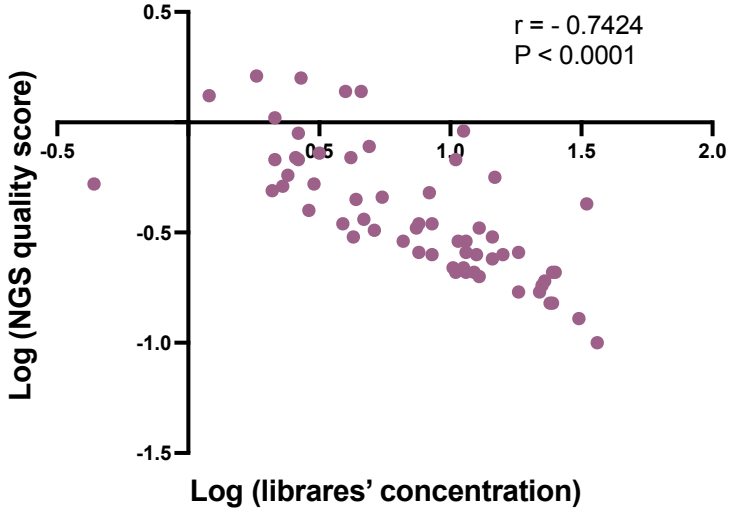


Figure 27. Correlation between libraries' concentration and NGS quality score. The r-value produced by the two-tailed Pearson r test shows a moderate negative correlation between the two log-transformed data sets with a significant P-value (95% confidence interval).

5.1.3.2. Chromosomal copy number analysis

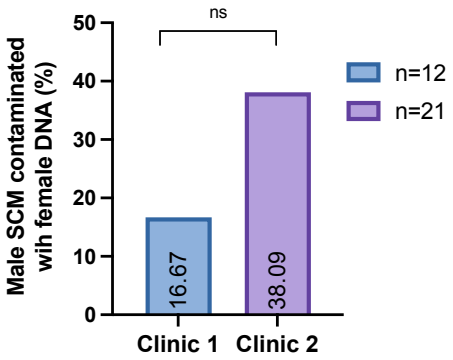
The karyotypes obtained from the SCM samples included for niPGT analysis and controls are found in **Appendix C**.

5.1.3.2.1. Female contamination in male SCM samples

When comparing the sex results of SCM with its corresponding TE biopsy, female DNA contamination was identified after gender mismatch in 29.73% (11/37) of the male samples. Figure 28(A) shows the individual contamination rates for Clinic 1 and Clinic 2. Clinic 3 was not included for lack of statistical representation due to the small number of

samples provided. The contamination rate shown by Clinic 2 (38.09%) is more than double that of Clinic 1 (16.67%). Nevertheless, the difference lacks statistical significance (P>0.05). Figure 28(B) gives more insight into the level of female DNA contamination. Even though the ratio of females to males is expected to be 1:1, Clinic 1 produced 62.07% females and Clinic 2 68.08%. These results agree with the higher rate of male samples contaminated with female DNA generated by Clinic 2. No gender mismatches were observed that indicated male DNA contamination in female SCM samples. With the data generated, it was impossible to determine which female samples were contaminated with female DNA and which male samples were contaminated with male DNA.

(A)



(B)

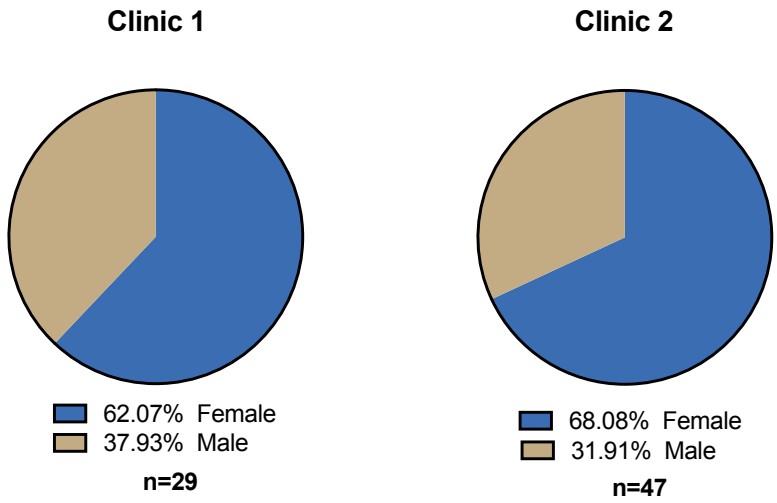


Figure 28. Female DNA contamination in SCM samples. 'ns' denotes not significant (P=0.259, unpaired *t*-test). CI, confidence interval.

5.1.3.2.2. Overall ploidy and full karyotype concordance

For three positive controls, the full karyotype concordance rate with the known karyotype was 100% (4/4) (Table 33).

Table 33. Full karyotype concordance rates for the four positive controls between niPGT and known karyotype.

Control ID	Known karyotype	SCM niPGT result
9P	46,XX	46,XX
D05	47,XX,+18	47,XX,+18
F03	46,XY,-10p14p12.1	46,XY, -10p14p12.1

The following analyses excluded the known contaminated male SMC samples with female DNA. When analysing male samples only, the present niPGT protocol presented concordance rates between SCM and complement TE biopsy of 50% for full karyotype and 69.23 % for overall ploidy (Table 34). Full karyotype and aneuploidy concordance rates increased when compared to the group that includes all samples. This is due to the effect of the karyotypically normal female contamination, likely present in female SCM samples.

Table 34. Concordance rates between SCM and complement TE biopsy chromosomal copy number analysis results.

Group	Full karyotype concordance rate, %	Euploidy concordance rate, %	Aneuploidy concordance rate, %	Overall ploidy concordance rate, %
All samples	44.78% (30/67)	72.97% (27/37)	56.66 % (17/30)	65.67% (44/67)
Male samples only	50% (13/26)	62.50% (10/16)	80% (8/10)	69.23% (18/26)

5.1.3.2.3. Sensitivity, specificity, PPV and NPV

One group in Table 35 included all samples and the other excluded female samples. Known contaminated male SCM samples with female DNA were excluded. When all samples were considered, the false positive rate decreased and the false negative rate increased compared to the male samples only. These results are likely because of karyotypically normal female DNA contamination in female samples. Consequently, the PPV decreased and the NPV increased for the male samples-only group. The sensitivity is higher for the male samples-only group (80% vs. 56.67%) but the specificity is lower (62.50% vs. 72.97%).

Table 35. Sensitivity, specificity, PPV and NPV results. NPV, negative predictive value; PPV, positive predictive value.

Group	False positive rate, %	False negative rate, %	Sensitivity, %	Specificity, %	PPV, %	NPV, %
All samples	27.03% (10/37)	43.33% (13/30)	56.67%	72.97%	62.96%	67.50%
Male samples only	37.50% (6/16)	20% (2/10)	80%	62.50%	57.14%	83.33%

5.1.3.2.4. Aneuploidy reciprocals between paired SCM and complement TE biopsy

Two potential reciprocals were detected between the chromosomal number analysis results of SCM and their complement TE biopsy (Table 36). The two SCM and TE samples were collected on day-6.

Table 36. Aneuploidy reciprocals findings between paired SCM and complement TE biopsy. Mos,mosaic.

SCM collection day	Clinic	Processing time (days)	SCM niPGT result	TE biopsy PGT-A result
6	2	1	46,XX, +4p16.3p15.32	45,XX, -4p16.3p15.33, -12
6	2	1	46,XY, +5q14.3q23.1, +12q23.3q24.23	46,XY, -5q13.3q23.1(mos), -5q23.1q35.3

5.1.4. Discussion

Studies exploring niPGT as an alternative for invasive PGT-A show discrepancies in the ploidy concordance rates between TE biopsy and SCM. It is likely that the embryo is being manipulated and this facilitates the release of nuclear DNA into the media (Rubio et al., 2019a, Leaver and Wells, 2020). This chapter evaluates the accuracy of a truly non-invasive PGT protocol where embryo manipulation is minimised before SCM collection.

5.1.4.1. DNA informativity in SCM

The concentration of libraries after WGA was known for 86 out of the 128 total day-5 and day-6 SCM samples. Amplification was successful for 70.93% of those libraries. From these 61 samples, 85.24% produced the maximum acceptable NGS quality score after sequencing. That is to say, only 60.46% of the initial 86 samples produced informative results.

The data on DNA amplifiability in this study differs from others. Capalbo et al. (2018) studied the potential of SCM as an alternative to TE biopsies for PGT-M. Amplification was successful if the targeted loci specific for each PGT-M case were

detected during RT-qPCR analysis. They reported that only 10.30% of the SCM failed to amplify. Using a similar approach to the one in this study, Xu et al. (2016) performed WGA on SCM from ICSI-conceived day-5 embryos using the Multiple Annealing and Looping Based Amplification Cycles (MALBAC) method. They then performed and NGS and obtained ploidy status information for 100% of the samples (42/42). However, in a more recent study by Yin et al. (2021) only 80% (60/75) of the blastocyst SCM samples showed successful WGA using a modified MALBAC method (Yikon Genomics, China) and 78.67% (59/75) produced interpretable NGS results. The authors did not specify if the blastocysts used for the study were conceived with conventional IVF or ICSI. The SCM included in the last two studies mentioned were retrieved from thawed blastocysts.

Regarding the chromosomal status results of the present niPGT protocol in the male samples-only group, 50% of SCM samples presented the exact same karyotype to their corresponding TE biopsies and there was an overall ploidy status concordance of 69.23%. Concordance rates between TE biopsies and SCM have varied greatly between studies since non-invasive strategies of embryo genetic analysis were first explored. For instance, Liu et al. (2017) performed MALBAC followed by NGS in day-5 SCM from ICSI-conceived embryos without previous freezing, obtaining a DNA detection rate in 90.90% of the samples. They found an identical karyotype concordance rate between SCM and TE biopsies was 64.52% and the concordance of ploidy classification was 83.87%. In contrast, Huang et al. (2019) reported much higher agreement between SCM and the whole embryo, after conducting MALBAC and NGS on ICSI-conceived and thawed day-5 and day-6 blastocysts. They identified an identical karyotype concordance rate of 83.30% and concordance of ploidy classification of 93.80%. They reported 92.30% of the media

samples produced interpretable results. TE biopsy was performed before the vitrification of the embryos.

Inconsistency in levels of DNA amplification, karyotype and ploidy classification concordance rates between studies may result from varying degrees of embryo manipulation during embryo culture and before SCM collection. For example, prior to media recovery Huang et al. (2019) vitrified, assisted-hatched and biopsied blastocytes. The study presented in this chapter only included SCM samples from embryos that had not been vitrified and biopsied previously. With a similar study design to this, Rubio et al. (2019a) included SCM from day-4 to day-7 embryos that were not subjected to vitrification or previous biopsy. SCM was subjected to WGA with the Ion Reproseq PGS Kit (Thermo Fisher Scientific, USA) and NGS for chromosomal status analysis. The reported overall karyotype and ploidy status concordance rate between SCM and TE biopsy was 63.90% and 78.70%, respectively, showing closer proximity to the values of this study. The same research group subsequently performed a multi-centre study including a total of 1,108 SCM samples and obtained the same overall ploidy concordance rate (78.20%) (Rubio et al., 2020). Thus, the comparisons of the level of embryo handling and the resulting concordance between studies support the hypothesis that embryo manipulation increases cell lysis and release of genetic material to the media, enhancing DNA amplification levels and concordance rates.

This study showed a false positive rate of 37.50% and a false negative rate of 20%. As reviewed by Leaver and Wells (2020), inflated aneuploidy levels reported in BF may result from studies attempting to define the chromosomal status in samples where WGA was inadequate, introducing artefactual aneuploidies during sequencing and analysis. Despite evidence suggesting that in humans, the quality of gDNA found in SCM is better

and in higher concentrations than in BF, the quality of the genetic material found in SCM is likely suboptimal. Rubio et al. (2019a) reported false positive rates of 29.60% and 8.60% for day-5 and day-6/-7, respectively. The authors did not indicate the numerical threshold during analysis, unlike the present study, but samples were assessed for their informativity depending on successful amplification and interpretable NGS results. In addition, informative samples that presented ≥6 aneuploidies were classified as chaotic. When chaotic samples were removed, their false positive rate was reduced by almost half. Yin et al. (2021) reported a false positive rate of 10.17% in informative samples. Similar to Rubio et al. (2019a), informativity was based on WGA success and interpretable NGS data. Huang et al. (2019) excluded 'noisy' profiles from analysis and reported a false positive rate of 20%. Evidently, there is also a discrepancy in the level of false positives amongst studies. This suggests more rigorous and standardised quantitative methods for sample inclusion for analysis are needed to eliminate the opportunity for subjectiveness. It is likely more stringent quantitative criteria for WGA and NGS quality control metrics would have led to improved concordance rates in this study but with a significantly further decrease in the informative sample size.

5.1.4.2. DNA origin in SCM

Statistically, samples collected on day-6 presented significantly higher library concentrations and lower (or superior) NGS quality scores. Other studies have also studied these effects. For instance, Rubio et al. (2019a) observed increased sample informativity when SCM was collected on day-6 or day-7 compared to day-5. Their results can be compared to those of this study due to the true non-invasiveness of their experimental design. These findings are in line with the reported increase in the levels of

nuclear and mitochondrial DNA in SCM at the blastocyst stage compared to cleavage stage embryos by Hammond et al. (2017). This can be explained by the apoptotic cells originating in the fast-growing blastocyst as a normal development process (Hardy et al., 2003)

In addition to introduced artefacts due to suboptimal WGA, the inflated aneuploidy reported in SCM could be explained by the potential source of DNA found in these samples. Early embryos possess mechanisms for eliminating commonly occurring aneuploid cells via arrest and apoptosis (Babariya et al., 2017, Yang et al., 2020). The nuclear products of the eliminated cells are likely to end in the surrounding embryo culture media. Babariya et al. (2017) reported a reduction in the frequency of segmental aneuploidies at the blastocyst stage as a result of the process of self-correction. These results further support transferring the embryo on day-4 of culture to a new drop of media if performing niPGT, which aims to eliminate traces of maternal contamination. This study followed this approach but the false positive and female contamination rate was very high. Lane et al. (2017a), in fact, studied the ability of cfDNA from SCM to inform about the ploidy of the blastocyst when the media was collected at different stages of development. The authors observed an increased concordance rate (>95%) when the embryos were transferred to a fresh drop of media on day-4 compared to day-3 with no difference between incubation periods (the time the embryo sits in the media) of 24 and 48 hours.

Evidence of self-correcting mechanisms in preimplantation embryos has led to hypotheses that aneuploid embryos release more DNA into the media than euploid embryos. However, this study found no statistical significance in the differences in library

concentrations and NGS quality scores between karyotypically normal and abnormal embryos, in line with other studies (Vera-Rodriguez et al., 2018).

Embryos included in this study were strictly conceived by ICSI to avoid cumulus cell and sperm contamination of the SCM samples. In addition, embryos were carefully washed and transferred to a fresh drop of media on day-4 and the SCM recovered was placed in UV-irradiated tubes. Despite the strict embryological measures followed to avoid it, female DNA contamination was observed in 29.73% of male SCM samples. The difference in percentages between clinics was not statistically significant (16.67% and 38.09% for Clinic 1 and Clinic 2, respectively). Nevertheless, one clinic presenting more than double the samples with confirmed female contamination suggests a potential effect of the clinic of origin, although the sample size was small. Regarding ploidy concordance, and in line with the above results, the multi-centre study from Rubio et al. (2020) found no significant differences between clinics. Another key finding of this thesis is that no gender mismatches were observed that indicated male DNA contamination in female SMC samples. This could mean that the female DNA contamination observed likely came from cumulus cells rather than external DNA contamination from handling reagents and materials during embryo culture and SCM collection. In contrast to the results in this chapter, Rubio et al. (2019a) reported contamination in only 9.26% of their informative samples. In addition, they found one female SCM sample contaminated with male DNA, providing evidence of variable sources between studies and clinics.

A critical limitation of this chapter was the inability to detect DNA contamination beyond that indicated by NGS gender mismatches between niPGT and PGT-A. For this reason, this study was unable to identify female or male contamination in female or male SCM samples. Despite this constraint, it is possible to obtain an insight into the female

DNA contamination impact on the assay sensitivity by comparing the reduction of more than half in the false negative rate between the groups of samples after removing confirmed male SCM contaminated with female DNA (43.33%) and the group where only male samples without confirmed female DNA contamination were included (20%). Vera-Rodriguez et al. (2018) extensively studied maternal DNA contamination in day-3 to day-5 SCM samples with SNP sequencing in addition to NGS chromosomal number analysis and, strikingly, discovered significant maternal contamination in all samples, even in samples where NGS failed to predict it. They established a link between SCM/TE biopsy discordant results and a higher percentage of maternal DNA found in the SCM. This information helps to better understand the origin and composition of the cfDNA in SCM and highlights the importance of including a methodology that analyses the origin of the cfDNA found in SCM to enhance chromosomal analysis accuracy. Future studies could assess the compatibility of methodologies such as DNA fingerprinting in cfDNA from SCM when maternal DNA is unavailable for SNP sequencing.

Although newly published data has provided a better insight into the clinical value of niPGT for aneuploidies, there are still many questions about the origin of the cfDNA found in SCM. Concordance with TE biopsies remains suboptimal for truly non-invasive PGT. Yet, interestingly, according to the followed-up clinical outcomes in Vera-Rodriguez et al. (2018), ongoing implantation rates of the corresponding transferred embryos were three times higher when both TE biopsy and SCM showed euploid results compared to when TE biopsy produced euploid results and the corresponding SCM showed chromosomal abnormalities. This data suggests additional clinical value indicators should be included for future niPGT studies.

5.1.4.3. Evaluation of the clinical application of the PG-Seq Rapid Non-Invasive PGT methodology

The PG-Seq Rapid Non-Invasive PGT kit manufacturer recommends laboratories determine their own PPV and NPV to understand their SCM concordance rate with TE biopsy. This study produced suboptimal concordance rates, which resulted in a PPV value of 57.14% and a NPV of 83.33%. Ultimately, the accuracy of the niPGT depends on the sample's DNA quality and quantity, and the methodology used for sample processing and analysis.

The PG-Seq Rapid Non-Invasive PGT uses degenerate oligonucleotide-primed PCR (DOP-PCR) for WGA as the first step for library preparation. Reports have agreed on the suitability of this method and MALBAC for accurate CNV determination at a single-cell level. While DOP-PCR has been shown to have the best cell-to-cell reproducibility in single-cell studies, the unmappable read rate is reported to be lower for MALBAC and Multiple Displacement Amplification (MDA) (Hou et al., 2015, Huang et al., 2015). Nevertheless, a higher unmappable read rate does not explain the difference in amplification levels observed among samples of this study compared to others. In fact, Yin et al. (2021) showed the closest level of WGA to the present study and the authors employed a modified MALBAC protocol (ChromInst). As already discussed, Rubio et al. (2019a) performed truly niPGT by minimising the interventions on the embryo. Their amplification levels are remarkably higher than those presented in this chapter (94.80% vs. 70.93%). The authors used a similar SCM volume input to this study (8 - 10µl vs. 7µl), and in an attempt to increase DNA yield, they modified the Ion ReproSeq PGS WGA protocol by extending the thermal cycles. This may explain the high successful WGA rates. However, it is unlikely that higher DNA yields improve quality control metrics,

reflected in the presence of chaotic chromosome plots (in 17 out of 108 samples) and suboptimal concordance rates with TE biopsy. This suggests that the low amplification, informative sample and concordance rates in this chapter may be elucidated by the reduced interventions on the embryo that facilitate the release of DNA to the media and not the method of WGA employed.

As indicated in this chapter's methods section, any apparent mosaicism was disregarded for karyotype comparisons between SCM and TE biopsies due to the high incidence of intermediate chromosome number profiles in SCM samples. In fact, the karyotypically normal commercial gDNA used as control also produced intermediate copy numbers for some chromosomes. This raises questions about the methodology's specificity, even when high quality gDNA is used as template. Many previous studies that explore the accuracy of niPGT have included mosaicism in their concordance analysis between SCM and TE biopsy or whole embryo (Xu et al., 2016, Liu et al., 2017, Huang et al., 2019, Yin et al., 2021) with varying levels of agreement. The PGT-A mosaicism result management in the clinic has recently been restated after reports of technical variation in the assay due to artefacts introduced by WGA and NGS in the results that sometimes make them indistinguishable from mosaicism leading to false positives (Capalbo et al., 2017). Pregnancy outcome results have shown that the level of whole chromosome mosaicism (or percent of aneuploid cells) and the number of affected chromosomes positively impact unfavourable outcomes (Viotti et al., 2021). The accumulation of supporting evidence that for some TE biopsies the presence of mosaicism can negatively affect the pregnancy outcome has led to the creation of evidence-based guidelines for ranking the level of mosaicism in embryos for transfer prioritisation. Conversely, the PG-Seq Rapid Non-Invasive PGT manufacturer recommends only calling for full chromosome

gains or losses when using SCM. According to Capalbo et al. (2017), performing a double biopsy that would show reciprocal aneuploidies is the most accurate analysis to identify real mosaicism in TE. This study identified two SCM samples with potential reciprocal aneuploidies to their corresponding TE biopsy. This could provide evidence of real mosaicism shown in the SCM samples. However, whole chromosome reciprocal aneuploidies would bring more compelling evidence, but most importantly, it is difficult to expect genuine reciprocity from two different types of samples and two different methodologies. In addition, the DNA origin composition findings in SCM from Vera-Rodriguez et al. (2018) suggest that NGS technologies may not be sensitive enough to detect mosaicism in media samples where, according to their study, the median fraction of embryonic DNA is only 8%. More data is needed from stringent niPGT validations to better understand intermediate chromosome number profiles in SCM as well as follow-up research to understand the clinical value of these findings.

5.1.4.4. Conclusion

Non-invasive alternatives to routine PGT-A, such as the analysis of cfDNA in SCM of embryonic origin, have been proposed. Nonetheless, there are discrepancies in ploidy concordance rates between TE biopsy and SCM reported by numerous studies. This chapter aims to evaluate the accuracy of a truly non-invasive PGT protocol where embryo manipulation is minimised before SCM collection.

The results obtained show a decrease in overall ploidy and karyotype concordance rates between TE biopsy and SCM when the embryo is minimally manipulated (no previous vitrification or biopsy). This data reinforces the hypothesis that interventions increase cell lysis and DNA release to the media, which results in an improved agreement

between TE biopsy and SCM. DNA origin, however, may also play a role. Due to the poor quality and quantity of embryonic DNA commonly found in the media, WGA can be suboptimal, introducing sequencing artefacts that sometimes lead to false positives. In addition, noisy profiles are difficult to assess accurately. Evidence from this truly niPGT study and others show increased informativity when SCM samples are recovered at later stages of embryo development (day-6 vs. day-5 or earlier). Cumulus cell carry-over remains a significant problem contributing to false negatives even when rigorous embryological methods are applied to avoid it.

Stringent and standardised criteria for DNA amplifiability and NGS quality control metrics for the inclusion of SCM samples for analysis could help to avoid false positive rates inflation. In addition, more studies are needed to better understand the role of mosaicism results with SCM, due to the value of this diagnosis in the clinic. Regarding niPGT validation studies, whole embryos would ideally be employed as a reference to assess SCM informativity about their chromosomal status, if accessible. Future studies exploring niPGT and SCM should also employ SNP sequencing or DNA fingerprinting, if compatible with cfDNA in SCM, to analyse DNA origin and composition during different stages of preimplantation development.

Even though, in some instances, niPGT has demonstrated encouraging results, its accuracy is currently inadequate to justify its routine clinical use. This chapter has provided some suggestions as to the further research needed to improve its applicability in clinical settings.

Chapter 6. Development of non-invasive methods for Y chromosome detection

6.1.1. Introduction

Whilst more desirable for patients, non-invasive methods for determining embryonic sex are more challenging due to the low abundance, and fragmented nature, of cfDNA. This chapter explores two different approaches to non-invasive sex determination. The first builds on the work of Zimmermann et al. (2005), who studied the presence of cfDNA of embryonic (or foetal) origin in the mother's blood. They measured circulating placental cfDNA in the maternal plasma from first trimester male gender pregnancies (ranging from week-11 to week-13 gestational ages) by targeting *SRY*, a single copy gene and an 84 bp long fragment from *DYS14*, a multicopy gene located in the human male-specific region of the Y chromosome (MSY). They found that *DYS14* had a 10-fold lower detection limit than the *SRY*, demonstrating that MSY multicopy genes can be used as Y chromosome markers for non-invasive foetal sex determination.

The second approach draws on Palini et al. (2013), who produced the first report of amplifiable embryonic cfDNA found in BF, suggesting its potential for comprehensive genetic testing. The study focused on targeting the multicopy gene, *TSPY1*, located in the human MSY and the single copy autosomal gene, *TBC1D3*.

This chapter separately introduces each study, the materials and methods, and results, before bringing them together for a comparative discussion.

6.2. Project 1. Y chromosome detection from circulating placental cfDNA in early pregnancy

6.2.1. Introduction

6.2.1.1. Prenatal sex determination

Prenatal sex determination can be critical to identifying X-linked genetic disorders in the foetus (Figure 29). If the pregnancy gender is male, further prenatal genetic testing, involving CVS or amniocentesis, are used to determine the genotype (Miura et al., 2011).

Traditionally, CVS and ultrasound have been employed for determining the foetal sex. However, these methods have significant drawbacks. Ultrasounds are only reliable after week-13 and CVS after week-11 of gestation. Pregnant women carriers of serious X-linked disorders naturally seek an early diagnosis. As the foetus becomes more developed, any procedure becomes potentially traumatic for the mother. Moreover, CVS is an invasive procedures with a miscarriage rate ranging from 0.50% to 1% (Tabor and Alfirevic, 2010).

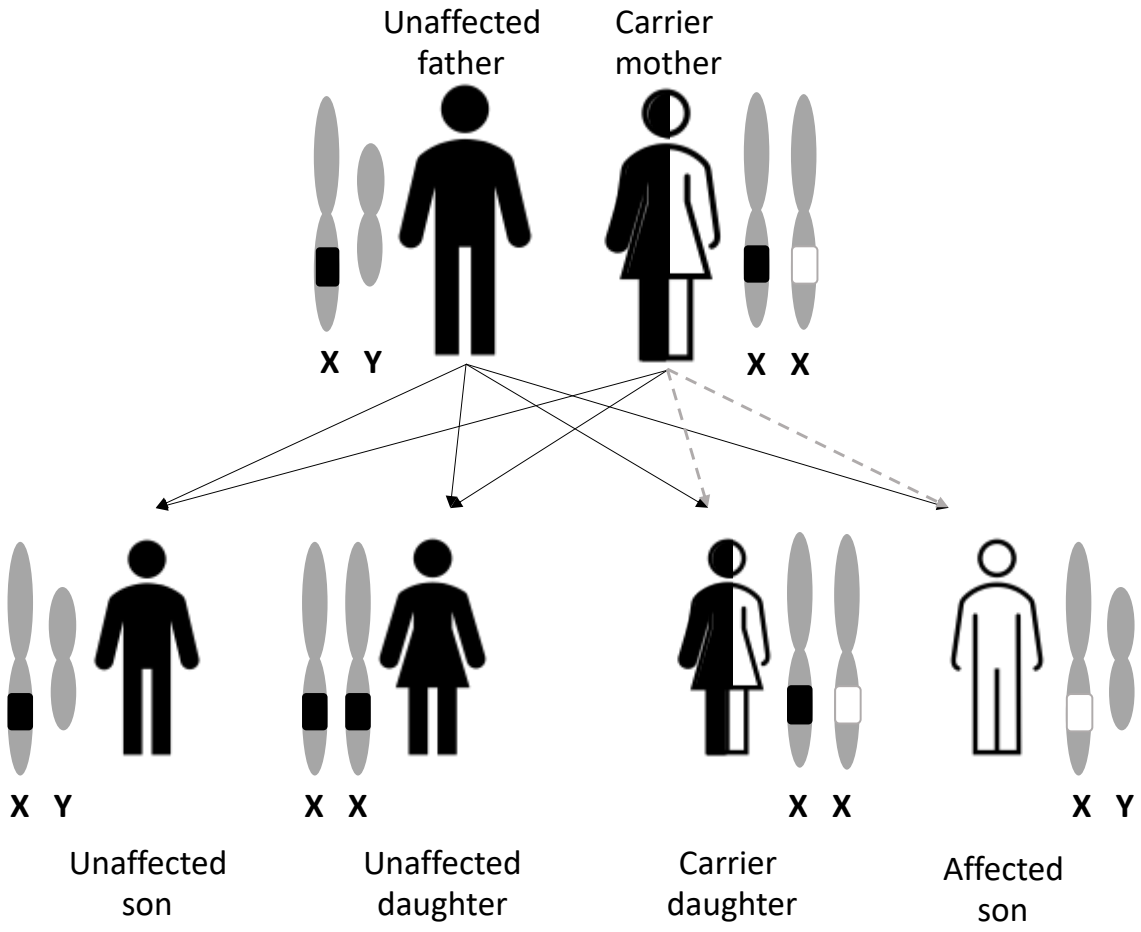


Figure 29. X-linked recessive inheritance. The results of prenatal sex determination show if the foetus is female and carries either the wild-type genotype or a heterozygous genotype (carrier of the X-linked mutation, as the mother), or if it is male and carries either the wild-type genotype or affected with the X-linked mutation.

6.2.1.2. Circulating placental cfDNA and non-invasive prenatal testing (NIPT)

Since the discovery of circulating cfDNA of foetal origin in the maternal plasma by Lo et al. (1997), non-invasive prenatal testing (NIPT) became a common service offered in clinics for 'first-line' aneuploidy screening because it has a high PPV in detecting trisomy of chromosomes 13, 18 and 21. It can also detect other autosomal aneuploidies, sex chromosome abnormalities and some microdeletions (Hui and Bianchi, 2020, Kypri et al., 2022). It can, therefore, be used as an alternative to CVS for early foetal sex determination. Routinely, NIPT is performed from week-10 (any earlier and it may affect

the assay reliability due to insufficient placental cfDNA in the plasma (Shi et al., 2015)) and employs NGS to analyse foetal cfDNA. Foetal circulating cfDNA mainly originates from trophoblast cells (Alberry et al., 2007) and is detectable from the first trimester of pregnancy onwards (Lo et al., 1998).

A major determinant of NGS-based NIPT assay's sensitivity is the foetal fraction. This fraction indicates the percentage of the total plasma cfDNA that is of foetoplacental origin. Higher foetal fractions provide greater confidence in the analysis results (Hui and Bianchi, 2020). This value varies significantly, but the average foetal fraction between week-10 and week-20 of gestation typically ranges from 10% to 15% (Canick et al., 2013).

Researchers have started to explore the opportunities for PCR-based non-invasive assays to detect Y chromosomes in maternal plasma (Ariga et al., 2001, Rijnders et al., 2003, Kolialexi et al., 2012, Qiu and Liu, 2019, Mahdavi et al., 2019). This leverages advances in PCR technology efficiency, specificity and sensitivity (Elnifro et al., 2000, Markoulatos et al., 2002) and better characterization of the human Y chromosome (Lahn and Page, 1997), for a simpler and cheaper approach to NIPT.

6.2.1.3. Study aims

This chapter aims to develop a highly sensitive PCR-based tool to detect DNA from foetal origin, present in multiple copies, in the mother's circulating cfDNA with a potential clinical application for sex determination.

The specific aims are:

1) To optimise a Y chromosome assay that can detect low amounts of target DNA present with significant amounts of maternal DNA background.

2) To assess the Y chromosome assay sensitivity to reduced volumes of maternal blood (as small as 0.5 ml).
3) To reveal whether the optimised protocol could be applied to stages of pregnancy earlier than week-10.

6.2.2. Materials and methods

6.2.2.1. Ethics statement

The blood samples were taken from patients undergoing NIPT who provided informed consent to donate any blood surplus to the research and development of new non-invasive prenatal genetic testing and screening technologies.

6.2.2.2. Sample characteristics and collection

Participants' gestational age was 10 weeks. Ten ml of blood were obtained through venepuncture in a Streck Cell-Free DNA BCT tube (Streck, USA). Samples were shipped at room temperature, placed at 4°C upon reception and processed within seven days.

6.2.2.3. Y chromosome assay

The Y chromosome assay and primer design were developed by Ph.D. Dhruti Babariya following **Chapter 2. General materials and methods**. The assay targets six MSY multicopy genes and amplifies them in a multiplex PCR (Table 37). Targeting a group of multicopy genes with short amplicon size increases the likelihood of successfully amplifying fragmented DNA. Amplicons differed in size to allow band distinction in the gel

electrophoresis. The assay was originally developed and the PCR conditions optimised to detect the Y chromosome in BF and SCM (Table 38).

Table 37. Human Y chromosome assay and primer design.

MSY gene	Number of copies	Amplicon size (bp)	Primer sequences
DAZ	4	77	F: 5' CATCAACAGCAGTGCAGACT 3' R: 5' AATGTGAGGGACCTTAGCCG 3'
RBMY	6	138	F: 5' GCACCTGCCACGCATATAGT 3' R: 5' TACCCAGAGAAGGCGGATTC 3'
TSPY	~35	101	F: 5' AGTGGAAGGAAGATGGTGGG 3' R: 5' GGAGTGGAATGAGAAGCCCT 3'
CDY	4	118	F: 5' ACACCCACTAATGTCGCAGA 3' R: 5' GTCTGTTGTCCCATTGGCTG 3'
BPY2	3	112	F: 5' TGCAGTTGTACCCAGGTGAT 3' R: 5' CAGGCAGAAAGGGAGACTCA 3'
FAM197Y9	~35	121	F: 5' TCTCCCAGGCCATTCATGTT 3' R: 5' TCCGTGTGAGGTGGACTAAC 3'

Table 38. Thermal program for the human Y chromosome assay for BF and SCM.

Step	Temperature	Time	Number of cycles
Initial denaturation	96 °C	1 minute	1
Denaturation	94 °C	15 seconds	55
Annealing	61 °C	15 seconds	
Extension	65 °C	45 seconds	
Final extension	65 °C	2 minutes	1
Hold	4 °C	∞	1

6.2.2.4. Experimental design for the Y chromosome NIPT method development

The optimisation of the Y chromosome NIPT was divided into three phases (Figure 30).

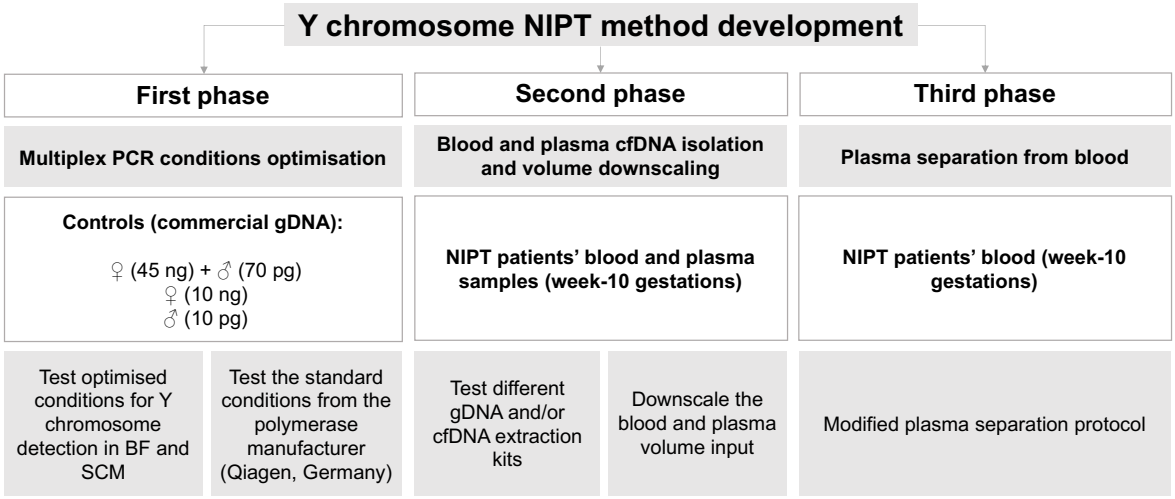


Figure 30. Experimental design for the method development of the Y chromosome NIPT. BF, blastocoel fluid; SCM, spent culture media.

First phase: The multiplex PCR conditions were optimised using commercial human female and male gDNA as controls. One control was a mix of 45 ng of female gDNA and 70 pg of male gDNA to mimic the circulating cfDNA composition found in maternal plasma. Ten ng of female gDNA and 10 pg of male gDNA were also included as two separate controls. Two different PCR thermal programs were used as a reference: the Y chromosome assay for BF and SCM and the polymerase manufacturer's universal, or standard, program.

Second phase: A surplus of pregnant participants' blood samples and separated plasma obtained for routine NIPT was used for testing two different DNA extraction kits and downscaling the blood and plasma volume input.

Third phase: A modified method was developed for plasma separation that can process blood using a standard microcentrifuge.

6.2.2.5. Plasma separation

The modified plasma separation protocol separates small volumes of blood (up to 2 ml). The 2 ml tubes containing the blood samples were centrifuged for 10 minutes at 1,600 x g speed in a microcentrifuge and the supernatant, or plasma, was collected using a micropipette and transferred to a 1 ml microcentrifuge tube (Eppendorf, Germany). The plasma was processed immediately for cfDNA isolation or stored at -80 °C for later use.

6.2.2.6. cfDNA extraction

The QIAmp DNA Micro and the MagMAX Cell-Free DNA Isolation Imprint Modification kits were tested for cfDNA isolation from 100 µl to 250 µl of blood or plasma, according to the protocols in **Chapter 2**. In all instances, DNA was eluted in 13 µl of elution buffer and measured for DNA concentration using the Qubit 1X dsDNA HS kit. The isolated DNA was immediately amplified or stored at -20 °C for later use.

6.2.2.7. Multiplex PCR

The six MSY loci in Table 37 were amplified in a multiplex, as described in **Chapter 2**. The optimised Y chromosome assay for BF and SCM (Table 38) was tested in parallel with the polymerase manufacturer's (Qiagen Multiplex PCR) universal thermal program, in **Chapter 2**. PCR protocols were tested in 30 µl reactions maintaining the annealing temperature of 61°C from the optimised protocol for BF and SCM as well as the high number of cycles (X55) since the targets were also expected to be in low concentrations. Ten ng of commercial female and male gDNA (Promega, USA) and a no-template control were used as controls in all reactions.

6.2.2.8. Data analysis and concordance rates

The products of the Y chromosome assay amplified in the multiplex PCR were analysed in a 2% agarose gel electrophoresis (2 g of agarose) at 100 volts for 1 hour and 20 minutes. Concordance rates of the gender determination between the Y chromosome NIPT and the complement routine NGS-based NITP were calculated with the results obtained in this study's second and third phases.

The Y chromosome NIPT assay results were classified as follows:

a) 'Male' was defined when there was amplification of the MSY genes.
b) 'Female' was defined when there was no amplification of the MSY genes but there were non-specific bands in the gel electrophoresis.
c) 'No-amplification' was defined when there were no bands in the gel electrophoresis.

The sensitivity and specificity of the test to detect the Y chromosome were calculated using the following equations:

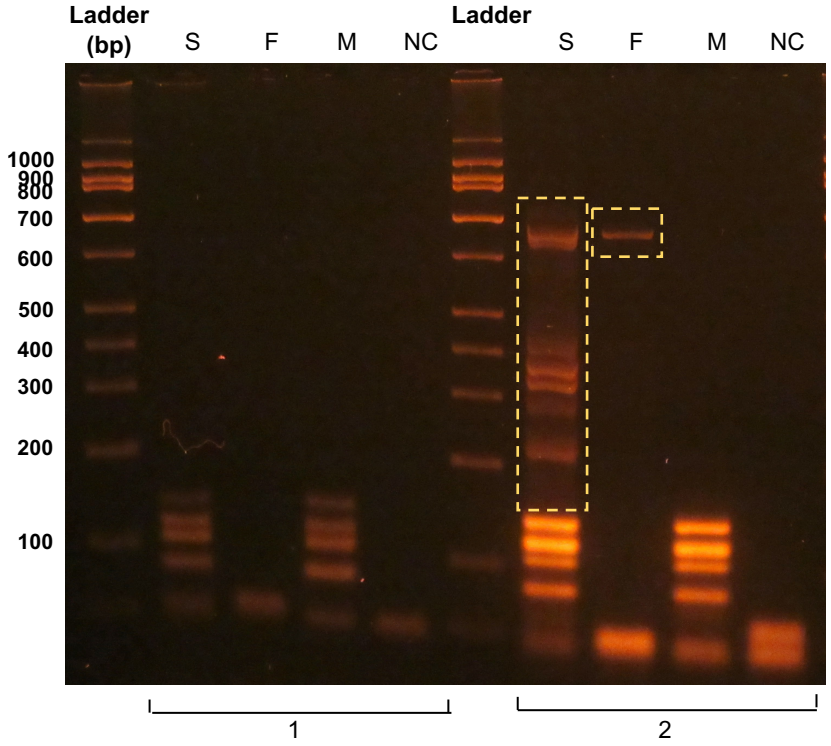
$$Sensitivity\ (\%) = \frac{True\ positives}{True\ positives + False\ negatives}$$

$$Specificity\ (\%) = \frac{True\ negatives}{True\ negatives + False\ positives}$$

6.2.3. Results

6.2.3.1. First phase: Multiplex PCR conditions optimisation

To mimic the constitution of the DNA found in the mother's circulating cfDNA, a mix of 70 pg of commercial male gDNA and 45 ng of commercial female gDNA was initially used as a 'sample'. Two different thermal protocols were tested: the protocol for detection of the Y chromosome in BF and SCM, and the Qiagen Multiplex PCR universal protocol.



Controls	PCR conditions (X55 cycles)
S: ♀ (45 ng) + ♂ (70 pg) **F:** ♀ (10 ng) **M:** ♂ (10 pg) **NC:** no-template control	**1:** Y chromosome detection in BF and SCM **2:** Standard for Qiagen Multiplex PCR polymerase

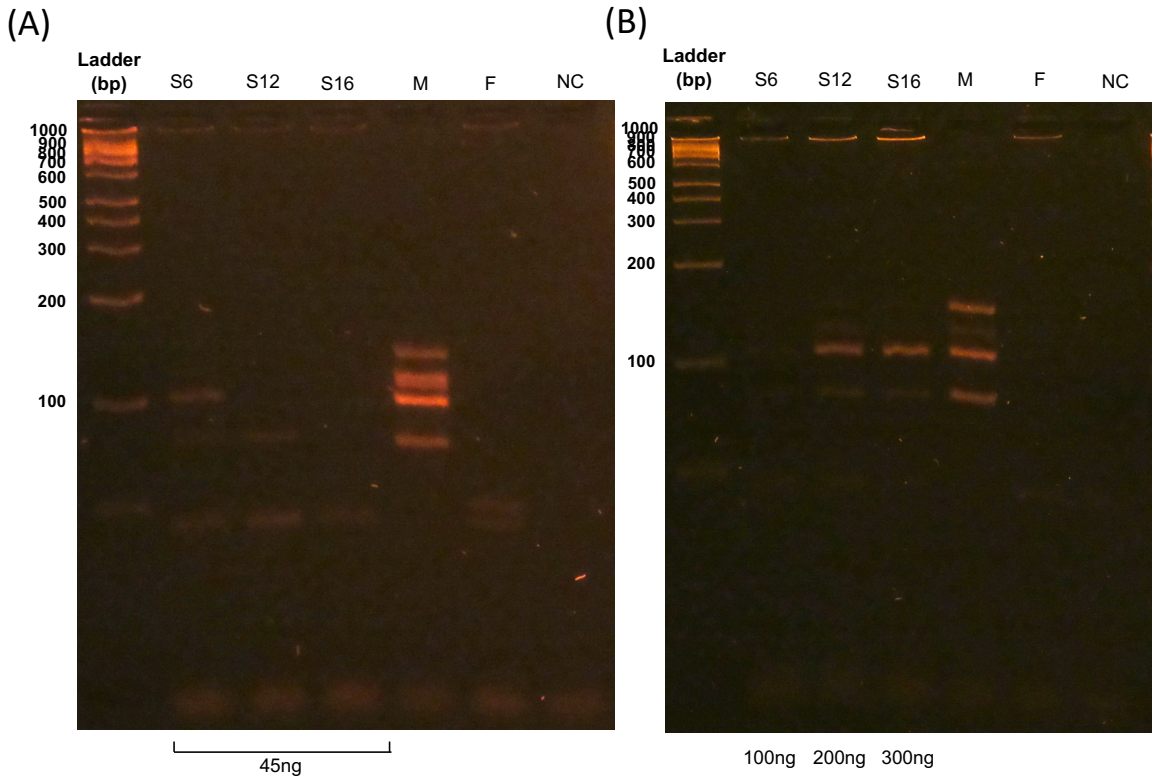
Figure 31. Agarose gel electrophoresis of the amplification of controls using two different thermal protocols. The yellow dashed boxes indicate non-specific amplification**.** Multiplex PCR products were run in a 2% agarose gel electrophoresis at 100 volts for 1 hour and 20 minutes.

The expected band sizes appeared in the gel electrophoresis (Figure 31), confirming the amplification of the targeted MSY loci. A band overlap was expected due to the similarities between some of the amplicon sizes (e.g., the similarity between the *CDY* amplicon's size of 118 bp and the one of *BPY2,* 112 bp). Even though the band intensity of the DNA amplified with the standard polymerase manufacturer's thermal protocol shows higher concentrations, there was a high degree of non-specific amplification in both the sample and the female control DNA. The no-template control also produced a more intense primer-dimer band. For this reason, the Y chromosome detection in BF and SCM protocol was used for the following amplifications. Due to the detection of non-specific amplification produced by female DNA background in the sample, the input female DNA control was increased from 10 to 45 ng regardless of the thermal protocol used.

6.2.3.2. Second phase: Blood and plasma cfDNA isolation and volume downscaling

6.2.3.2.1. QIAmp DNA Micro Kit - Protocol: Isolation of Genomic DNA from Small Volumes of Blood

This protocol is indicated for isolating gDNA from small volumes of blood. Initially, 200 µl of blood were used from mostly male gender pregnancies (confirmed by routine NIPT). The concentration of DNA after isolation for samples S6, S12 and S16 were 108 ng/µl, 33 ng/µl and 46 ng/µl, respectively. Increasing the input DNA from 45 ng to 100 ng, aiming to mitigate the low amplification levels, did not improve band intensity levels for S6 (Figure 32).

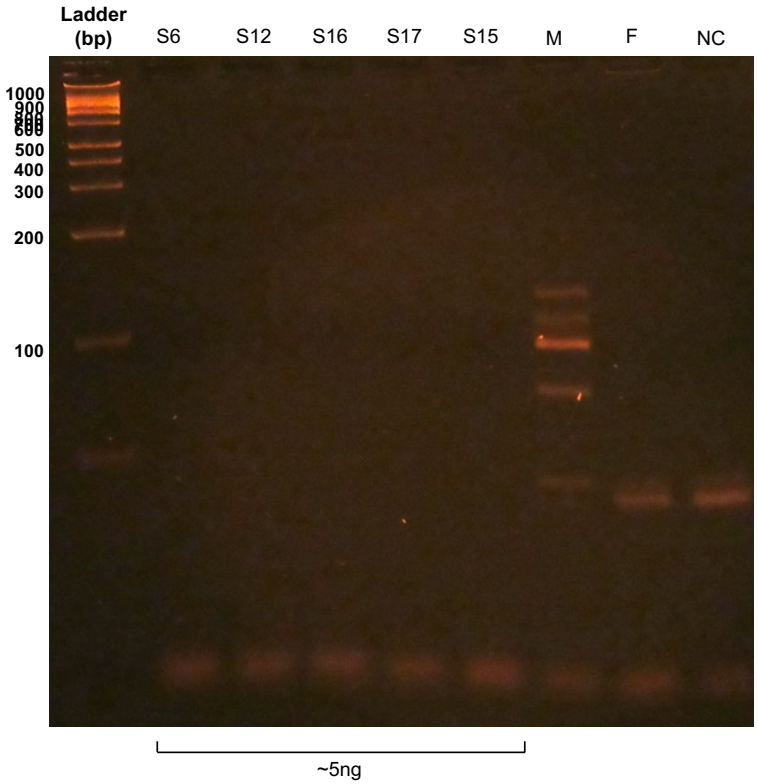


DNA isolation kit	Blood samples (200 µl) week-10 gestational age	Controls	PCR conditions (X55 cycles)
QIAmp DNA Micro	**S6:** ♂ pregnancy **S12:** ♂ pregnancy **S16:** ♂ pregnancy	**M:** ♂ (10 pg) **F:** ♀ (45 ng) **NC:** no-template control	Y chromosome detection in BF and SCM

Figure 32. Agarose gel electrophoresis of the amplification of the DNA isolated from 200 µl of blood from patients at week-10 gestational age using the QIAmp DNA Micro Kit. (A) The input DNA was initially normalised to 45 ng for the samples. **(B)** Input DNA was increased to 100 ng, 200 ng and 300 ng to compare results against 45 ng.

One hundred µl of plasma, separated by the standard clinical blood separation protocol, was then employed directly instead of whole blood, aiming to obtain circulating cfDNA in greater concentrations and to decrease the mother's DNA background originating from blood cells. The concentration of DNA after isolation for samples S6, S12, S16, S17 and S15 (female gender pregnancy) were 0.50 ng/µl, 0.39 ng/µl, 0.35 ng/µl, 0.38 ng/µl and 0.52 ng/µl, respectively. DNA concentrations after isolation from plasma were

expected to be lower compared to DNA isolation from whole blood. Approximately 5 ng of input DNA was used for the multiplex PCR. None of the male gender pregnancies' plasma samples showed amplification (Figure 33).



DNA isolation kit	Plasma samples (100 µl) week-10 gestational age	Controls	PCR conditions (X55 cycles)
QIAmp DNA Micro	**S6:** ♂ pregnancy **S12:** ♂ pregnancy **S16**: ♂ pregnancy **S17:** ♂ pregnancy **S15**: ♀ pregnancy	**M:** ♂ (10 pg) **F:** ♀ (45 ng) **NC:** no-template control	Y chromosome detection in BF and SCM

Figure 33. Agarose gel electrophoresis of the amplification of the DNA isolated from 100 µl of plasma from patients at week-10 gestational age using the QIAmp DNA Micro Kit.

6.2.3.2.2. MagMAX Cell-Free DNA Isolation Kit – Protocol: Isolation of cfDNA from small volumes of plasma and serum samples

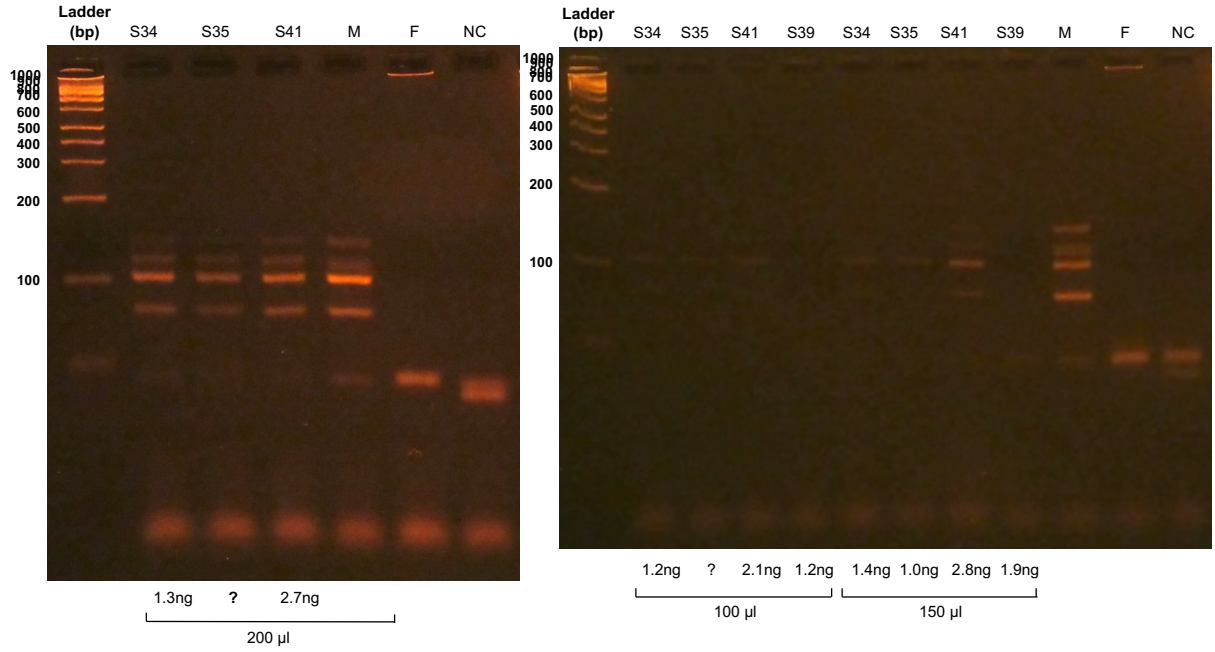
After the no-amplification results from the QIAmp DNA Micro Kit, this study tested the MagMAX Cell-Free DNA Isolation kit, specialised in isolating cfDNA from plasma. The kit was tested with 100 µl, 150 µl and 200 µl of plasma, separated from whole blood by the standard clinical blood separation protocol. The DNA concentrations after isolation are shown in Table 39 and amplification results in Figure 34.

Table 39. DNA concentration after cfDNA isolation with MagMAX Cell-Free DNA Isolation Kit. S39 was not included when cfDNA was isolated from 200 µl. Out-of-range readings had concentrations lower than 100 pg/µl since this is the lower detection limit of the Qubit 1X dsDNA HS kit, used for DNA concentration quantitation. OOR, concentration reading out of range.

Sample	Foetal fraction	200 µl of plasma	150 µl of plasma	100 µl of plasma
		Concentration after cfDNA isolation (pg/µl)		
S34	15%	130	140	120
S35	8%	*OOR*	100	*OOR*
S41	5%	270	280	210
S39	7%	-	190	120

As expected, the DNA concentration decreased as the plasma volume decreased. Sample S35, which produced an out-of-range reading, was still able to show amplification when the cfDNA was extracted from 200 µl. Additionally, there was a significant decrease in placental cfDNA amplification levels when the input plasma volume was decreased to 150 µl and 100 µl. It is difficult to accurately study a correlation between the theoretical concentration for circulating placental cfDNA, calculated with the foetal fraction, and the level of observed amplification of the MSY genes due to the small sample size. Nevertheless, S34 and S41 produced a theoretical concentration for circulating placental cfDNA of 19.5 pg/µl and 13.5 pg/µl, higher than the theoretical concentration (<8 pg/µl)

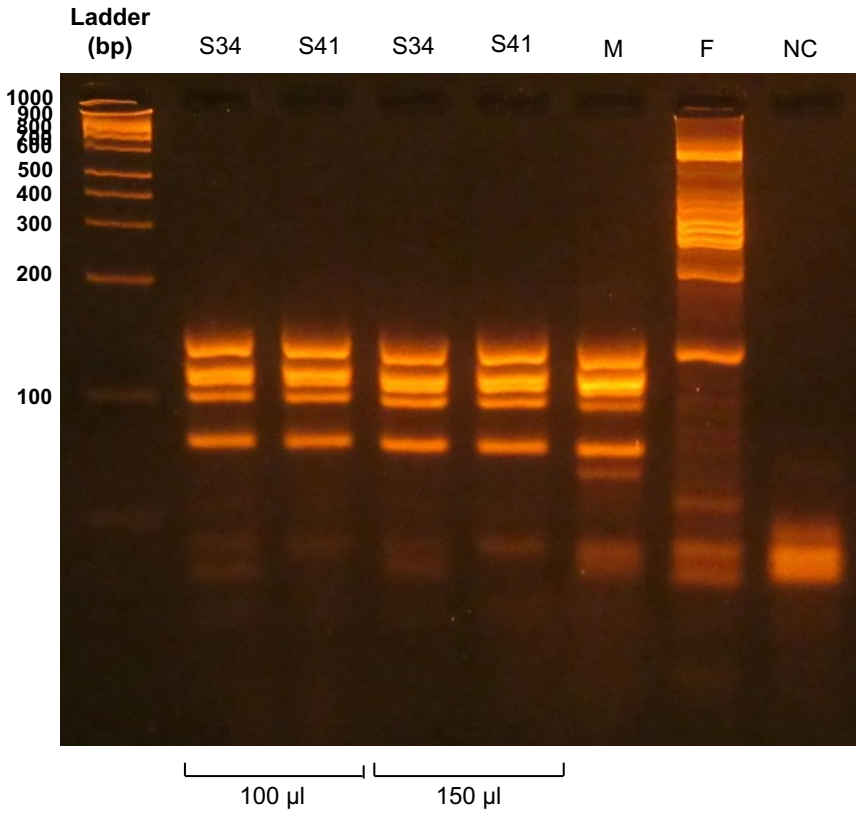
for circulating placental cfDNA of S35 and the observed amplification levels seemed higher in S34 and S41 than S35.



DNA isolation kit	Plasma samples (100 µl, 150 µl and 200 µl) week-10 gestational age	Controls	PCR conditions (X55 cycles)
MagMAX Cell-Free DNA	**S34:** ♂ pregnancy **S35:** ♂ pregnancy **S41**: ♂ pregnancy **S39:** ♀ pregnancy	**M:** ♂ (10 pg) **F:** ♀ (45 ng) **NC:** no-template control	Y chromosome detection in BF and SCM

Figure 34. Agarose gel electrophoresis of the amplification of the DNA isolated from 100 µl, 150 µl and 200 µl of plasma from patients at week-10 gestational age using the MagMAX Cell-Free DNA Isolation Kit and the Y chromosome detection in BF and SCM PCR thermal protocol.

Due to the low levels of amplification shown in the gel electrophoresis when decreasing the DNA input, the polymerase manufacturer's universal thermal program was tested, maintaining X55 cycles. The results in Figure 35 show that the amplification of the MSY targets improved considerably. The amplification bands for the MSY targets in the plasma can be distinguished even with a volume of 100 µl.



DNA isolation kit	Plasma samples (100 and 150 µl) week-10 gestational age	Controls	PCR conditions (X55 cycles)
MagMAX Cell-Free DNA	**S34:** ♂ pregnancy **S41**: ♂ pregnancy	**M:** ♂ (10 pg) **F:** ♀ (45 ng) **NC:** no-template control	Standard for Qiagen Multiplex PCR polymerase

Figure 35. Agarose gel electrophoresis of the amplification of the DNA isolated from 100 and 150 µl of plasma from patients at week-10 gestational age using the MagMAX Cell-Free DNA Isolation Kit and the standard PCR thermal protocol for Qiagen Multiplex PCR polymerase.

6.2.3.2.3. Concordance rate, sensitivity, specificity and foetal fraction

The amplification and concordance rates between the detected sex by routine NIPT and the Y chromosome NIPT assay when isolating cfDNA from 100 µl and 150 µl of plasma using the MagMAX Cell-Free DNA Isolation Kit and the standard Qiagen Multiplex PCR polymerase thermal protocol was 100% (6/6 male and 2/2 female gender pregnancies).

The sensitivity and specificity were 100%.

Figure 36 shows the distribution of the foetal fractions, the mean of which was 8.50%, as determined by the NGS-based NIPT.

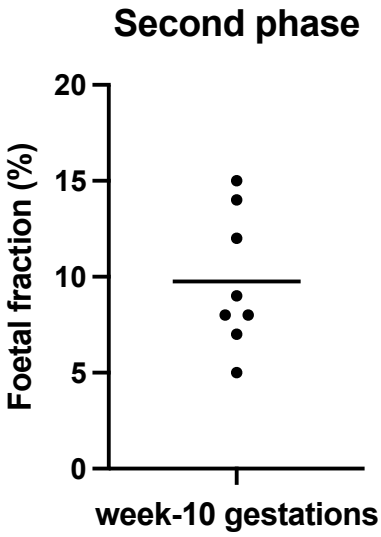
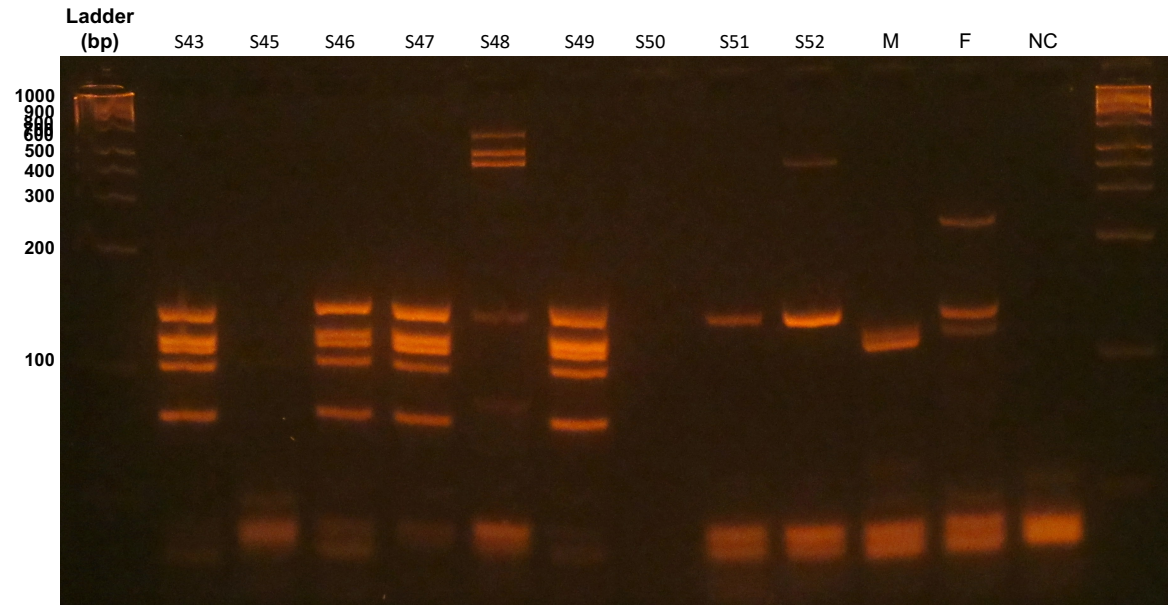


Figure 36. Foetal fraction distribution for eight samples from the second phase of the Y chromosome NIPT study. The horizontal line represents the mean of the data set (8.50%).

6.2.3.3. Third phase: Plasma separation from blood

6.2.3.3.1. Plasma separation from blood

Plasma was separated from 500 µl of aliquoted blood using a standard microcentrifuge (1,600 g x 10 minutes). Fifty to 75% of the blood volume was recovered of plasma volume. Figure 37 depicts an example gel electrophoresis of the Y chromosome assay amplification from the separated plasma. Amplification was successful for all the samples displayed except for S50 (female gender pregnancy), which showed no-amplification and did not display primer-dimer accumulation.



DNA isolation kit	Plasma separated from 500 µl blood samples week-10 gestational age	Controls	PCR conditions (X55 cycles)
MagMAX Cell-Free DNA	**S43:** ♂ pregnancy **S45**: ♀ pregnancy **S46**: ♂ pregnancy **S47**: ♂ pregnancy **S48:** ♀ pregnancy **S49**: ♂ pregnancy **S50:** ♀ pregnancy **S51:** ♀ pregnancy **S52:** ♀ pregnancy	**M:** ♂ (10 pg) **F:** ♀ (45 ng) **NC:** no-template control	Standard for Qiagen Multiplex PCR polymerase

Figure 37. Agarose gel electrophoresis of the amplification of the DNA isolated from plasma separated by the modified protocol from 500 µl of patients at week-10 gestational age using the optimised cfDNA separation and PCR protocols.

6.2.3.3.2. Concordance rate and foetal fraction

The amplification rate for the separated plasma samples with the modified protocol was 96.65% (22/23) and the concordance rates between the sex detected by NIPT and by the Y chromosome NIPT assay was 100% (13/13 male and 9/9 female gender pregnancy).

The sensitivity and specificity were 100%.

Figure 38 shows the distribution of the foetal, where the mean was 8.59%, as determined by the NGS-based NIPT.

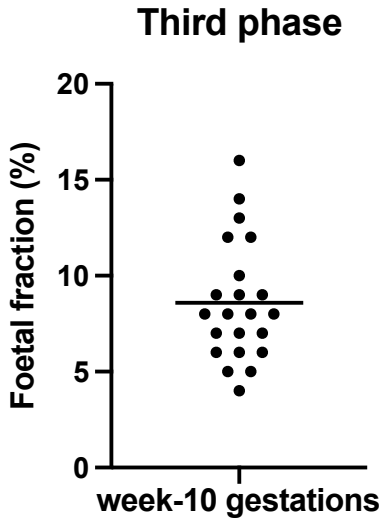


Figure 38. Foetal fraction distribution for 22 samples from the third phase of the Y chromosome NIPT study. The horizontal line represents the mean of the data set (8.59%).

Before moving to discuss the results, this chapter will now turn to introduce the other study on Y chromosome detection in horse BF.

6.3. Project 2. Y chromosome detection in horse BF

This chapter is based on joint work with Ph.D. Nuno Costa-Borges from Embryotools in Barcelona, Spain.

6.3.1. Introduction

6.3.1.1. Equine sex determination

As previously discussed in this thesis, another possibility for non-invasive foetal sex determination lies in the use of amplifiable embryonic cfDNA found in BF. This section of the chapter details a study that explored the possibility of using cfDNA found in BF in the context of the equine industry. There is a strong potential commercial benefit to developing such a method, since there is a strong interest in sex selection for sporting purposes and the current approaches of equine sex determination (transrectal ultrasonography and PGT) have their limitations (Aurich and Schneider, 2014). Although transrectal ultrasonography is non-invasive, it is not a sex selection tool in and of itself. Rather it identifies the sex and the pregnancy can then be terminated, although there are costs to maintaining a pregnant mare and terminating the pregnancy (Aguilar et al., 2012, Crișan et al., 2016). Sex determination before transfer to a recipient mare is a more desirable alternative.

Invasive PCR-based methods for the amplification of equine MSY sequences for embryo sex determination have long been explored. Huhtinen et al. (1997) attempted an equine embryo TE biopsy with needle aspiration, but encountered difficulties when performing the day-6 and day-7 embryo capsule penetration (which is distinctive to the early equine embryo). They achieved pregnancies in 21% of the embryos dissected with a

microblade. However, Choi et al. (2010) later showed increased viability outcomes with day-6 and day-7 equine embryos biopsied with a Piezo drill. Herrera et al. (2014) successfully determined the sex in 50% of biopsied embryos by PCR amplification of the equine *SRY* MSY gene and *AMEL* located in the X chromosome. Concordance of the sex determination between the biopsied cells and ultrasonography was 84.30% for females and 100% for males. Despite the encouraging results, equine embryo TE biopsy requires highly sophisticated equipment and specialised staff, making it difficult to scale. The inability to cryopreserve horse embryos also, until recently, reduced the feasibility of this approach.

6.3.1.2. Equine ARTs

Conventional IVF has been unsuccessful in horses. Although the reasons remain poorly understood, some hypothesise that in vitro capacitation of stallion spermatozoa is inadequate and that the oocyte-sperm binding chemistry is more complex, compared to other species (Leemans et al., 2016). To overcome this, ICSI can be performed. Multiple immature oocytes, which will later be subjected to IVM, can be recovered from live mares as well as a single MII oocyte from a dominant gonadotropin-stimulated follicle. Oocytes can also be recovered from post-mortem ovaries (Hinrichs, 2018). Currently, blastocyst production is successful in 20% to 35% of ICSI-conceived MII oocytes (Sansinena, 2020). The indication for ICSI is generally for mares that cannot produce an embryo by standard reproductive management for embryo transfer programs (Hinrichs, 2018).

Embryo transfer, where an in vivo produced embryo is recovered non-surgically from a donor mare, typically on day-7 or day-8 post-ovulation, and transferred to a

recipient mare, is nowadays broadly employed in the equine industry. Embryo transfer programs allow breeders to obtain more than one foal per donor per year. Superior mares, mares that are in competition or even older mares unable to carry a foal to term can be used as donors (Hinrichs, 2018, Squires, 2020). Most mare embryo donors ovulate only one oocyte per breeding season and the average embryo recovery rate per flush is around 70%. Currently, commercially available equine follicle-stimulating hormone (FSH) for superovulation is unavailable (McCue and Squires, 2015).

6.3.1.2.1. Equine embryo freezing

The introduction of embryo cryopreservation to the equine industry has simplified the reproductive management of mares for donor–recipient synchrony, allowed for the distribution of embryos internationally and for the storage of embryos after TE biopsy, whilst awaiting PGT results (Stout, 2012). However, freezing equine embryos is more complex than freezing human embryos. Slow freezing and vitrification are used on equine morula or early blastocyst (day-5 or day-6 post-ovulation) when the embryo generally has a <300 μm diameter. Although this approach requires that the donor mare is closely monitored with ultrasonography and embryos to be recovered earlier than day-7 or day-8 post-ovulation, when embryo recovery tends to be more successful, the techniques produce a pregnancy rate >55% (Lascombes and Pashen, 2001, Eldridge-Panuska et al., 2005). Hinrichs and Choi (2016) introduced a new equine vitrification approach where larger blastocysts are 'collapsed' prior to freezing by aspirating the BF with micromanipulation. They found that the ongoing pregnancy rates after blastocyst collapse and vitrification were the same as smaller blastocysts that did not undergo collapsing.

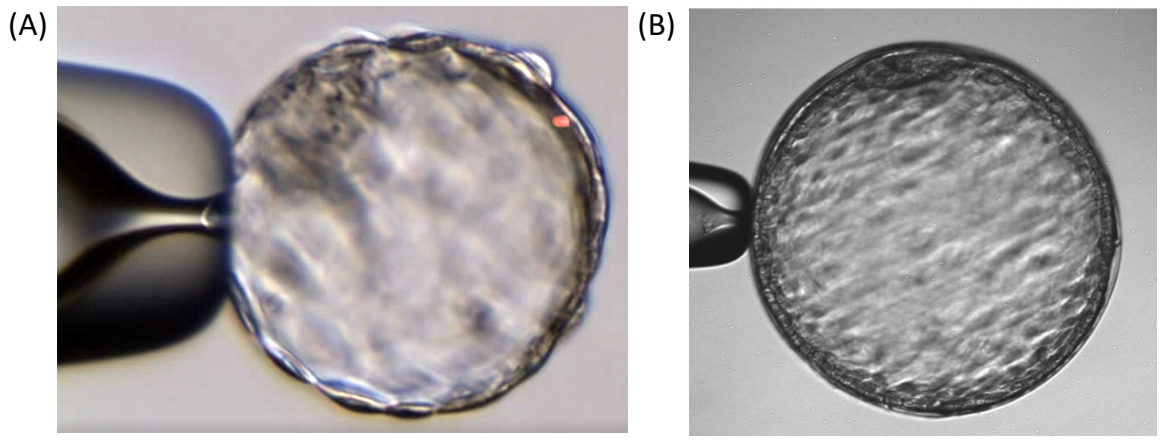


Figure 39. Size comparison between human and equine expanded blastocysts against a micromanipulation pipette. (A) Human day-5 expanded blastocysts present a diameter that ranges between 150 and 200 µm (Gardner and Schoolcraft, 1999). **(B)** Equine day-7 to day-8 expanded blastocysts present a diameter that ranges between 136 and 3,980 µm (Vanderwall and Woods, 2007). The human blastocyst image was obtained from Jakobsson (2018) and the equine blastocyst image from Herrera et al. (2015).

6.3.1.3. Equine BF

The difference in size between human and equine blastocysts is remarkable (Figure 39). Equine early blastocysts (day-6) have 117 to 417 cells and expanded blastocysts (day-7) range between 272 and 2,217 (Colchen et al., 2000). To compare, an expanded human blastocyst (day-6) cell count ranges from 24 to 332 (Hardarson et al., 2003). The three main reported reasons for the low DNA amplification rates observed in human BF are the degraded nature of DNA, the low quantities of template and a small BF volume (Magli et al., 2019). Herrera et al. (2015) successfully amplified DNA in aspirated BF from 83.33% of the in vivo and in vitro produced horse embryos by targeting an X-linked and a Y-linked single copy gene in a duplex PCR. The larger blastocoel cavity volume in horses, enclosed by a significantly higher number of cells, could harbour greater quantities of embryonic cfDNA compared to humans. This could potentially

enhance observed DNA amplifiability and, therefore, Y chromosome PCR-based sex determination.

6.3.1.4. Study aims

The main aim of this chapter is to develop a highly sensitive PCR-based protocol that detects DNA of embryonic origin present in multiple copies in the equine BF, with a potential application for commercial sex determination in embryo transfer programs.

The specific aims are:

1) To design and optimise an equine Y chromosome assay that detects minute quantities of fragmented DNA.
2) To reveal a correlation between the blastocyst diameter and DNA amplifiability.

6.3.2. Materials and methods

6.3.2.1. Sample characteristics and collection

Nineteen blastocysts were collapsed prior to vitrification by Ph.D. Nuno Costa-Borges at Embryotools. The horse blastocyst ID and size are found in Table 40. The embryos were fertilised for commercial purposes after insemination and collected by flushing the mare's uterus at day-6 or day-7 post-ovulation. Gentle suction of 0.3 to 1 µl of BF was subsequently performed using a micromanipulator (Figure 40). The BF was then transferred to a UV-irradiated 0.2 ml tube and shipped to Oxford at -20 °C. The samples were stored at -80 °C until processing.

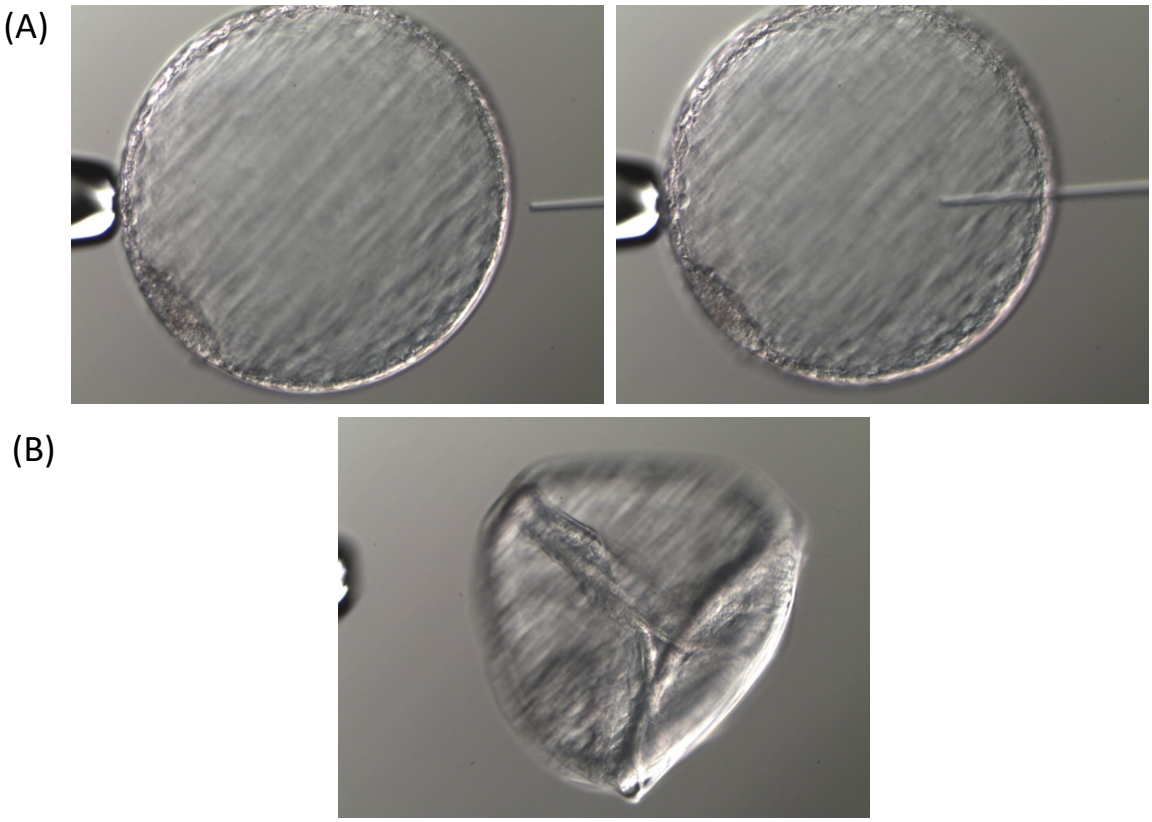


Figure 40. Equine blastocyst collapsing prior vitrification. (A) Micromanipulation is employed to extract the BF with a fine needle. **(B)** The blastocyst cavity collapses after removing the BF. The micromanipulation shown was performed by Ph.D. Nuno Costa-Borges in Embryotools, Spain.

Table 40. Horse blastocyst ID and size. The size of the blastocysts for this study ranged from 187 µm to 2,000 µm in diameter.

Blastocyst ID	Blastocyst size (diameter in µm)	Blastocyst ID	Blastocyst size (diameter in µm)
S1	340	S11	2,000
S2	440	S12	2,000
S3	228	S13	230
S4	187	S14	1,300
S5	237	S15	442
S6	323	S16	700
S7	476	S17	850
S8	230	S18	285
S9	522	S19	1,271
S10	300		

6.3.2.2. Experimental design for the method development

The development of the Y chromosome detection method in horse BF was divided into three phases (Figure 41).

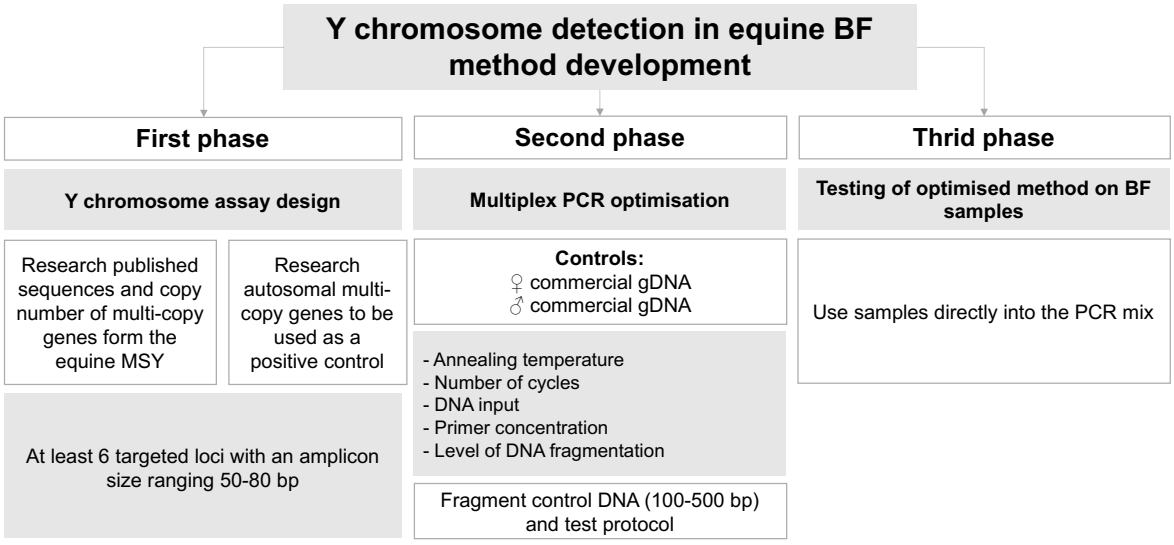


Figure 41. Experimental design for the method development of the Y chromosome detection in equine BF.

First phase: Y chromosome assay was designed.

Second phase: The multiplex PCR conditions, including annealing temperature, number of cycles, DNA input and primer concentration were optimised using commercial equine female and male gDNA. The commercial gDNA was digested to obtain fragments of 100 bp to 500 bp in length to mimic the quality of DNA found in BF.

Third phase: The optimised PCR conditions were tested on the BF samples.

6.3.2.3. Development of the Y chromosome assay

A multiplex PCR assay was designed that targeted loci from various multicopy genes found in the equine MSY. The assay included two positive control candidates that amplified multicopy genes from autosomal chromosomes with the aim of confirming the

presence of detectable nuclear DNA in the BF sample. The amplicons differed in size but kept within the 50 bp to 90 bp in range so that the targets could be distinguished in the agarose gel electrophoresis. The aim was to maintain a similar number of copies among the targeted multicopy genes, including the positive control, to avoid preferential amplification amongst targets. The primer design was carried out following the methods in **Chapter 2**.

6.3.2.4. Multiplex PCR

Unless otherwise indicated, the equine MSY loci were amplified in a multiplex PCR following the Multiplex PCR polymerase manufacturer's universal conditions in **Chapter 2**. Throughout optimisations, different primer pool constitutions were tested and commercial horse ovary and testis gDNA (AMSbio, UK) were used. All the targets in the Y chromosome assay were initially amplified individually using 10 ng of commercial male gDNA in 15 µl reactions, X40 cycles, and an annealing temperature gradient ranging from 50.10 °C to 58.20 °C to find a common temperature. All reactions included a no-template control. Ten ng of commercial female gDNA were then amplified to confirm the specificity of the MSY targets to the male gDNA and the autosomal controls to both the male gDNA and the female gDNA. An annealing temperature gradient was repeated to amplify the targets in a multiplex reaction.

After selecting a common annealing temperature and confirming specificity, the DNA input for the multiplex PCR was decreased to a single-cell level (5 pg to 10 pg) and the number of cycles increased, if necessary.

For the final stage of the multiplex PCR optimisation, the male and female gDNA were fragmented to a size of 100 bp to 500 bp to simulate the quality of the DNA found in BF. To obtain these fragments, gDNA was digested with NEBNext dsDNA Fragmentase for 25 minutes, following **Chapter 2**. The fragments of interest were first purified with a reverse clean-up after mixing them with AMPure XP beads in a 1:0.6 (DNA:beads) dilution factor. The mix was placed in a magnet, the supernatant recovered and then transferred to a new tube and mixed with the magnetic beads in a 1:2.1 ratio for conventional clean-up. The beads were washed away with 80% ethanol. The purified fragments were then eluted in 24 µl of nuclease-free water. The DNA was, subsequently, visualised on the TapeStation using the Genomic DNA and D1000 TapeStation Reagents, following **Chapter 2,** to confirm the size of the recovered fragments. The protocol was tested on the fragmented gDNA.

Finally, the 19 horse BF samples were processed using the optimised protocol. A master mix was prepared with all the components of the multiplex reaction and the corresponding volume was transferred to the 0.2 ml tubes containing the samples and amplified with the optimised PCR conditions.

6.3.2.5. Data analysis and concordance rates

The products of the equine Y chromosome assay amplified in a multiplex PCR were analysed in a 2% (2 g of agarose) or a 4% (4 g of agarose) agarose gel electrophoresis at 100 volts for 1 hour and 20 minutes. Concordance rates between the gender obtained from the equine Y chromosome assay and the known gender at birth were calculated.

The equine Y chromosome assay results were classified as follows:

d) ‘Male’ was defined when there was amplification of the autosomal control and the MSY genes.

e) ‘Female’ was defined when there was amplification of the autosomal control but no amplification of the MSY genes.

f) ‘Female (non-specific amplification)’ was defined when there was no amplification of specific targets (MSY and autosomal genes), but there were non-specific bands in the gel electrophoresis.

g) ‘No-amplification’ was defined when there were no bands in the gel electrophoresis.

The sensitivity and specificity of the test to detect the Y chromosome were calculated using the following equations:

$$Sensitivity\ (\%) = \frac{True\ positives}{True\ positives + False\ negatives}$$

$$Specificity\ (\%) = \frac{True\ negatives}{True\ negatives + False\ positives}$$

6.3.3. Results

6.3.3.1. First phase: Equine Y chromosome assay design

The Y chromosome assay candidates were selected from the equine MSY sequence published on GeneBank (NCBI, USA) by Janečka et al. (2018), who also published a list of the MSY multicopy genes and their copy numbers. The genes with the highest number of copies were selected. According to this list, only five multicopy genes from the equine MSY have a copy number greater than 10. More than one locus was

targeted in the same gene for this reason. Sequence nucleotide mismatches between copies were found in the sequences extracted from GeneBank. These mismatches were avoided from the oligonucleotide sequences limiting further the options for candidates. For the positive control, the *APOBEC3-Z2* multicopy gene located in chromosome 28 was selected. The candidate primers for the Y chromosome assay were subsequently designed and are described in Table 41.

Table 41. Equine Y chromosome assay candidates.

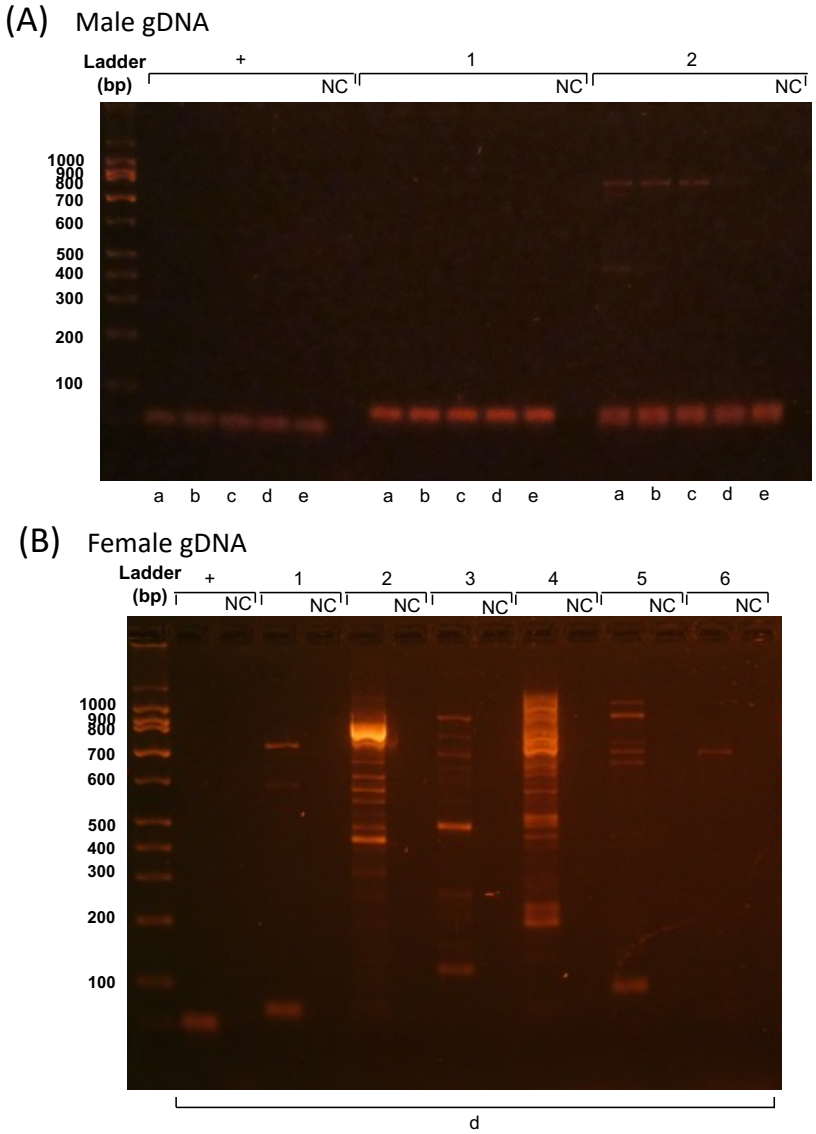
MSY gene	ID	Number of copies	Amplicon size (bp)	Primer sequences
APOBEC3-Z2	**+**	5	46	F: 5' GGCATCCCAAACTCAAGTCG 3' R: 5' CGCATTTGGACAAGGACTGG 3'
APOBEC3-Z2	**2+**	5	90	F: 5' CCCAAACTCAAGTCGTTATCTGA 3' R: 5' GCATTTGGACAAGGACTGGA 3'
ETY1	**1**	22	63	F: 5' TCCCTTTCTGCCGTCTCTTT 3' R: 5' GGCAAAGACCCAGCTAAACC 3'
ETY1	**2**	22	67	F: 5' TTTCAAGCCGGCAAAAGCAC 3' R: 5' CGGGCAGGCTGATGTGTA 3'
ETY4	**3**	17	65	F: 5' AGAAAGGGATGTGGCGAGTG 3' R: 5' GGCAGAGGATGAAGCTTTGT 3'
ETY4	**4**	17	68	F: 5' CAGAGTTCCCCTGCAGCC 3' R: 5' CCTGGGAGCATGTCTGTATCAT 3'
YIR2	**5**	13	74	F: 5' AACGCTTAACTATGCCTGGC 3' R: 5' TCTATGTCTTATTGCACCAACGT 3'
YIR2	**6**	13	69	F: 5' AGAGTGAGTACGTTGACCCA 3' R: 5' TGTAGCAGTTTCATGTTGAGCA 3'

6.3.3.2. Second phase: Multiplex PCR optimisation

6.3.3.2.1. Annealing temperature gradient for the singleplex PCR

All the primer pairs were initially amplified using the commercial male gDNA in a singleplex with an annealing temperature gradient to find a common temperature and confirm specificity. The singleplex reactions were then carried out on the commercial

female gDNA. Figure 42 shows two representative examples of the agarose gels. All the MSY primers were specific for the male gDNA with the expected sizes. The autosomal control amplified for male and female gDNA with the expected size.

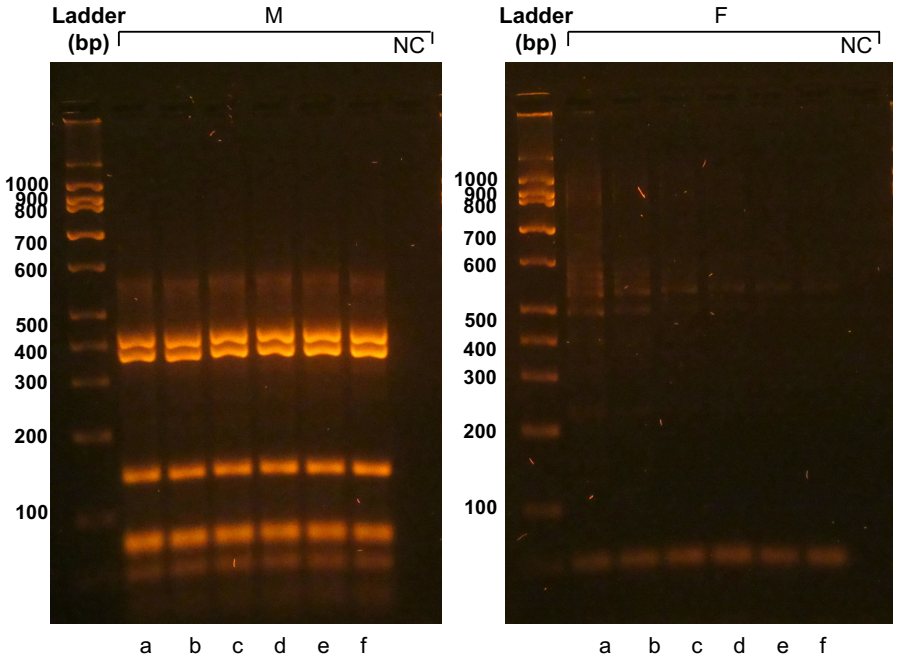


Primers shown in this figure	Controls	Annealing temperatures
+, 1, 2, 3, 4, 5 and 6	**NC:** no-template control	**a:** 50.1 °C **d:** 56.9 °C **b:** 51.7 °C **e:** 58.2 °C **c:** 54.4 °C

Figure 42. Agarose gel electrophoresis of the individual amplification of primer pairs for the equine Y chromosome assay. (A) The primers targeting the MSY loci showed specificity for the male gDNA. **(B)** The primers targeting the MSY loci did not show specificity for the female gDNA, except for primer 1. Nevertheless, this band did not re-appear in further experiments with female control gDNA. Therefore, it was concluded this band was non-specific amplification. Not all the results obtained from the singleplex reactions are included in this figure.

6.3.3.2.2. Annealing temperature gradient for the multiplex PCR

Various primer pools with different combinations of primers were used throughout this study for optimisation purposes. To avoid the non-specific amplification observed in the previous experiment, the second annealing temperature gradient used for the first multiplex used higher temperatures than the gradient in the singleplex reactions (Figure 43). Although the primers showed MSY specificity for the male gDNA, products with greater sizes than expected were produced. This was due to the proximity of the targeted loci, as some of the primer pairs were designed to target the same multicopy gene. Consistent with the singleplex reactions, higher annealing temperatures produced less non-specific amplification with the female gDNA. The annealing temperature 62.9 °C was used for the following multiplex PCR experiments.



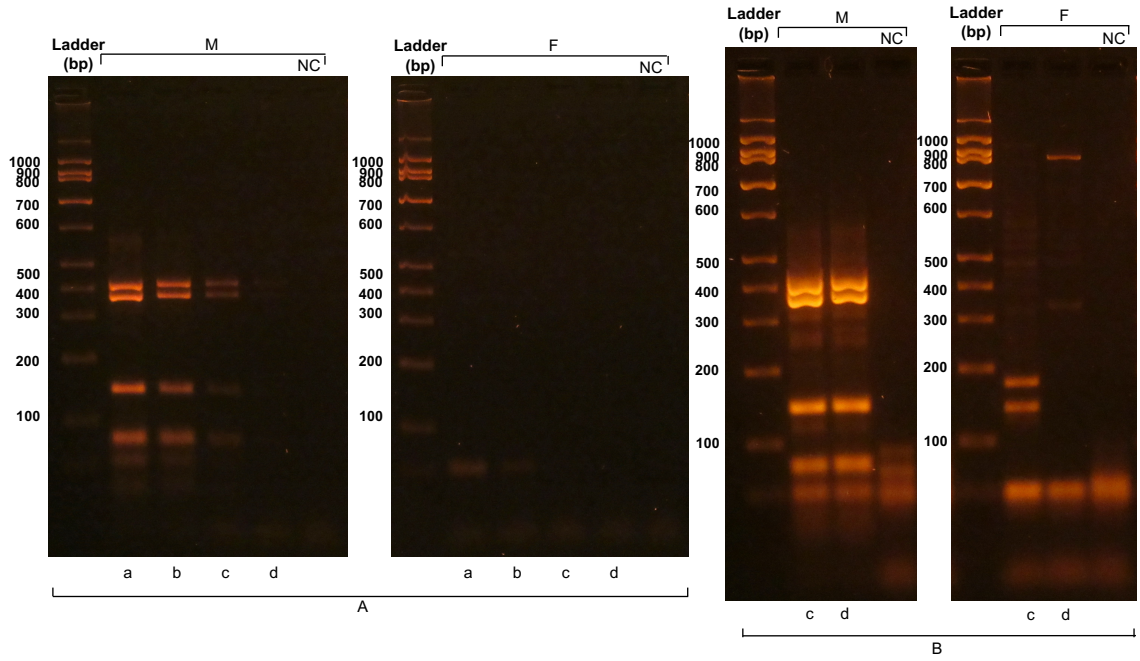
Primer pool 1	Controls	Annealing temperatures
+, 1, 2, 3, 4, 5 and 6	**M:** ♂ (10 ng) **F**: ♀ (10 ng) **NC:** no-template control	**a:** 58.2 °C. **d:** 61.2 °C **b:** 59.4 °C **e:** 62 °C **c:** 60.3 °C **f:** 62.9 °C

Figure 43. Agarose gel electrophoresis of the multiplex PCR annealing temperature gradient with primer pool 1.

6.3.3.2.3. Primer pool constitution, DNA input, number of cycles and primer concentration

Once the Y chromosome assay proof of principle was confirmed on male and female gDNA, the input gDNA was decreased to a single-cell level and the number of cycles adjusted.

Primer pool 1 was used first (Figure 44). The increase in cycles from X30 to X45 mitigated the lack of amplification of 100 pg and 10 pg of DNA. A band was produced in the no-template control with the same size as the autosomal control amplicon, which was likely the accumulation of primer-dimers due to the increased number of cycles.

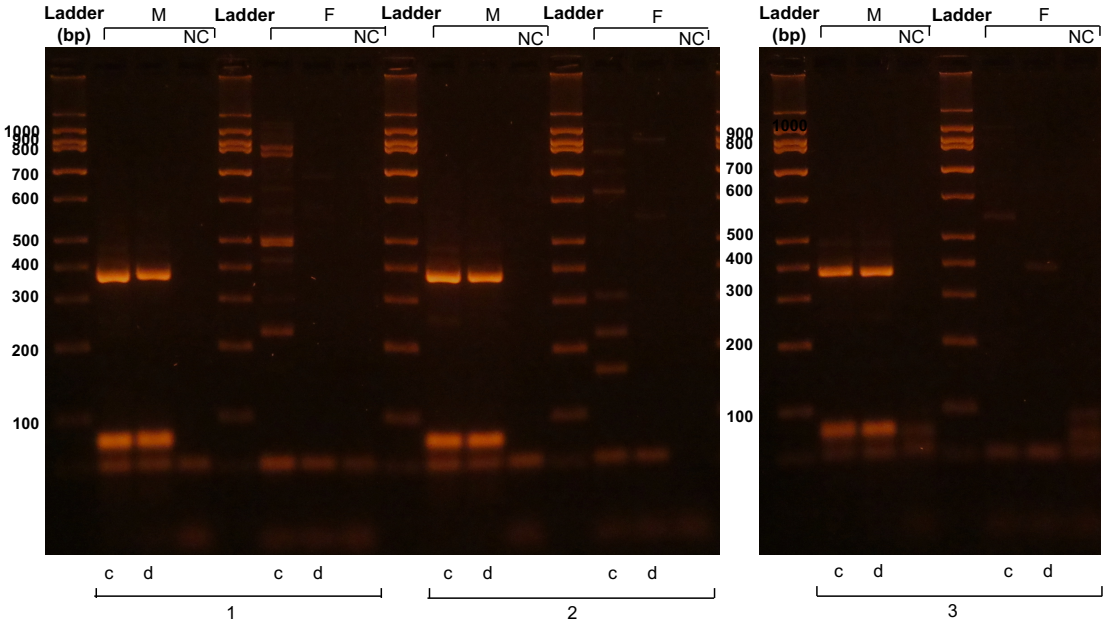


Primer pool 1	Controls	DNA input	Number of cycles
+, 1, 2, 3, 4, 5 and 6	**M:** ♂ (10 ng) **F**: ♀ (10 ng) **NC:** no-template control	**a:** 10 ng **d:** 10 pg **b:** 1 ng **c:** 100 pg	**A:** X30 **B:** X45

Figure 44. Agarose gel electrophoresis of the multiplex PCR with various DNA inputs, number of cycles and primer pool 1.

The number of cycles was then adjusted to X40 and primer pool 2 was used after removing primer pair 1. Lower primer concentrations were also tested, aiming to avoid

the accumulation of primer-dimers (Figure 45). Similar amplification levels of the targets were obtained with different primer concentrations, the lowest of which, 0.1 µM, did not avoid the accumulation of primer-dimers. It was, therefore, impossible to distinguish between a female DNA sample and the no-template control. Using primer concentrations of 0.15 µM and 0.2 µM increased the non-specific amplification for the female gDNA.

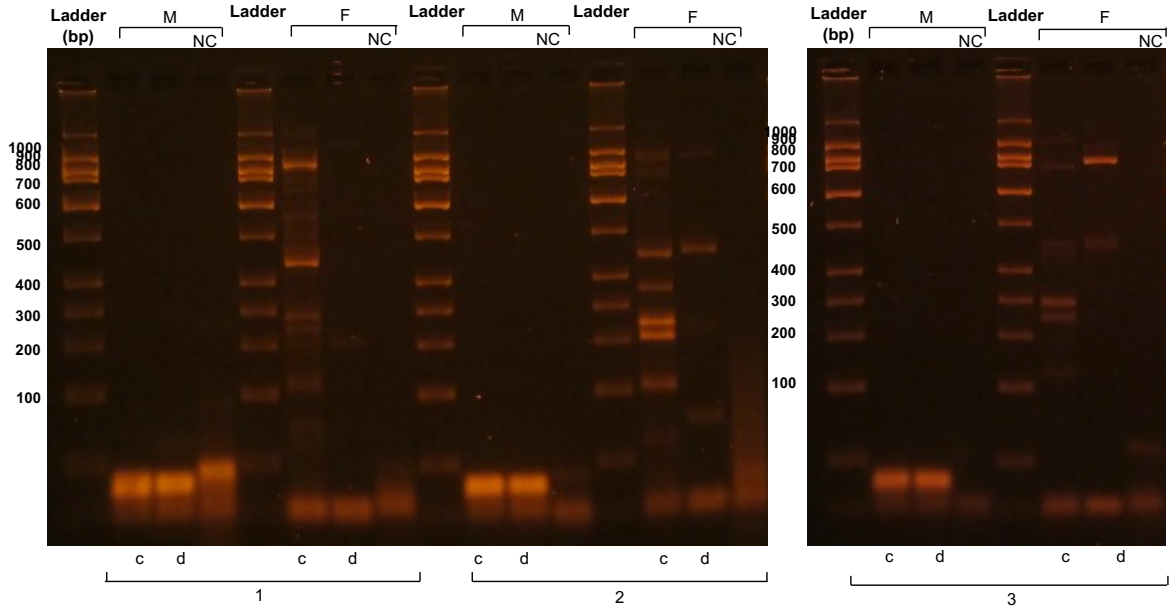


Primer pool 2	Controls	DNA input	Number of cycles	Primer concentration
+, 2, 3, 4 and 6	**M:** ♂ (10 ng) **F**: ♀ (10 ng) **NC:** no-template control	**c:** 100 pg **d:** 10 pg	X40	**1:** 0.2 µM **2:** 0.15 µM **3:** 0.1 µM

Figure 45. Agarose gel electrophoresis of the multiplex PCR with various DNA inputs, primer concentration and primer pool 2.

Primer pool 3 was tested leaving only the autosomal control and primer pairs 2, 4 and 6. Different primer concentrations were used. The cycles increased to X50, aiming to increase the target band intensity (Figure 46). The band intensity for the MSY targets increased when compared to X40 cycles. The brightest Y chromosome loci band was produced when the primer concentration was 0.2 µM. Increasing the number of cycles

increased the non-specific amplification for the female gDNA with all primer concentrations, despite this pool comprising only four primer pairs. A band with the same size as the autosomal control was produced for the no-template control.

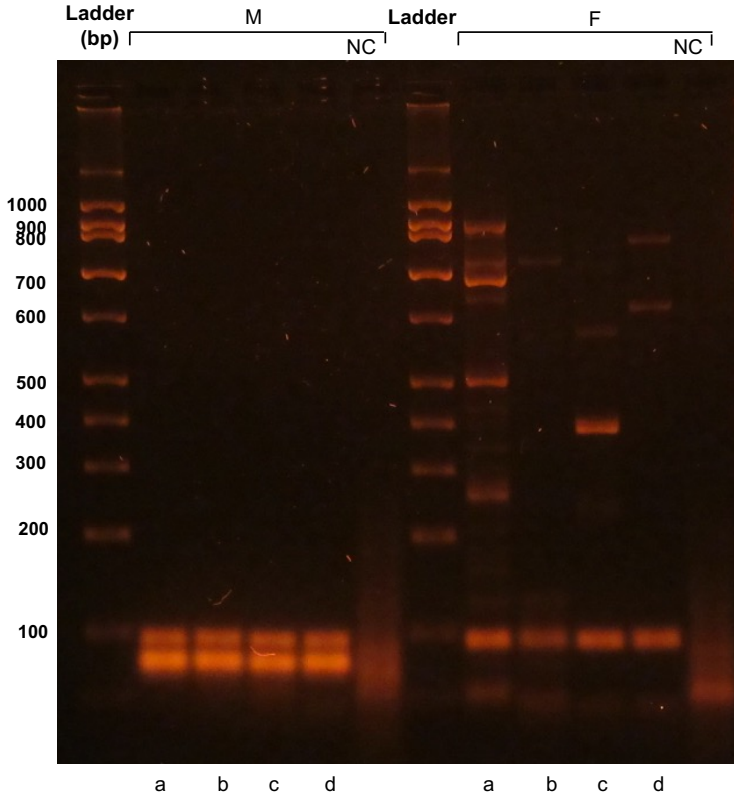


Primer pool 3	Controls	DNA input	Number of cycles	Primer concentration
+, 2, 4, and 6	**M:** ♂ (10 ng) **F**: ♀ (10 ng) **NC:** no-template control	**c:** 100 pg **d:** 10 pg	X50	**1:** 0.2 µM **2:** 0.15 µM **3:** 0.1 µM

Figure 46. Agarose gel electrophoresis of the multiplex PCR with various DNA inputs, primer concentration and primer pool 3.

After the persistent primer-dimer accumulation, a final pool was tested, keeping primers 2, 4 and 6 but replacing the autosomal control '+' with '2+'. The autosomal control 2+ target is 90 bp in size. The results of the multiplex PCR amplification using primer pool 4, 0.2 µM of primer concentration and X50 cycles are shown in Figure 47. The input DNA was also decreased to sub-single cell level when testing primer pool 4. Despite the smear on the no-template control, this new primer pool constitution made it possible to distinguish between the primer-dimer accumulation band on the negative control and

the amplification of the autosomal control for the female gDNA. The male gDNA amplified the autosomal control and the MSY targets, which are shown as a thick brighter band below the autosomal control amplification band.



Primer pool 4	Controls	DNA input	Number of cycles	Primer concentration
2+, 2, 4, and 6	**M:** ♂ (10 ng) **F**: ♀ (10 ng) **NC:** no-template control	**a:** 100 pg **b:** 15 pg **c:** 100 pg **d:** 5 pg	X50	0.2 µM

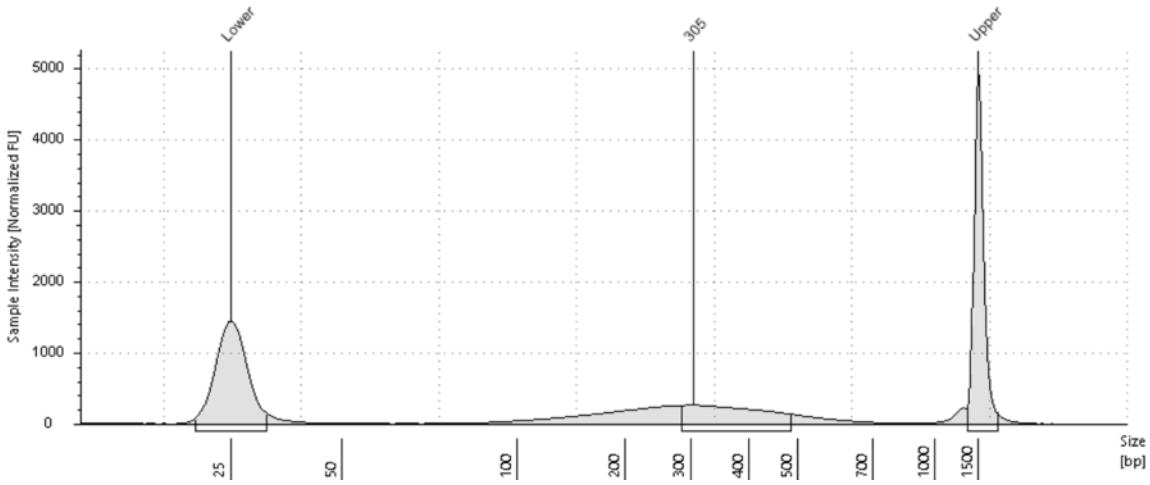
Figure 47. Agarose gel electrophoresis of the multiplex PCR with various DNA inputs and primer pool 4.

6.3.3.2.4. DNA fragmentation

The following step for the protocol optimisation was to fragment the commercial female and male gDNA into sizes ranging between 100 bp and 500 bp. The products were visualised on the TapeStation after treating the DNA with the NEBNext dsDNA

Fragmentase. The fragments obtained after purification ranged from 75 bp to 700 bp peaking in concentration at around 300 bp (Figure 48). The concentration of the recovered male and female DNA fragments was 7.03 ng/µM and 3.12 ng/µM, respectively.

(A) Male gDNA



(B) Female gDNA

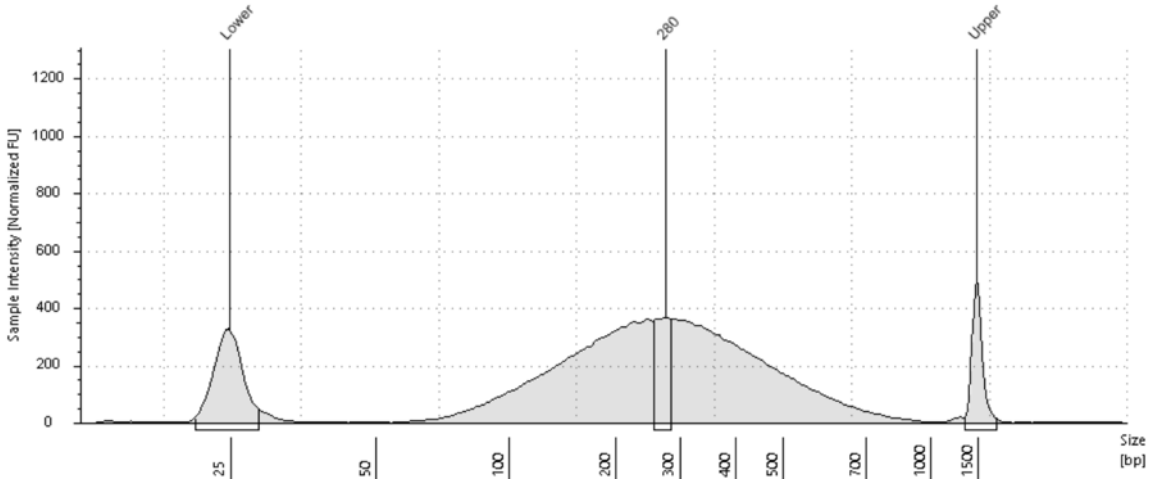
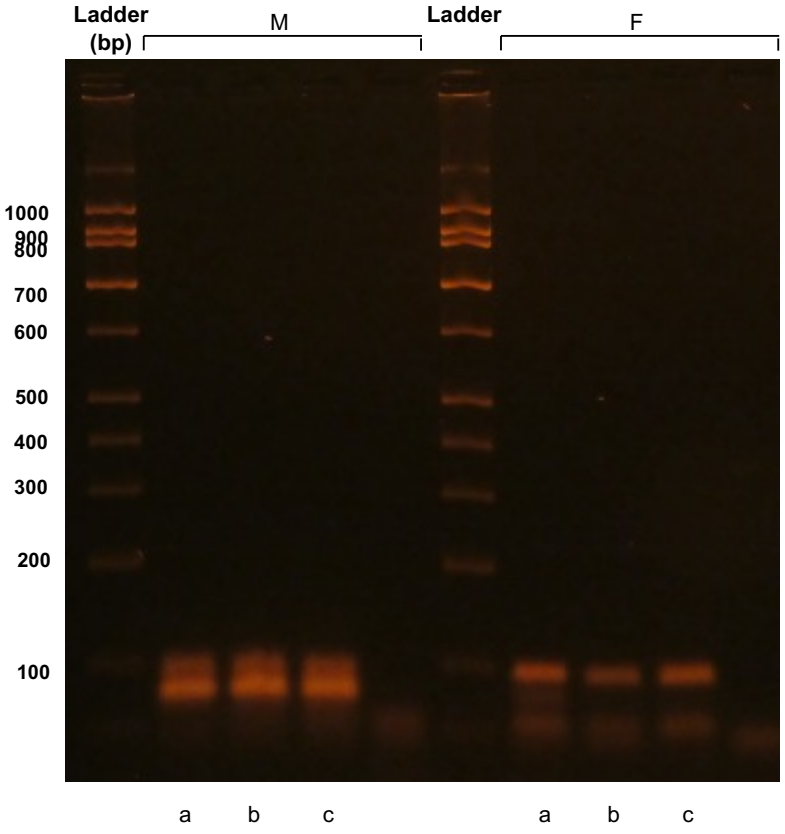


Figure 48. Products obtained after fragmentation of equine commercial female and male gDNA visualised with the TapeStation. Fragment size concentration peaked at 305 bp for the male DNA and 280 bp for the female DNA.

An amplification using primer pool 4 and the fragmented DNA was carried out after confirming the fragments had the desired size range. The results and the PCR

conditions are shown in Figure 49. These reactions included sub-single cell level amounts of DNA input. There was an absence of non-specific amplification, contrary to previous multiplex reactions, due to the reduction of fragment size in the DNA input. These results also clearly distinguished the no-template control and the amplification of female DNA since only the band of primer-dimer accumulation is shown for the negative control, whilst the autosomal control amplified the female DNA. Specific bands amplified for the male DNA. After obtaining these results, the PCR configuration used for this last experiment was determined to be the final optimised multiplex PCR conditions.

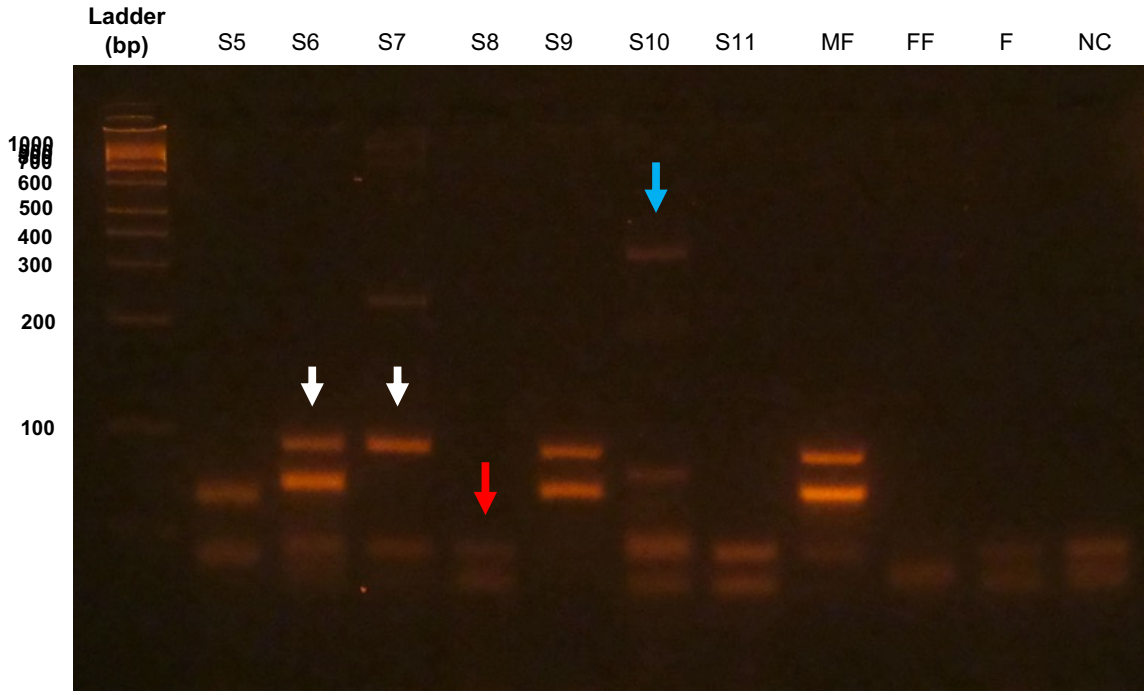


Primer pool 4	Controls	DNA input	Number of cycles	Primer concentration
2+, 2, 4, and 6	**M:** ♂ (fragmented) **F**: ♀ (fragmented) **NC:** no-template control	**a:** 15 pg **b:** 10 pg **c:** 5 pg	X50	0.2 µM

Figure 49. Agarose gel electrophoresis of the multiplex PCR with various inputs of fragmented DNA and primer pool 4.

6.3.3.3. Third phase: Testing of optimised method on horse BF samples

This study tested the optimised protocol for equine Y chromosome detection in BF on the 19 samples obtained from Embryotools. The summary of the results generated and the band amplification description is shown in Table 42. The experiment found three samples from female blastocysts (15.79%), five from male blastocysts (26.31%) and seven samples that did not produce any amplification (36.84%). Four samples were identified as female due to the presence of non-specific amplification (21.05%), but these samples failed to amplify the autosomal control. A representative example of the gel electrophoresis presented in Figure 50. As it is observed, in this gel none of the female gDNA controls, fragmented and unfragmented, amplified the autosomal control. Three out of the five total multiplex PCR runs that were performed to analyse the 19 BF samples failed to amplify the autosomal control for the female DNA controls, likely due to degradation of the commercial gDNA.



Primer pool 4	Controls	Number of cycles	Primer concentration
2+, 2, 4, and 6	**MF:** ♂ (10 pg, fragmented) **FF**: ♀ (10 pg, fragmented) **F**: ♀ **NC:** no-template control	X50	0.2 µM

Figure 50. Representative example of an agarose gel electrophoresis of the results obtained from testing the optimised Y chromosome detection in BF in samples. One of the white arrows indicates a sample identified as male due to the presence of the autosomal control band and the MSY loci band. The other white arrow indicates a sample identified as female due to the presence of the autosomal control band only. The red arrow indicates a sample with no amplification and the blue arrow indicates a sample identified as female due to non-specific amplification despite the absence of the autosomal control band. Multiplex PCR products were run in a 4% agarose gel electrophoresis at 100 volts for 1 hour and 20 minutes.

Table 42. Band amplification description for each BF sample. 'M' stands for 'male', 'F' for 'female', 'NSA' for 'non-specific amplification', 'NA' for 'no amplification', 'INA' for 'information not available', 'NT' for 'embryo not transferred' and 'NP' for 'no pregnancy stablished'.

ID	Blastocyst diameter (µm)	MSY genes amplification	Autosomal control amplification	Non-specific amplification	Y chromosome assay result	Confirmed sex at birth
S1	340	Yes	Yes	No	**M**	**M**
S2	440	No	Yes	Yes	**F**	**F**
S3	228	No	Yes	No	**F**	**F**
S4	187	No	No	No	**NA**	*INA*
S5	237	No	No	Yes	**F (NSA)**	*INA*
S6	323	Yes	Yes	No	**M**	**M**
S7	476	No	Yes	Yes	**F**	**F**
S8	230	No	No	No	**NA**	*NT*
S9	522	Yes	Yes	No	**M**	*NT*
S10	300	No	No	Yes	**F (NSA)**	*NT*
S11	2,000	No	No	No	**NA**	*INA*
S12	2,000	No	No	No	**NA**	*INA*
S13	230	Yes	Yes	No	**M**	*NT*
S14	1,300	No	No	No	**NA**	*INA*
S15	442	No	No	No	**NA**	*INA*
S16	700	No	No	No	**NA**	*INA*
S17	850	No	No	Yes	**F (NSA)**	**F**
S18	285	No	No	Yes	**F (NSA)**	**M**
S19	1,271	Yes	Yes	No	**M**	*NP*

Three out of the six no-amplification results came from BF from blastocysts with a diameter above 1,000 µm. Only one BF sample, of the total four samples from blastocysts with a diameter above 1,000 µm, produced a male result.

Eight out of the 19 embryos analysed in this study were transferred to recipient mares during the breeding season of 2022, see Table 42. Of the eight embryo transfers, one failed to achieve a pregnancy.

The sex was confirmed from the seven embryos that reached birth and the concordance rate with the Y chromosome assay was 85.71% (6/7) with a test sensitivity of 66.67% and a specificity of 100%.

When considering only specific amplification, the concordance is 100% (5/5) with a test sensitivity and specificity of 100%.

For seven embryos, at the time of writing it was not known if a pregnancy had been achieved.

6.3.4. Discussion

Non-invasive methods for PGT are challenging due to the low abundance, and fragmented nature, of cfDNA. This chapter explores two different PCR-based approaches for non-invasive sex determination that target multicopy genes of the MSY. One, in placental circulating cfDNA and the other in horse BF.

6.3.4.1. Development of a PCR-based method for the detection of uniquely foetal DNA in maternal blood during early pregnancy

Placental circulating cfDNA found in the maternal bloodstream during pregnancy originates from trophoblast apoptotic and necrotic cells (Tjoa et al., 2006). For this reason, the DNA molecules of foetal origin are of fragmented nature with a peaking length around 166 bp (Lo et al., 2010). In terms of quantities, NGS studies show that the total circulating cfDNA foetal fraction during the first and second trimesters ranges from under 3% to over 30%, with increasing absolute quantities as gestation advances (Lo et al., 1998, Canick et al., 2013). Thus, it is imperative to employ a highly sensitive assay for the analysis of circulating cfDNA of foetal origin. A typical routine NIPT workflow consists of maternal blood collection, plasma separation, cfDNA isolation and cfDNA analysis.

6.3.4.1.1. CfDNA isolation

This study first compared the cfDNA yield between the QIAmp DNA Micro and the MagMax Cell-Free DNA kits. Both use spin column-based methodologies for the isolation of DNA. The feasibility of obtaining informative foetal cfDNA with a widely used spin column methodology for the isolation of gDNA from blood (the QIAmp kit) was assessed. Unsurprisingly, placental cfDNA isolated from plasma using the MagMax kit showed better amplification for the PCR-based Y chromosome NIPT. Spin column-based DNA isolation technologies use silica membranes to which nucleic acids bind to later be detached with low salt concentration buffers (Melzak et al., 1996). However, these commercial kits have focused on retrieving high-molecular-weight DNA. Kirsch et al. (2008) compared the yield of highly fragmented DNA between the QIAmp DNA Blood Mini kit and a the NucleoSpin Plasma XS (Macherey-Nagel, Switzerland), a kit developed for the isolation of cfDNA. While the NucleoSpin kit successfully recovered fragment sizes as small as 50 bp long, the QIAmp kit recovered only a fraction of the 100 bp, 150 bp and 250 bp fragments recovered by the NucleoSpin kit and failed to recover 50 bp long fragments.

The concentrations of total DNA isolated from 100 µl, 150 µl and 200 µl of plasma with the MagMAX kit in this study ranged from readings that were out of the reading range (<100 pg/µl) to 280 pg/µl. As expected, higher concentrations of DNA were obtained when the volume of plasma was increased. In fact, amplification of the MSY genes was detected in all 200 µl plasma sample inputs. However, data from Jorgez et al. (2006), where the performance of various cfDNA isolation methods was analysed, showed that methods that produced the highest concentration of total isolated DNA (boiling-lysis approach) did not necessarily produce the highest rate of detection of male

gender pregnancies from maternal blood, magnetic-beads approach did. This is likely due to the difference in fragment size between placental cfDNA (<200 bp) and normally circulating cfDNA (>400 bp) (Lo et al., 2010, Mouliere et al., 2011) and the difference in fragment size yield recovery between methodologies. Therefore, evidence suggests that the total concentration of isolated DNA from plasma does not indicate foetal Y chromosome detection performance. High foetal fraction variability between samples may also be at play. In fact, results in this study showed a positive correlation between the theoretical concentration of isolated circulating placental cfDNA, calculated with the total isolated DNA concentration and the known foetal fraction, and observed levels of Y chromosome amplification of male gender pregnancy samples. This correlation is expected, but statistical significance could not be proven due to the small sample size.

Although spin column-based protocols are widely used, these are complex procedures that involve a high degree of manual handling, increasing the risk of cross-contamination (Tan and Yiap, 2009). Studies have shown the superior performance of magnetic bead-based and microfluidics methodologies in the isolation of circulating foetal cfDNA (Jorgez et al., 2006, Diefenbach et al., 2018, Lee et al., 2020).

6.3.4.1.2. PCR conditions

The main difference between the multiplex polymerase manufacturer's universal thermal protocol and the Y chromosome assay for BF and SCM is the reduced times for each PCR step in the latter, aiming to reduce by almost one-third the overall PCR time (~72 vs. ~218 minutes). The manufacturer's thermal protocol achieved high levels of foetal MSY amplification even with plasma volumes of 100 µl. Nevertheless, some female gender pregnancy plasma samples and female gDNA controls showed non-specific

amplification, likely due to the longer extension time (45 seconds vs. 90 seconds). Future optimisation experiments could test a range of extension times.

6.3.4.1.3. Plasma separation

A crucial step for NIPT is the separation of plasma from blood. The routine clinical method for 10 ml of blood employs a refrigerated centrifuge and a program with a speed of 1,600 x g for 10 minutes at a temperature of 4°C that allows a no-break deacceleration at the end of the centrifugation. The protocol avoids the disruption of the buffy coat, the fraction between the plasma and the erythrocyte fractions containing leukocytes and platelets, to prevent the aspiration of cellular components of maternal origin (Luo et al., 2022). Maternal blood cell carryover and mechanical injury of these cells (haemolysis) during sample collection and storage can lead to a reduction of the foetal fraction impairing the sensitivity of the NIPT for aneuploidies (Wald et al., 2018, Nishimura et al., 2019).

This study tested a protocol where a standard laboratory microcentrifuge was used to separate plasma from up to 2 ml of blood without no-break configuration or refrigeration. Although the likelihood of maternal carryover increased with the modified plasma separation protocol, the Y chromosome assay amplification rate was still 96.65% (22/23) and the concordance rates with sex determination by routine NIPT was 100%.

These results demonstrate a high sensitivity of the optimised assay application for detecting the foetal Y chromosome in maternal blood at week-10 gestations and the suitability of the modified plasma separation protocol. Considering the observed variation in foetal fraction between samples (4% to 16%) and that, on average, 55% of

the blood volume is plasma, the data indicates that it is feasible to detect foetal Y chromosome with a PCR-based approach from blood volumes as low as 0.5 ml, a remarkable reduction of the standard 10 ml of venous blood volume needed for routine NIPT. Although the present methodology successfully demonstrated the proof-of-principle, the results produced are not a definitive indication of the accuracy of the present Y chromosome NIPT protocol and further validation studies are needed.

6.3.4.1.4. Detection of foetal Y chromosome in circulating cfDNA and gestational age

Studies have analysed the presence of Y chromosome sequences of foetal origin in the maternal blood at different stages of gestation with a consistent conclusion that the circulating placental DNA increases as pregnancy progresses. Lo et al. (1998) reported a mean foetal fraction in early pregnancy (up to week-12 of gestation), measured by targeting a 137 bp sequence of the *SRY* gene with RT-qPCR, of 3.40% ranging from 0.39% to 11.90%. The DNA was isolated from plasma with the QIAmp Kit. Ariga et al. (2001) targeted a 73 bp sequence of the *SRY* gene with RT-qPCR and found a transient increase in foetal DNA between week-7 and week-16 of gestation in eight of the 11 male gender pregnancies studied. Plasma DNA was also extracted with the QIAmp Kit. Rijnders et al. (2003) evaluated the earliest gestational age at which *SRY* can be detected in the mother's blood of male gender pregnancies. A 137 bp long sequence was targeted with RT-qPCR and plasma DNA was isolated with the Ultrasens Virus (Qiagen, Germany) and QIAmp kits. The *SRY* gene was detected in 50% of week-5, 80% of week-8, and 100% of week-9 gestations. Recent studies using the QIAmp kits to isolate DNA from plasma and to target a single sequence from *SRY* have obtained similar results (Qiu and Liu, 2019).

Mahdavi et al. (2019) co-amplified the *SRY* and *ZFY*, reporting a gender determination assay sensitivity and specificity of 100% from week-10 to week-12 gestations.

This study optimised a foetal Y chromosome detection protocol that achieved high amplification levels of the six multicopy MSY short-length sequences in male gender gestations. The results show that higher placental circulating cfDNA yields are obtained when cfDNA isolation commercial kits are used. Foetal fractions of the week-10 gestations included ranged from 4% to 16%, according to the routine NGS-based NIPT. The above, along with the often-successful reports of foetal Y chromosome amplification from gestational ages as early as week-5, suggest that the present Y chromosome assay could amplify Y chromosome multi-copy sequences from gestational ages earlier than week-10.

Future experiments would include venous blood from week-5 to week-9 gestations and an increased sample size to validate the gestational age and foetal fraction threshold of male gender pregnancy determination of the present Y chromosome NIPT methodology. The high detection success in venous blood volumes of as little as 400 µl opens the way to test the present assay in capillary blood samples obtained from finger-prick collection. Gyanchandani et al. (2018) carried out a pilot study where finger-prick blood samples were collected onto a paper-based microfluidic device for centrifugation-free plasma separation. CfDNA was separated and the authors reported a 100% concordance rate with the tumour mutations detected with routine venous blood cfDNA analysis. This same approach can be tested with placental circulating cfDNA.

6.3.4.1.5. Ethical considerations

This study provides promising evidence that low-cost prenatal sex determination can be used very early into pregnancy. However, the increased availability of NIPT, with reduced oversight from clinicians, creates potential risks that the technology could be used for sex-selective termination of pregnancy for social reasons. There is indeed a historical precedent for how new technologies can be used for non-medical sex selection. After the introduction of the one child policy in China in 1980, ultrasounds were commonly used for sex determination and contributed significantly to the disproportionate increase in male births compared to female births (Chen et al., 2013). More demographic data from the late 2000s is needed to study if the introduction of NIPT has influenced a change in the sex ratio at birth in China or elsewhere (Bowman-Smart et al., 2020).

Responding to these ethical concerns, in 2017, the UK´s Nuffield Council on Bioethics attempted to limit the scope of NIPT use domestically. It recommended that the National Health Service (NHS), and private providers, only offer NIPT to patients who seek to determine if their foetus has a serious medical condition or impairment (Nuffield Council on Bioethics, 2017). Ultimately, NIPT is an emerging tool, soon to become widespread, and as discussed by Bowman-Smart et al. (2021) and Zhang et al. (2023), NITP stakeholders should engage in the global governance of the technology to ensure that it is consistently applied ethically.

6.3.4.2. Development of a PCR-based method for the detection of embryonic DNA in horse BF

Molecular analyses in human BF report a majority population of DNA fragments sized 160 bp – 220 bp (peaking at 169 bp) and a smaller population of fragments sized 300 bp – 400 bp (Zhang et al., 2016) making this DNA population very similar in size to the placental circulating cfDNA. This aligns with the hypothesis that the DNA found in the BF, or extraembryonic DNA, originates from apoptotic cells of the growing blastocyst (Hardy et al., 2003, Palini et al., 2013). This study developed a PCR-based equine Y chromosome assay that targets small fragments of the equine MSY multicopy genes.

6.3.4.2.1. Primer pool constitution and PCR conditions

Unlike the human MSY, the equine MSY was recently comprehensively assembled for the first time (Janečka et al., 2018). However, the horse Y chromosome remains sub optimally characterised. The current horse reference genome, EquCab3.0, is female and considered 'incomplete' due to lacking the Y chromosome sequence (Raudsepp et al., 2019). The number of the equine MSY genes published by Janečka et al. (2018) that presented a copy number greater than 10 is small. Therefore, two loci in the three genes that showed the highest copy number were targeted. Primers were designed for six MSY target regions (in the genes *ETY1*, *ETY4* and *YIR2*) and two autosomal candidates (both loci in the *APOBEC3-Z2* gene). The six MSY targets and the autosomal target with the smallest fragment size of the entire target group were included in primer pool 1 aiming to differentiate the autosomal control amplification band from the MSY genes. The second autosomal control candidate was designed with the same aim but with the largest fragment size.

The design of primer pool 1 presented two main limitations. First, larger-than-expected bands were produced when amplifying commercial male horse gDNA, due to the proximity of the targeted sequences within the same gene. This has the potential to hinder the consistency in the band patterns produced by a male sample due to the different populations of DNA fragment size ranges and their proportions in BF. Consequently, identifying the target bands would be more difficult.

Second, the number of cycles was increased (up to X50), once optimisations reduced the control gDNA input to single-cell levels, and at a higher number of cycles, a primer-dimer band became apparent in the no-template control. This was problematic due to the similarity in size to the first autosomal control (46 bp) tested in the assay and the inability to differentiate between female horse DNA and insufficient template in the sample.

Successful multiplex PCR assays primarily rely on the cycling temperatures, Taq polymerase input and primer concentration (Markoulatos et al., 2002). Different primer pools with a reduced number of pairs and lower primer concentrations were tested, aiming to mitigate the high accumulation of primer-dimers and non-specific bands. The annealing temperature remained unchanged throughout optimisations (62.9°C). Nevertheless, the lowest primer concentration tested (0.1 µM) did not eliminate the primer-dimer band in the no-template control, and neither did the most minimalistic pool targeting only one locus per gene.

The equine Y chromosome assay developed was technically challenging due to the inconsistency in the volumes of BF samples (0.3 – 1 µl) and the minute quantities of template DNA that resulted in a suboptimal primer-to-template DNA ratio. However, according to Markoulatos et al. (2002), the range of primer concentrations for multiplex

PCR should fall between 0.3 µM to 0.5 µM for low sequence repeat numbers or high-complexity DNA. Therefore, the primer concentration was not decreased below 0.1 µM. The final primer pool targeting a single locus per gene and the autosomal control with the largest size (90 bp) successfully differentiated single-cell level commercial male horse DNA, female horse DNA and the no-template control. Nevertheless, non-specific amplification of varying sizes was present in the female DNA control, due to a low PCR stringency caused by the lack of template, a high number of cycles and excess primer concentrations (Lorenz, 2012).

6.3.4.2.2. Equine Y chromosome assay in horse BF samples

Studies of BF as a non-invasive alternative for human embryo genotyping initially explored its potential for sex determination. Palini et al. (2013) reported 26 of 31 human BF samples (89.87%) successfully amplifying with RT-qPCR the chromosome 17 multicopy gene *TBC1D3* with approximately 53 copies in the genome. They targeted a 66 bp long fragment. From the 26 samples, 65.38% also amplified the *TPSY1* multicopy gene located in the Y chromosome with approximately 35 copies targeting a 60 bp long fragment. From the 19 horse BF samples, this study reported amplification in 12 of them (63.16%) and specific amplification, defined as successful amplification of the autosomal control, in 8 of the total samples (42.10%). The sex determination concordance with the sex at birth was 85.71% when also including samples that only showed non-specific amplification but not MSY genes amplification (identified as female). The concordance including only specific amplification was 100%. Palini et al. (2013) only confirmed sex determination concordance, reported to be 100%, with six of the samples that amplified *TPSY1* with cells lost during vitrification and with SCM.

When the autosomal control successfully amplifies, the present Y chromosome assay shows a high level of specificity. This was expected since the purpose of the autosomal control in the experimental design is to confirm the presence of sufficient DNA and of suitable quality in the BF. In fact, all samples that amplified MSY genes also showed amplification of the autosomal control, suggesting that despite the difference in the copy number in the horse genome (13 – 22 for *ETY1*, *ETY4* and *YIR2* and five for *APOBEC3-Z2*), there is no preferential amplification of MSY genes due to PCR selection in the multiplex reaction (Wagner et al., 1994, Markoulatos et al., 2002). The small number of paired BF samples with information of the sex at birth is a limitation for determining if non-specific amplification is an indicator of a BF sample from a female horse. However, from the two samples that only showed non-specific amplification from which the sex at birth is known, one is a male. This suggests that non-specific amplification, when there is lack of amplification of the autosomal control, is an indicator of a low amount of, and bad quality, DNA, and only samples that show amplification of the autosomal control should be considered informative for the horse Y chromosome assay. Palini et al. (2013) did not report any instances where *TPSY1* amplification was successful and amplification of *TBC1D3* was not.

6.3.4.2.3. Horse blastocyst size and BF sample informativity

This study hypothesised that the equine blastocyst provides greater sample informativity for sex determination through Y chromosome amplification than humans, due to its larger size and the greater number of cells harboured. The blastocysts used in this study ranged from 187 µm to 2,000 µm in diameter. Hence, this study also aimed to

reveal if there was a positive correlation between horse blastocyst size and equine Y chromosome assay amplification.

With a similar experimental design to the present study, Herrera et al. (2015) performed a conventional duplex PCR on horse BF samples collected from day-8 post-ovulation blastocysts targeting a 184 bp - 200 bp long sequence of *AMEL*, a single copy gene located in chromosome X and a 280 bp long sequence of *TSPY* with approximately 13 copies in the equine genome, according to Janečka et al. (2018). The gender was determined for 11 of the 13 samples (84.60%). The two samples that failed to amplify were female. Herrera et al. (2015)'s amplification rate is double that of this study. This could be due to the stage of the blastocysts. This study used samples from day-6 and day-7 post-ovulation, whereas Herrera et al. (2015) used day-8 post-ovulation embryos. A study by Colchen et al. (2000) reported a doubling in the number of cells in horse blastocysts every six hours between day-6 and day-7 post-ovulation (the day-6 total number of cells mean was 275 and for day-7 was 1,093). In line with these findings, an early study on equine embryo development by Betteridge et al. (1982) observed a considerable increase in diameter between day-7 post-ovulation fresh retrieved horse embryos (ranging 238 µm – 768 µm) and day-9 (ranging 1,700 µm – 4,000 µm). The diameter of day-8 post-ovulation equine embryos could range from 120 µm to 3,980 µm (Vanderwall and Woods, 2007).

Unfortunately, Herrera et al. (2015) did not report the diameters of the blastocysts included for BF aspiration. Nevertheless, it can be speculated that they used larger blastocysts that harboured a significantly greater number of cells, compared to this study. The small sample size in this study is a limitation to draw accurate conclusions. But, contrary to expected, there was no observable positive correlation between blastocyst

diameter and Y chromosome amplification success. Thus, when speculating that the BF recovery was optimal and there was no loss of significant fluid volume, BF from larger blastocysts were not more informative in this study.

Only a few studies target multicopy genes in BF to compare amplification levels between different experimental designs and sample characteristics. With the data gathered, however, it can only be hypothesised that the difference in levels of amplification between this study, Palini et al. (2013) and Herrera et al. (2015) could be due to differences in the embryo developmental stage at the moment of sample collection, in the handling and storage of samples or in the experimental design. Future studies with horse BF should include a larger MSY pool targeting a single locus per gene and samples obtained at different stages of blastocyst development (day-6, day-7 and day-8 post-ovulation) with statistically representative sample sizes to detect a real impact of blastocyst developmental stage in sample informativity. Additionally, samples can be stored in EDTA buffer to preserve the DNA, following Palini et al. (2013)'s methodology, since cfDNA is highly unstable.

6.3.4.3. Conclusion

The mammalian Y chromosome harbours multicopy genes located in the MSY region. These genes can be targeted in multiplex PCR for gender determination, particularly in instances where the template DNA is found in low abundance and fragmented. This chapter aimed to develop two highly sensitive non-invasive methods for the detection of the embryonic (or foetal) Y chromosome in human maternal blood, during early human pregnancy and in horse BF.

This study demonstrated the suitability of a simplified plasma separation methodology for the detection of the Y chromosome in maternal blood. It shows that the strict separation of maternal cellular components is not needed, due to the sensitivity of the Y chromosome assay, despite the high levels of maternal DNA background. In addition, fragments of placental origin are the smallest from the maternal circulating cfDNA population, and this study indicated the requirement of specialised kits for the isolation of cfDNA to achieve optimal recovery yields. This chapter also showed that, unlike the Y chromosome assay for maternal blood, the method for the detection of MSY genes in BF requires the inclusion of an autosomal control as an indicator of the DNA quality and quantity found in samples. Contrary to expectations, this study provided evidence that the larger size of the equine blastocyst, compared to the human, did not show higher sample informativity rates. The larger size within the blastocyst fluid sample cohort did not demonstrate this correlation either, but later blastocyst developmental stages (day-8 vs. day-6 or day-7) may have a positive impact on DNA amplification. For both methodologies, non-specific amplification was highly prevalent, likely due to the high number of cycles, maternal DNA background (in the maternal blood study) and insufficient DNA template (in the horse BF study). Nevertheless, both Y chromosome assays demonstrated high sensitivity and specificity, in the case of the equine Y chromosome when amplification of the autosomal control was successful.

This study successfully determined the gender of pregnancies with small volumes of venous blood, opening the possibility of testing microfluidic devices for capillary blood sample recovery and plasma separation. In addition, the high levels of MSY genes amplification from samples with low foetal fractions prompts future studies where the human chromosome assay is tested in gestational ages younger than week-10 to

determine a gestational age limit detection. For the horse Y chromosome assay, day-8 post-ovulation embryos (late blastocysts) could be tested to determine the effect of developmental stage in horse BF sample informativity. Additionally, further optimisations could be performed on Y chromosome assay thermal protocols, with particular attention to extension times, to reduce non-specific amplification.

Non-invasive PCR-based sex determination with maternal blood is highly sensitive at week-10 gestational ages. In addition, there is a lack of evidence supporting the hypothesis that the equine BF produces a higher informativity rate than the smaller human blastocyst for PCR-based sex determination.

Experimental section 3. Epigenetic avenues for embryo genetic testing

Chapter 7. Development of a targeted methylation analysis protocol for 46 imprinted genes

The study design of this chapter was developed by Ph.D. Dhruti Babariya. In a previous pilot study, Miss Janvi Patel optimised the library preparation conditions with an initial panel that included eight of the 46 targeted DMRs in this study. Those conditions were used as a baseline.

7.1.1. Introduction

Epidemiological studies have consistently suggested a correlation between ART-conceived children and an increased incidence of rare imprinting disorders (Vermeiden and Bernardus, 2013, Cortessis et al., 2018). Although reports have hypothesised that it is the infertility of the parents that predisposes progeny to such epigenetic abnormalities (Doornbos et al., 2007), subsequent molecular studies using animal models found an altered monoallelic expression of *H19*, a widely studied imprinted gene, in embryos conceived after ovarian hormonal stimulation compared to controls (Fauque et al., 2007). Loss of methylation was also found in maternally imprinted genes including *Snrpn*, *Kcnq1ot1* and *Peg3* (Market-Velker et al., 2010a). Fauque et al. (2007) showed that IVF

per se can affect the methylation of *H19,* while Market-Velker et al. (2010a) found varying degrees of loss of methylation in *Snrpn*, *Kcnq1ot1* and *Peg3,* depending on the embryo culture media used.

Although recent publications have consistently showed the impact of ARTs on genomic imprinting, mainly reporting on *H19*, *SNRPN*, *KCNQ1OT1,* and *PEG3* (Kindsfather et al., 2019), studies produce varying results on the analysis of the DNA methylation. For instance, using mice, Market-Velker et al. (2010a) employed BS for analysis and distinguished maternal from paternal alleles. Therefore, the expected methylation level of the corresponding imprinted allele for *H19* was 100%. The group found values ranging from 82% to 100% for the in vivo derived embryos and from 21% to 94% for the in vitro conceived. Kindsfather et al. (2019) also performed BS in mice blastocysts to analyse the imprinting of *H19* and found, for the naturally conceived group, imprinted allele methylation percentages ranging from 75% to 98% and in the in vitro conceived embryos, values from 30% to 80%. Thus, robust methods for DNA methylation analysis of imprinted loci, in addition to the employment of appropriate controls, are needed to overcome these variations to determine the presence of hyper or hypomethylation accurately.

7.1.1.1. Mechanisms of genomic imprinting

Genomic imprinting was discovered during experiments with mice where maternal or paternal pronuclei were removed and reintroduced into recipient zygotes to produce embryos with two maternal genomes (gynogenetic) or two paternal genomes (androgenetic). Gynogenetic gestations exhibited underdeveloped placentas, while androgenetic gestations presented extraembryonic tissues almost in totality (McGrath

and Solter, 1983, McGrath and Solter, 1984). This shed light, for the first time, on the importance of non-equivalent expression of a specific set of genes during development.

7.1.1.1.1. Genomic imprinting function

Phenotypic studies in mammals have shown that imprinted genes have essential roles in foetal growth, lineage development, brain function and postnatal homeostatic control of energy balance (Bartolomei and Ferguson-Smith, 2011). Early experiments with mice established a correlation between the disruption of imprinting in *H19* and *Igf2* and defective placental function and development (Leighton et al., 1995, Eggenschwiler et al., 1997). Recent research in humans has confirmed that the abnormal expression levels of *H19/IGF2,* caused by hypomethylation of the germline-derived differentially methylated region (gDMR) controlling this imprinted gene cluster, can lead to pregnancy complications (Yamaguchi et al., 2019).

Understanding the function of genomic imprinting in humans has partially been attributed to the study of the phenotypes exhibited in patients with disorders caused by imprinting aberrations (Bartolomei and Ferguson-Smith, 2011). For instance, Angelman syndrome, a neurogenetic disease, is caused by the loss of the UBE3A transcript, an imprinted gene expressed from the maternal allele (therefore, paternally imprinted). *UBE3A* belongs to the *SNRPN/UBE3A* imprinted gene cluster located on human chromosome 15. This cluster is known as the Prader-Willi syndrome/Angelman syndrome domain. The *SNRPN/SNURF* gene is also found in this cluster, which, contrary to *UBE3A*, is maternally imprinted. A gDMR at the 5' end of the *SNRPN/SNURF* gene controls the monoallelic expression of several maternally expressed imprinted genes, including *SNRPN/SNURF*. *UBE3A*, however, is imprinted in the opposite direction (Nicholls and

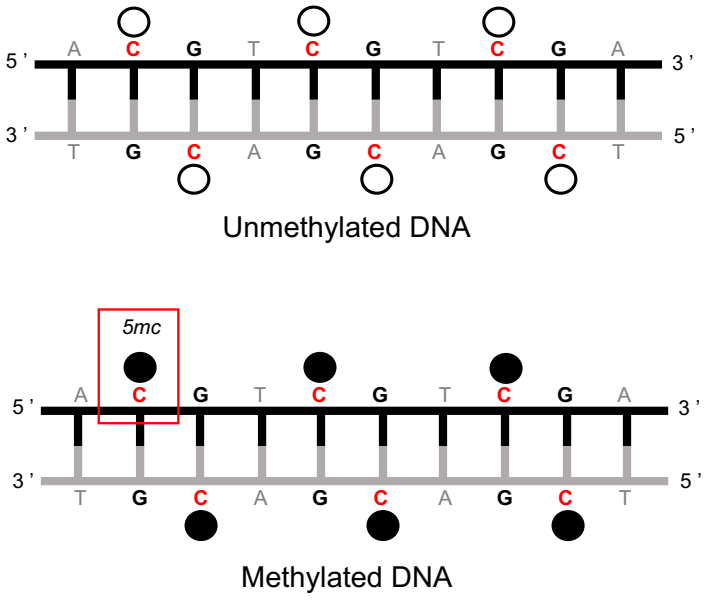
Knepper, 2001). Thus far, around 100 imprinted genes have been identified in humans (Monk et al., 2019).

7.1.1.1.2. Molecular mechanisms of genomic imprinting

The principal epigenetic mark involved in genomic imprinting is DNA methylation. In mammals, DNA methylation occurs almost exclusively in the context of CpG sites (CpGs). CpGs are dinucleotides where a cytosine (C) is followed by a guanine (G). When a CpGs is methylated, the cytosines on both strands present the 5mC molecule (Figure 51(A)) (Paulsen and Ferguson-Smith, 2001). The CpGs methylation percentage analysis of DMRs is expected to be close to 50%.

GDMRs, or primary DMRs, are differentially marked during gametogenesis. Most imprinted genes are gathered in clusters regulated by a shared gDMR. The most widely studied imprinted gene cluster is the *H19/IGF2* (Figure 51(B)). Secondary DMRs that acquire their parent-of-origin DNA methylation post-implantation assist in the maintenance of imprinted expression by primary DMRs.

(A)



(B)

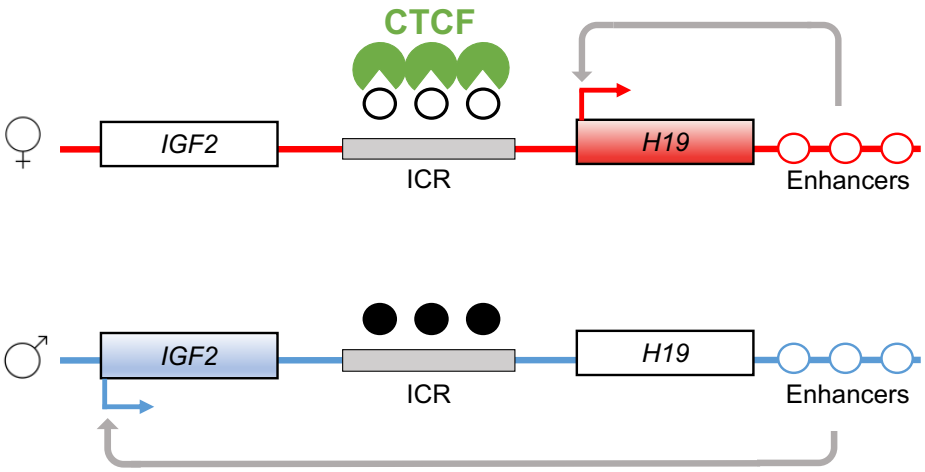


Figure 51. Molecular mechanisms of genomic imprinting. (A) Unmethylated and methylated CpGs in double-stranded DNA**.** Areas of the genome that are CpGs-rich are known as CpG islands and are found in DMRs that control the expression of imprinted genes. Typically, the CpGs of the imprinted allele are methylated, while the CpGs of the allele from where the gene is expressed are unmethylated (Bartolomei and Ferguson-Smith, 2011). **(B)** *H19/IGF2* cluster gene expression regulation mechanism. In normal circumstances, the ICR (or gDMR) on the maternal allele is unmethylated, allowing the binding of the CTCF transcriptional repressor that blocks the activity of downstream enhancers to access the *IGF2* promoter. Thus, the *IGF2* expression in the maternal allele is repressed while *H19* is expressed. The ICR on the paternal allele is methylated, CTCF cannot bind, allowing the access of downstream enhancers to *IGF2* promoters. Therefore, *IGF2* is expressed in the paternal allele and the expression of *H19* is repressed. CTCF; CCCTC-binding factor, ICR; imprinting control region, 5mc; 5-methylcytosine.

7.1.1.2. Technical challenges in CpGs methylation analysis

The current gold standard for methylation analysis is BS. The basis of BS is the sodium bisulfite ($NaHSO_3$) treatment of DNA to convert unmethylated cytosines (C) to uracils (U). Methylated cytosines are not converted. In addition, BS uses high-throughput sequencing allowing for single nucleotide resolution during analysis. BS is generally used for whole genome analysis, although targeted approaches are employed when the loci of interest are amplified through PCR to interrogate only a fraction of the genome. In a standard targeted BS protocol, DNA is bisulfite-converted first, followed by PCR amplification for library preparation. The resulting libraries' uracil (U) is represented as thymine (T). Libraries are subsequently sequenced and analysed. An accurate recall of

5mC molecules depends on various elements, including bisulfite conversion and amplification efficiency and sequencing depth (Wreczycka et al., 2017).

Even though BS is widely used for CpGs methylation analysis in epigenetics research, it has some technical limitations. Firstly, bisulfite conversion is an aggressive chemical treatment that can lead to excessive DNA degradation and incomplete unmethylated cytosine conversion. This is a challenge when working with minute quantities of starting DNA obtained from single-cells (Pappas et al., 2013). Secondly, the complexity of the DNA sequences abruptly decreases after bisulfite treatment, due to the highly elevated level of adenine (A) and thymine (T) base pairs in the converted DNA (Reinders and Paszkowski, 2010). Consequently, primer specificity reduces substantially, affecting PCR amplification efficiency during library preparation and quality base calls during NGS. Finally, BS can be particularly prone to PCR amplification bias leading to overrepresentation of the methylated or unmethylated allele. This significantly interferes with the biologically true frequency of alleles, affecting the accuracy of the CpGs methylation analysis (Olova et al., 2018).

Initial studies on DMRs in human gametes and embryos employing targeted BS approaches failed to provide apt controls to assess the accuracy of the CpGs methylation analysis (Ibala-Romdhane et al., 2011, Khoueiry et al., 2013). However, publications that followed used heterozygous SNPs to distinguish the two parental alleles (White et al., 2015, Denomme et al., 2016). Heterozygous SNPs, like DMRs CpGs methylation, are inherited in a parent-of-origin-specific manner. Thus, when these polymorphisms are in the vicinity of, or within the DMR, they can be employed either to identify the maternal and paternal alleles and perform CpGs methylation analysis separately or as controls to identify differences in allele amplification. However, the presence of heterozygous SNPs

varies among individuals. Moreover, apart from allele amplification bias, PCR duplicates also contribute to the distortion of true allele frequency caused by sequence overrepresentation. Therefore, robust methods to overcome amplification bias in BS must be developed.

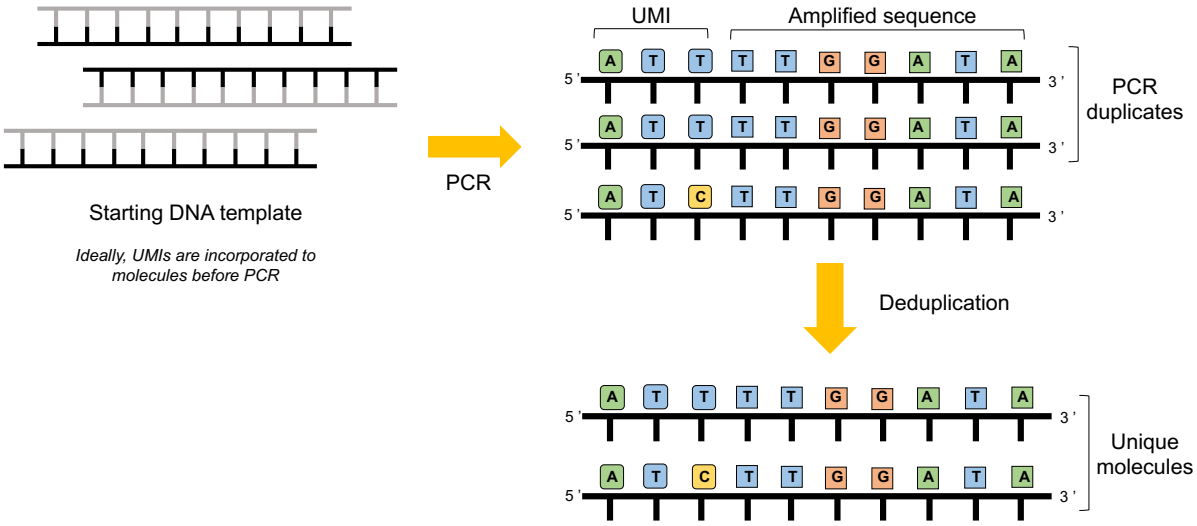


Figure 52. Removal of PCR duplicates with unique molecular identifiers (UMIs). UMIs are DNA fragments of randomly allocated nucleotides incorporated into the sequencing read, ideally, before any PCR step of the library preparation. During bioinformatical analysis, PCR duplicates are identified if the reads map to the same genomic coordinates with the same UMI sequence (Clement et al., 2018). In this example, the length of the UMI is three nucleotides. Reads that have the same amplified sequence, or map to the same genomic coordinates, and have the same UMI sequence, are most likely PCR duplicates. Deduplication is performed computationally, so only unique molecules are employed for downstream analysis to draw biologically accurate conclusions when allele frequencies are paramount for quantification, such as DMRs CpGs methylation analysis.

In recent years, unique molecular identifiers (UMIs) have been used to count molecules directly in single-cell RNA-seq experiments to correct PCR-induced artefacts (Islam et al., 2014). Some of these artefacts include PCR duplicates, which occur when two reads are produced from the same original DNA fragment during amplification, thereby distorting the true abundance of molecules in the biological sample and affecting the accuracy of downstream analyses (Rochette et al., 2023). UMIs have mostly been

successfully utilised for RNA-seq studies and have yet to be employed in the context of BS to investigate their utility for CpGs methylation analysis of imprinted genes (Figure 52).

7.1.1.3. Study aims

This chapter aims to develop a targeted BS approach methodology that simultaneously detects methylation perturbations in 46 DMRs in human preimplantation embryonic cells using a multiplex PCR during library preparation as well as directly counting molecules with the use of UMIs.

The specific aims are:

1) To optimise the PCR conditions of the library preparation using commercial human DNA and clumps of buccal cells.
2) To uniformly amplify the 46 targeted DMRs during library preparation.
3) To determine the average CpGs methylation percentages in the targeted DMRs after PCR duplicate removal with UMIs.
4) To employ, when present within the DMR, heterozygous SNPs as reference to assess the performance of the UMIs in increasing the accuracy of CpGs methylation quantification in the present protocol.

7.1.2. Materials and methods

7.1.2.1. Experimental design for the method development

7.2.2.1.1. Study design

Figure 53 presents this chapter's study design. GDNA is first extracted from TE biopsies and subsequently subjected to bisulfite conversion. In this chapter, however,

clumps of eight to nine buccal cells were used as representative of TE biopsies. The library preparation protocol is a modification of the Illumina TruSeq (Illumina, USA) workflow. During the first PCR (PCR1) the 46 DMRs are amplified in a multiplex reaction with a primer pool. A UMI is included in the forward primer of all the 46 targets to allow for its incorporation into every strand copied. During the second PCR (PCR2) forward and reverse universal Illumina primers and sample multiplexing indices are attached to the products of PCR1. The libraries are then sequenced and the data bioinformatically analysed where UMIs are used for PCR duplicate removal. The unique reads are subsequently mapped to the reference genome for CpGs methylation calling.

7.2.2.1.2. Differentially methylated regions (DMRs) panel and Illumina indices

The specific primer design was performed by Ph.D. Dhruti Babariya, as set out in **Chapter 2**. Primer sequences and characteristics of the target loci are presented in Table 43. Details of the DMRs included in this study can be found in **Appendix D**. Figure 53 shows the final sequences of the primers for PCR1. The UMIs used comprised eight randomly allocated nucleotides.

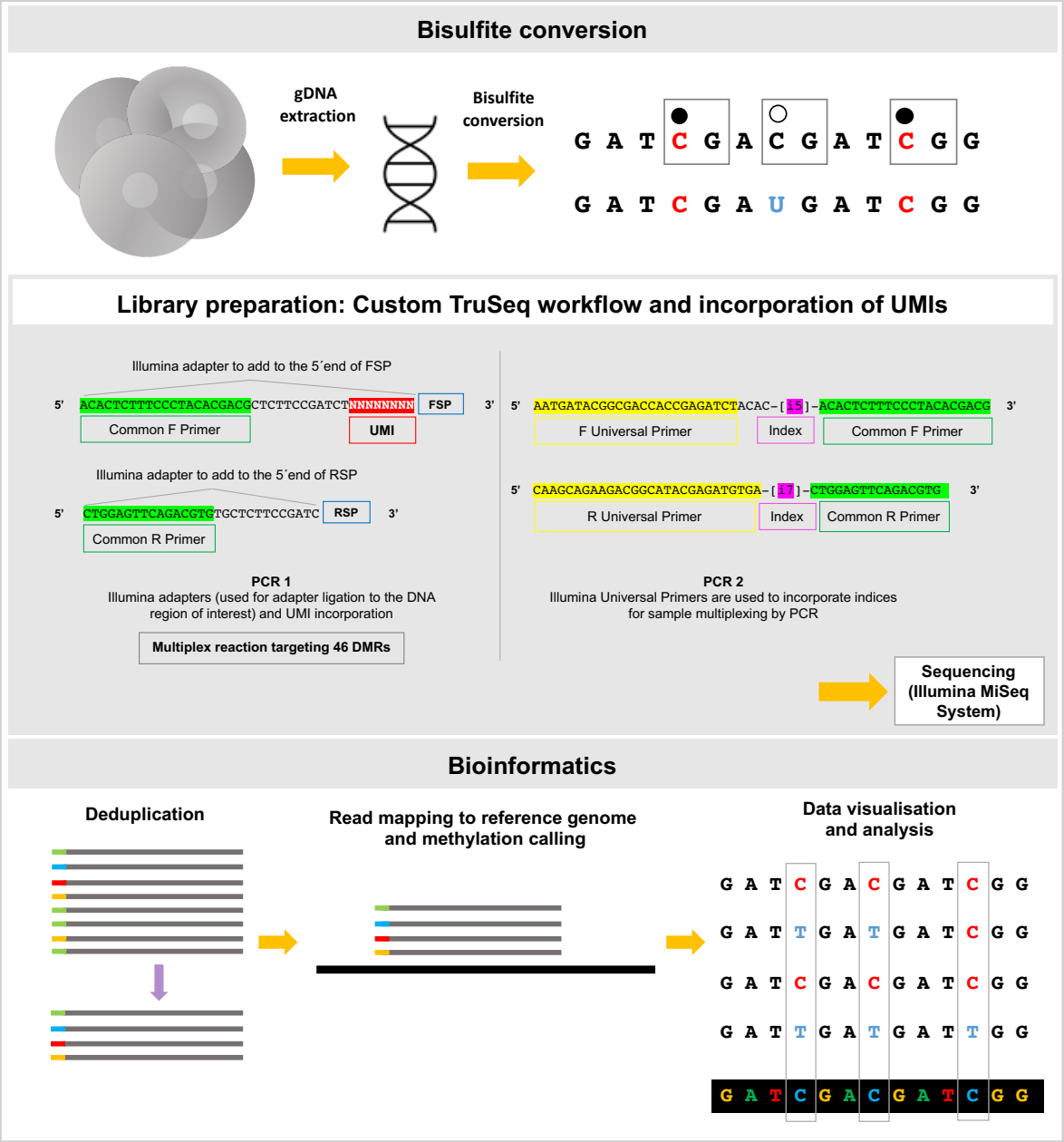


Figure 53. Study design for the development of a targeted BS protocol for the methylation analysis of 49 DMRs. FSP, forward specific primer; RSP, reverse specific primer.

Table 43. Specific primers sequences for the 46 targeted DMRs.

Gene/DMR	Primer sequences	Amplicon size	Number of CpGs
PPIEL F:	5’ TTAATGGATTAGAGTTGTGGAGTTTG 3’	387	20
PPIEL R:	5’ ACCCCTTCTTTCCTTAATCTAATACC 3’		
DIRAS3 F:	5’ TTTGGTTTTAGTTTTAGGGATAATT 3’	334	13
DIRAS3 R:	5’ CAACCCCTCACAACTAATTTCTTAC 3’		
DIRAS3 Ex2 F:	5’ GTTGTAGTTTAGTAATTGGTAGT 3’	322	18
DIRAS3 Ex2 R:	5’ CACTTTCTTCCTAAAAATCT 3’		
GPR1-AS F:	5’ GTATTAATTTGGTTAAGGTTGGATG 3’	365	14
GPR1-AS R:	5’ CCAAAACAAAATTATCTAAAAAAAA 3’		
ZDBF2 F:	5’ GTATGTTTTAGTTTTTTTGTTGTGGG 3’	375	4
ZDBF2 R:	5’ AACATAAACTACCCCTAAAAATTCCA 3’		
NAP1L5 F:	5’ GTATTTTTTGTAGGGGTTTGGG 3’	358	19
NAP1L5 R:	5’ AATAATAAAAAACCAATCAAATACC 3’		
VTRNA2-1 F:	5’ TTAGTTTTAGGTGTAGTAGGGGGTT 3’	378	18
VTRNA2-1 R:	5’ ACAAACCCAAAAACAATAAATATTC 3’		
FAM50B F:	5’ GGAGTTTGGTATTTTTTTTAGGGTT 3’	313	21
FAM50B R:	5’ AACACCAATTTCCAACATTTTATTTA 3’		
WDR27 F:	5’ TATTATTAGGTATTTTGGGAAAGGAG 3’	336	10
WDR27 R:	5’ CTAACCAAAAATTACTCCAACAACAC 3’		
HTR5A F:	5’ GGAGGTTGTGTTAGTTTTGGAT 3’	348	25
HTR5A R:	5’ CAACCTTATAAATCTTCCAATACAC 3’		
ERLIN2 F:	5’ GGTATGGTTTTTAGGAGTAATTTGAAA 3’	383	13
ERLIN2 R:	5’ ACCAAAAATCAACCAATTTAACAAC 3’		
INPP5F F:	5’ GTGTAGGTAGTTGTTTGAGTTT 3’	324	16
INPP5F R:	5’ AATTACCAACACCACATAAAACC 3’		
MKRN3 F:	5’ GTTTTTTGTTAATGTTTTGTTGGTG 3’	393	14
MKRN3 R:	5’ CCAACCCCTTTTATCTACCTTCTAC 3’		
MAGEL2 F:	5’ GTTGGGTTTTTTTTTTTAGAGAG 3’	380	9
MAGEL2 R:	5’ TCTTAAATACCTTCCATATTTCC 3’		
NDN F:	5’ AGATTTTTATTTTGTTTTGATATGTTTG 3’	379	34
NDN R:	5’ AAAATATCTTTCCCTTAAACTAACCC 3’		
IGF1R F:	5’ AAGGTTTTTTATTTGTTTTTGATTA 3’	326	14
IGF1R R:	5’ ACTCTAAAAACCTCCTCCTACCTAC 3’		
ZNF597/NAA60 F:	5’ TTGGTTAGAGTAAGGTGATTGAGAG 3’	328	19
ZNF597/NAA60 R:	5’ CCAAATACCTACACCAAAACACTC 3’		
ZNF331 variant 7 F:	5’ GGTTTTTAGTTGGGGTGATATAGTTT 3’	301	10
ZNF331 variant 7 R:	5’ CACCCCTAAATTCCTCCTAAAAA 3’		
MCTS2P F:	5’ TAGTTTGTTAAGTTTTATTGTTTTAG 3’	326	25
MCTS2P R:	5’ ACAATTAAACACACTTTCCTTTTCAT 3’		
WRB F:	5’ GGTGTGGGTTGAGTGGTTATA 3’	334	9
WRB R:	5’ CCAAACCTAATCCATAACAAAAACT 3’		
NHP2L1 F:	5’ TTTATGTAGGGAGTATATTGGGTGAGT 3’	306	10
NHP2L1 R:	5’ TAAATAAATAATCCCACCAAACCAT 3’		
PLAGL1 F:	5’ GGTAGTTAAGAGGATGGTTG 3’	337	32
PLAGL1 R:	5’ TTTCTCATATATAATTAAACTCTAAC 3’		
IGF2R F:	5’ TGGTTTTTGAATTTGTTAGGAAATAT 3’	363	10
IGF2R R:	5’ ACACAAAAATAAAAAAATATAAAAC 3’		

GRB10 F:	5’ TAATTTGGAATTTTTTTTATTTTTTTATGT 3’	303	6
GRB10 R:	5’ AAATTTTCTTTCTCTCTATCACTCCAA 3’		
PEG10 F:	5’ GGATTGTTTTTGTTGTATGAAATGA 3’	393	14
PEG10 R:	5’ AAACAAAACCTAAACAACTCACCAC 3’		
MEST F:	5’ TGGTGGGTTTAATAGAGTTTGTTGT 3’	290	16
MEST R:	5’ CCTAAATACCCCAACTCTTTCCTTA 3’		
IGF 3’UTR F:	5’ AAGAGGTTTTAGAGGTTATAGGAAA 3’	325	18
IGF 3’UTR R:	5’ AAAAATAATTTACTCCCCCTTCAAC 3’		
IGF DMR0 F:	5’ GGTTATTTTGTTAGGATTTGGGTAG 3’	322	6
IGF DMR0 R:	5’ TTTACTAAAACCCCCTCCTTATCTC 3’		
KCNQ1OT1 F:	5’ AGGTTATTTATTTGGTAAAGGGTAG 3’	322	23
KCNQ1OT1 R:	5’ CCAAAATCCTACATTCCAACACTAT 3’		
RB1 F:	5’ GATTATGTTTTGAGGGTTAAAA 3’	379	24
RB1 R:	5’ CACCCAACTAAAAAAAACCAATAAC 3’		
MEG3 F:	5’ TGTAGTTTTGGATTTTTGGTTAAAT 3’	300	22
MEG3 R:	5’ CTCAAAACTATTCCCTCTTAAATAAAC 3’		
SNRPN variant 4 F:	5’ GTTTTTTATTTTTGTATTTTAAAATGTTAT 3’	339	3
SNRPN variant 4 R:	5’ ATTCTATCTACACAAATTAATTCCTTATAC 3’		
ZNF597 F:	5’ GGTTAAATTTTATTATTGGATTTTTAT 3’	383	4
ZNF597 R:	5’ CAAAAACAAAAACCTCTATCTTCAC 3’		
PEG3 F:	5’ AGTAAAGATGATGTTATAGAGTGGG 3’	341	10
PEG3 R:	5’ ATAAAACTAAAAACACCCTCACCTC 3’		
NNAT F:	5’ GGTATAAAGAATTTTATTGGTTAAGGT 3’	359	17
NNAT R:	5’ AATTTTAAACCCAAATCCTCTACTTC 3’		
L3MBTL1 F:	5’ GGTAGGGTTATTAGATTTATTTGTGT 3’	396	15
L3MBTL1 R:	5’ TAAAACCCATCATCTAAACCAACTT 3’		
GNAS-AS1 F:	5’ GTGTTTGTGGTATTTTGGTTTTAAGA 3’	380	15
GNAS-AS1 R:	5’ ATACCCACACCACCTAATAACCTTA 3’		
GNAS-XL F:	5’ TTTAAGGTTAAGAAGGTATTTTTGG 3’	313	17
GNAS-XL R:	5’ TAAAACTTTCTAACAAATTCTTCCC 3’		
*GNAS A/*B F	5’ GTGTTTAAGAGTTAAATTAAGGAG 3’	381	17
GNAS A/B R:	5’ ATAACCATCTTCAACATAATAACCC 3’		
GNAS-NESP F:	5’ GGAGTTTTAGAGTTTTAGGGAAGG 3’	355	16
GNAS-NESP R:	5’ CTAACCAACTAAACCTCCTCTCTT 3’		
H19 F:	5’ TTGGTTATTTAGGGTTGAGGAGAT 3’	300	8
H19 R:	5’ AAAATCATCTAAAAATAAAACACTCATAAA 3’		
IG-DMR F:	5’ GGAGGAAAGTTGTTGTTAAATTGTT 3’	316	2
IG-DMR R:	5’ TACAAAAATACCACACCCACCTAAT 3’		
MEG8 F:	5’ GGTTGTTTGGGTTATATTTTTGAGTT 3’	317	22
MEG8 R:	5’ CTAAAAATACTCCCTCCCTCATCTC 3’		
SNRPN/SNURF F:	5’ AGGAGGTTATGGTAGTGGATTAGG 3’	383	19
SNRPN/SNURF R:	5’ CACCACAATAAACAAACCAAATAAC 3’		
PEG3_UMI F:	5’ TGAGTTTGGTTTTGGATTTTATTTT 3’	249	11
PEG3_UMI R:	5’ ATTCATTCATTCTATACCATCTTTACC 3’		
ZNF331 variants *4-6* F:	5’ TTGTTTATAATGTAGTAAGTAAGGGATAT 3’	333	15
ZNF331 variants *4-6* R:	5’ ACTCCTAAATAATCTCTAAAAAATATAATC 3’		

SNPs were found during the specific primer design for seven targets. The DMRs and their corresponding SNPs are shown in Table 44. The SNPs were searched and selected, according to the process in **Chapter 2**. However, during the 5th phase of this study, additional heterozygous SNPs were identified and included as controls. The existence of these SNPs was corroborated using the NCBI SNP database (https://www.ncbi.nlm.nih.gov/snp/).

Table 44. SNPs included in the targeted region of interest of seven DMRs. Coordinates are from the GRCh37/hg19 genome assembly.

Gene/DMR	SNP (ID, coordinates and population frequency)
H19	rs11042167 (chr11:2,020,901) (A: 29.29%; G: 70.71%)
SNRPN/SNURF	rs12916854 (chr15:25,200,902) (A: 62.90%; G: 37.10%)
ZNF331 variants 4-6	rs67529716 (chr19:54,040,861) (C: 16.87%; G: 83.13%)
GNAS-NESP	rs1800905 (chr20:57,415,995) (A: 40.59%; G: 59.40%)
GRB10	rs6958750 (chr7:50,848,873) (C: 39.68%; G: 60.32%)
IG-DMR	rs1884539 (chr14:101,277,786) (A: 49.30%; G: 50.70%)
MEG8	rs10135782(chr14:10,1371,131) (C: 48.24%; G: 51.76%)

Both forward and reverse primers for PCR2 contained an index to allow for the inclusion of multiple samples per sequencing run (Table 45). Five different indices were used for the forward primer and two for the reverse to allow multiplexing for up to 10 samples.

Table 45. Illumina indices i5 and i7 sequences.

i5 indices (forward primer)	
ID	**Sequence**
[N/S/E]501	5' TAGATCGC 3'
[N/S/E]502	5' CTCTCTAT 3'
[N/S/E]503	5' TATCCTCT 3'
[N/S/E]504	5' AGAGTAGA 3'
[N/S/E]505	5' GTAAGGAG 3'
i7 indices (reverse primer)	
N701	5' TCGCCTTA 3'
N702	5' CTAGTACG 3'

7.2.2.1.3. Library preparation optimisation

Overall, the library preparation optimisation experiments were divided into five phases, set out in Figure 54 and described below:

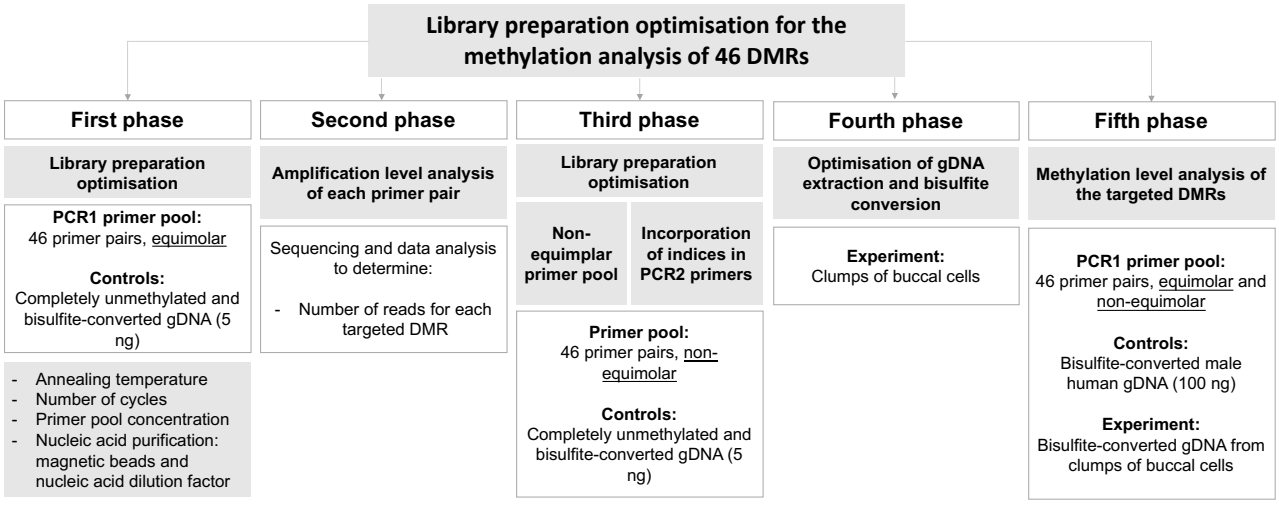


Figure 54. Library preparation optimisation for the methylation analysis of 46 DMRs.

First phase: An initial PCR1 and PCR2 were optimised with commercial completely unmethylated and bisulfite-converted human gDNA (Qiagen, Germany). The primary purpose of this initial phase was to find the optimal annealing temperature for PCR1, the

number of cycles and primer pool concentrations for both PCR1 and PCR2, and an optimal magnetic beads:nucleic acid dilution factor for maximal DNA recovery with minimal primer-dimer carryover after nucleic acid purification of PCR1 and PCR2 products. All amplifications for PCR1 were performed with an equimolar primer pool containing the 46 primer pairs.

Second phase: Libraries were sequenced aiming to determine the level of amplification for each primer pair by counting the number of reads mapped to the reference genome without deduplication. The primer concentrations in the pool were subsequently adjusted based on the level of amplification of each target in a non-equimolar primer pool.

Third phase: PCR1 conditions were optimised with the non-equimolar primer pool and PCR2 conditions with the incorporation of the indices to the primer sequence.

Fourth phase: Two kits were used to assess the level of bisulfite-converted DNA recovery of clumps of eight to nine buccal cells. The study then proceeded to perform library preparation optimisation with this material.

Fifth phase: Libraries from clumps of buccal cells and from commercial human male gDNA (Promega, USA) were sequenced. For the commercial DNA, libraries were prepared using only the equimolar primer pool and libraries for the clumps of cells used the equimolar and non-equimolar primer pools. The sequencing reads were deduplicated for CpGs methylation calling.

7.1.2.2. Buccal cell collection and isolation

Buccal cells were collected from a karyotypically normal and healthy volunteer using the Catch All sample collection swab kit. The buccal cells were then isolated by Ph.D. Dagan Wells and Mr Ayman Haj Ali following the instructions in **Chapter 2**.

7.1.2.3. Genomic DNA (gDNA) extraction

The clumps of cells were subjected to gDNA extraction using Proteinase K, as per **Chapter 2.**

7.1.2.4. Bisulfite conversion

The EpiTect Bisulfite and the Imprint Modification kits were tested on the extracted gDNA to convert unmethylated cytosines according to the protocols in **Chapter 2**. Ten ng of commercial human male gDNA was used as positive control for each conversion reaction. All bisulfite-converted DNA was eluted in 20 µl of elution buffer.

7.1.2.5. Library preparation

7.1.2.5.1. PCR1 and PCR2

The 20 µl of eluted bisulfite-converted DNA was used directly for PCR1 to amplify the 46 targets with the Multiplex PCR kit, as per **Chapter 2**. For optimisation purposes, commercial completely unmethylated and bisulfite-converted human gDNA was used as template DNA for a significant portion of this study. As mentioned, previously optimised conditions for a panel that included eight of the 46 targets were used as baseline (Table 46). The recovered fragments were visualised on the TapeStation with D1000, D1000

High Sensitivity or Genomic DNA reagents, following **Chapter 2**. Singleplexes were run for some experiments to confirm specific target amplification after PCR2, following the instructions in **Chapter 2,** in a 4 % agarose (4 g of agarose) gel electrophoresis at 90 volts for 1 hour and 20 minutes. Libraries were stored at -20°C until sequencing.

Table 46. Previously optimised conditions of the library preparation for eight DMRs. Different values for primer concentration, PCR1 input DNA, PCR1 annealing temperature and the number of cycles were tested throughout the library preparation optimisation. '*' denotes values that were optimised or modified throughout this study.

PCR1			
Table of components			
Component	**Volume (µl)**	**Final concentration/input**	
2X QMM	7.50	1X	
Primer pool (2.14 µM, equimolar *)	1.40 *	0.20 µM *	
Bisulfite-converted DNA (control) (10 ng/µl)	0.50 *	5 ng *	
Nuclease-free water	5.60 *		
Total volume	15 *		
Thermal protocol			
Step	**Temperature**	**Time**	**Cycles**
Initial activation	95°C	10 minutes	1
Denaturing	94°C	30 seconds	X20 *
Annealing	56°C *	90 seconds	
Extension	72°C	1 minute	
Final extension	60°C	10 minutes	1
Hold	4°C	∞	1
1:1 beads:DNA clean-up ratio after PCR1 *			
PCR2			
Table of components			
Component	**Volume (µl)**	**Final concentration/input**	
2X QMM	7.5	1X	
Universal Primers (10 µM, equimolar)	0.30 *	0.2 µM *	
Bisulfite-converted DNA (control) (10 ng/µl)	3 *		
Nuclease-free water	4.20 *		
Total volume	15 *		
Thermal protocol			
Step	**Temperature**	**Time**	**Cycles**
Initial activation	95°C	10 minutes	1
Denaturing	94°C	30 seconds	X20 *
Annealing	51.2°C	2 minutes	
Extension	72°C	1 minute	
Final extension	60°C	10 minutes	1
Hold	4°C	∞	1
1:1 beads:DNA clean-up ratio after PCR2 *			

7.1.2.5.2. Nucleic acid purification

The products of the PCRs were purified using the AMPure XP magnetic beads. The beads were washed away with 80% ethanol, the purified nucleic acids of PCR1 were eluted in nuclease-free water and the products of PCR2 in Elution Buffer (Illumina, USA). Different dilution factor, or ratios, between magnetic beads and DNA were tested to maximise the recovery of DNA fragments with a size ranging from 302 bp to 449 bp after PCR1 and from 372 bp to 519 bp after PCR2 and to minimise the presence of primer-dimers, which size was below 200 bp.

7.1.2.6. NGS

After purification, the libraries were normalised and pooled. Four nM (for the first sequencing run) or 30 pM (for the second) of libraries were subsequently denatured with 0.2 N of NaOH or at 96°C for 10 minutes in the thermocycler (for the first and second sequencing run respectively). Then 10 pM (first sequencing run) or 12 pM (second sequencing run) of the denatured libraries were loaded with the MiSeq Reagent Kit V2 (600 cycles) using single-read sequencing. The base calling and data acquisition was performed in the MiSeq Instrument.

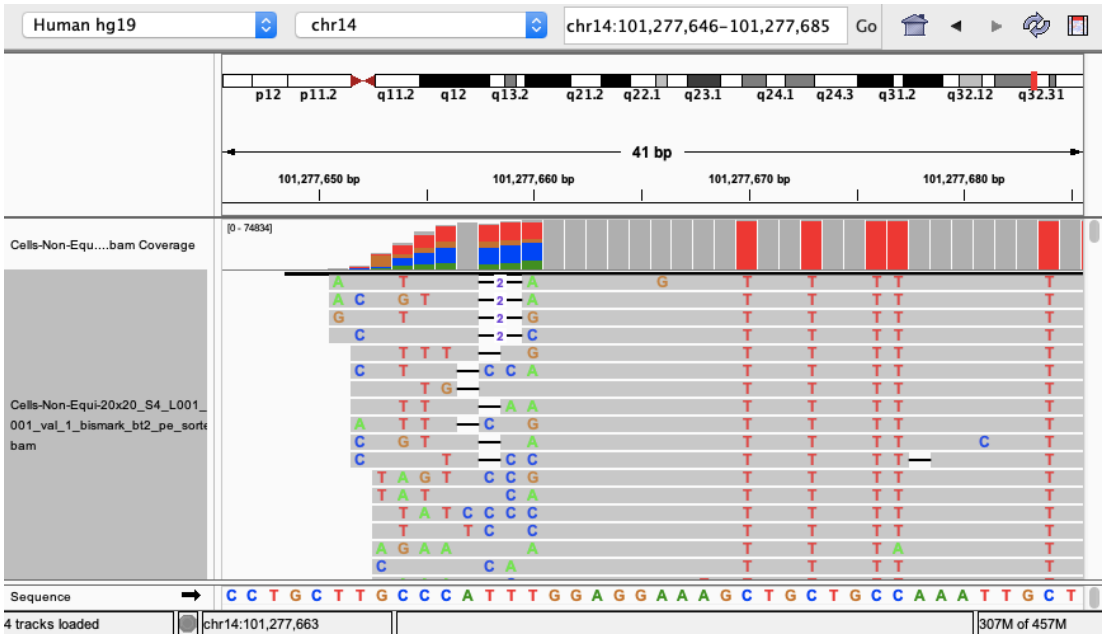
7.1.2.7. Data analysis

FASTQ files obtained after sequencing were converted into BAM files and aligned to the human genome with the Bismark tool (Babraham Institute, UK) for methylation calling of CpGs. When deduplication was performed, the duplicated reads were

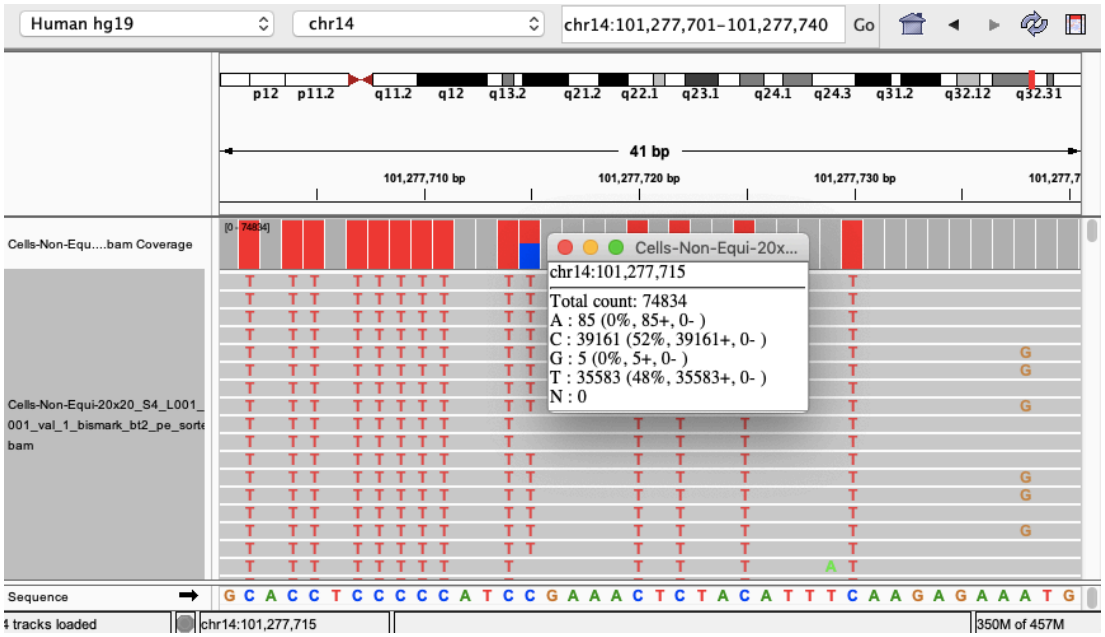
eliminated from analysis prior to genome alignment using UMI-tools (Smith et al., 2017). This analysis was performed by Ph.D. Alejandro Aleman.

The alignment of the DNA sequencing reads to their genomic location was visualised on the Integrated Genome Viewer (IGV) software (Broad Institute, USA). CpGs were identified in each DMR to determine the percentage of methylated cytosines. The presence of a thymine (T) instead of a cytosine (C) in the CpGs was detected as an unmethylated cytosine, whereas the presence of a cytosine was detected as a methylated cytosine. An average was calculated of the methylation percentage of the CpGs included in the targeted region (Figure 55(B)). Additionally, the coverage was determined for each DMR by calculating the average number of reads aligned to each CpGs in the targeted region. Allele frequency was also determined for heterozygous SNPs (Figure 55(C)). Methylation and allele percentages were not included for analysis for sites where less than two reads aligned.

(A)



(B)



(C)

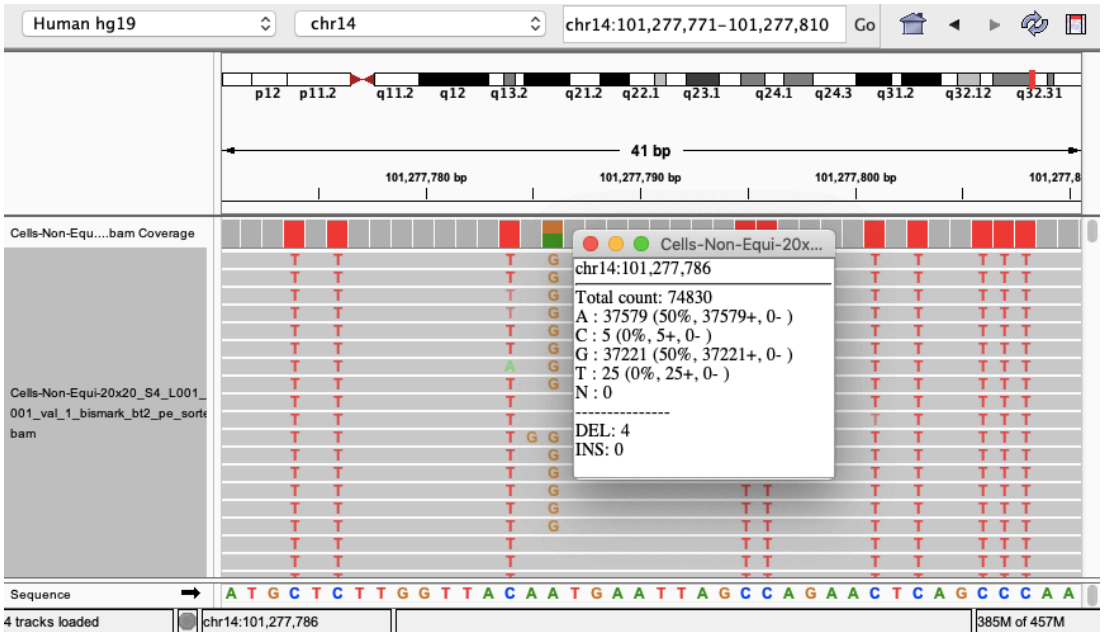


Figure 55. Alignment of the *IG-DMR* DMR to the human genome visualised on the Integrative Genome Viewer (IGV). (A) Eight bp long UMI **(B)** A representative example of a CpGs with 52% of cytosines methylated **(C)** Heterozygous SNP with a frequency A:50%, G:50%. The alignment presented was without previous deduplication.

7.1.2.8. Statistical analysis

A Chi-square or Fisher's exact test with a 95% confidence level was performed to compare the average CpGs methylation percentage of the 46 DMRs between number of

PCR1 cycles. The number of reads mapped for clumps of buccal cells was substantially smaller compared to commercial human gDNA impeding the statistical analyses in this group.

The statistical analyses and descriptive statistics were performed with GraphPad Prism Software. GraphPad Prism Software and Microsoft Excel were used for graph drawing.

7.1.3. Results

7.1.3.1. First phase: Library preparation optimisation with an equimolar primer pool for PCR1

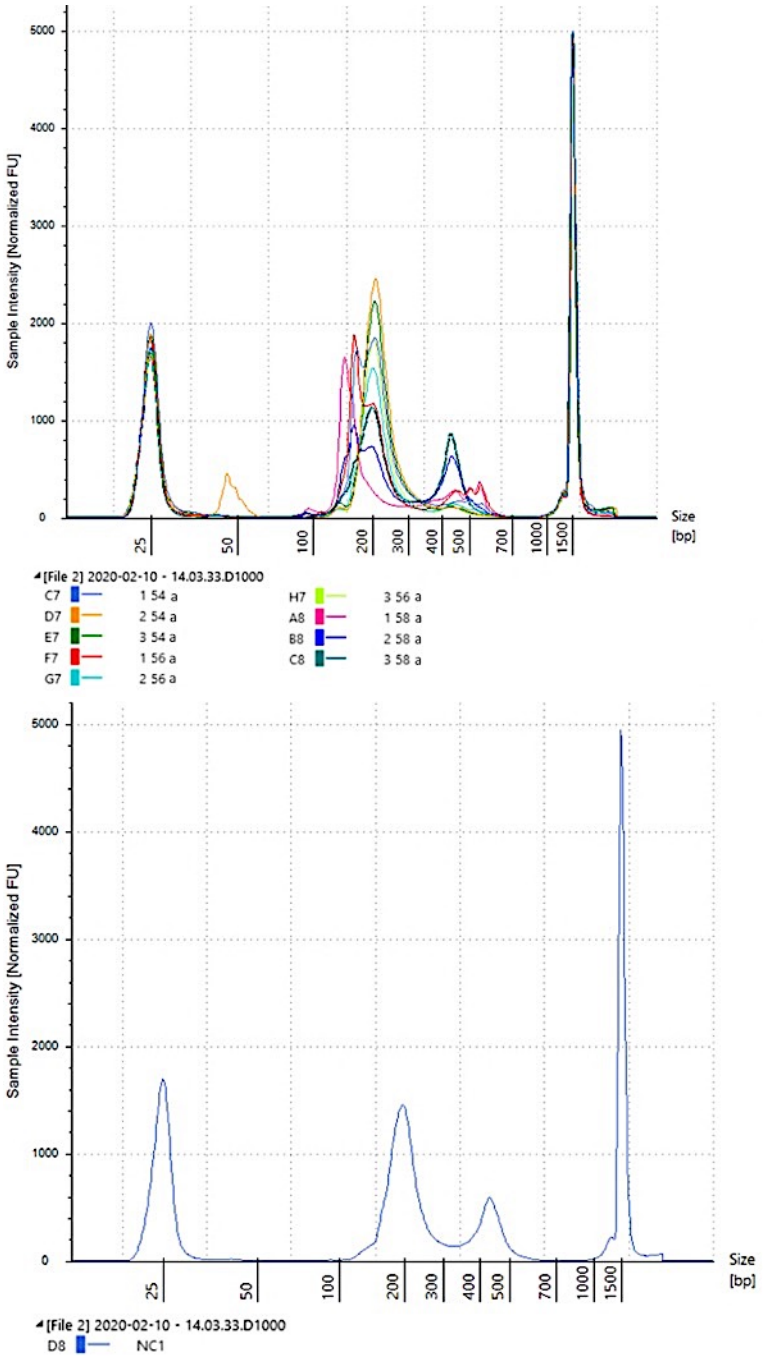
7.1.3.1.1. Primer concentration and annealing temperature gradient for PCR1

Library preparation with the completely unmethylated and bisulfite-converted gDNA (control DNA) was performed with three annealing temperatures (54°C, 56°C and 58°C) and three primer pool concentrations (0.10 µM, 0.20 µM and 0.30 µM) for PCR1 to compare libraries and primer-dimer accumulation after PCR2. The products of PCR1 were cleaned-up with a 1:1 beads:DNA ratio and the products of PCR2 with a 0.7:1 ratio. Eight DMRs (*PEG3, IG-DMR, GRB10, MEG8, ZNF, H19, GNAS* and *SNRPN*) were amplified in singleplex with the products of PCR2 to confirm specific amplification.

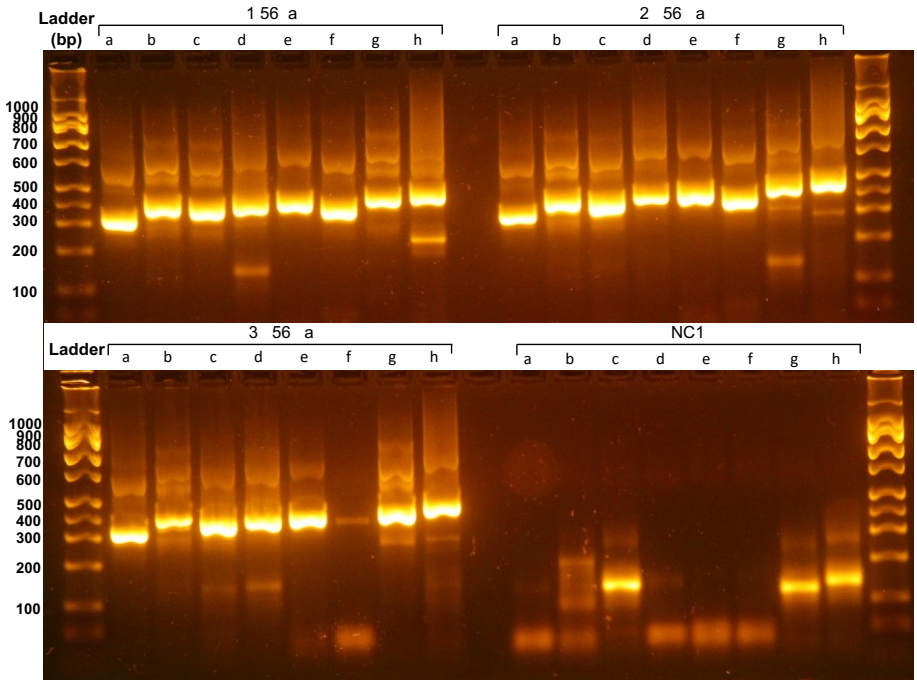
Libraries showed higher concentrations of primer-dimers with lower PCR1 annealing temperatures (54°C) and preferential amplification of targets at higher temperatures (58°C), as seen in the change of sample intensity curve shape. With a PCR1 annealing temperature of 58°C, 0.10 µM concentration of PCR1 primers accumulated the

most primer-dimers (Figure 56(A)). Figure 56(B) confirmed the specific amplification of the eight DMRs after PCR2 clean-up.

(A)



(B)



DNA	PCR primer concentration	PCR annealing temperature
Control DNA (5 ng of completely unmethylated and bisulfite-converted gDNA)	PCR1 (equimolar): **1:** 0.10 µM **3:** 0.30 µM **2:** 0.20 µM	PCR1: 56 °C, 57°C and 58 °C

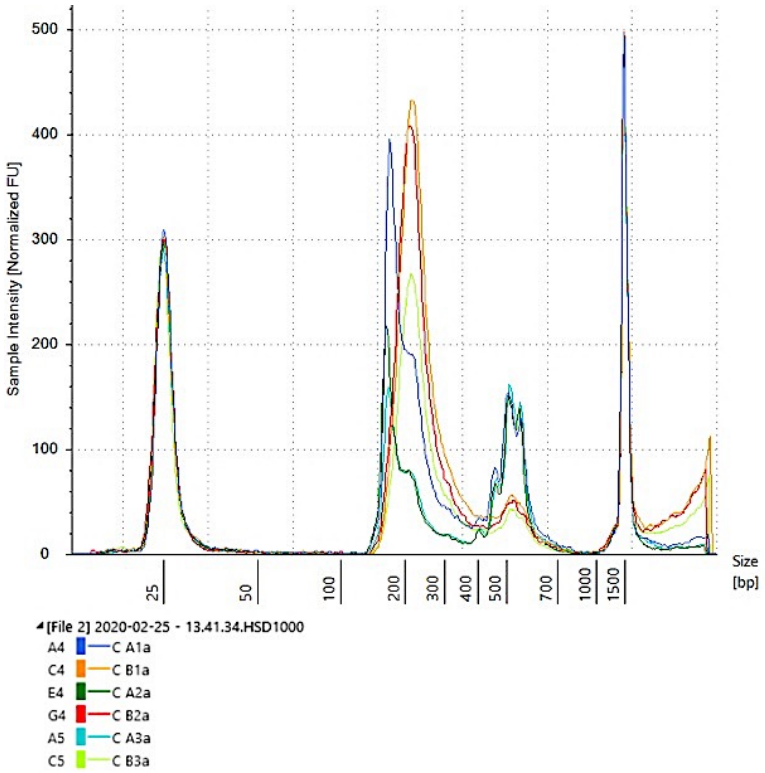
Figure 56. Primer concentration and annealing temperature gradient for PCR1. All experiments were run in duplicates (a and b, b not shown) and two no-template controls were included (NC1 and NC2, NC2 not shown) **(A)** The libraries of replicate a were visualised on the TapeStation before 0.7:1 beads:DNA clean-up. **(B)** Eight DMRs (a=*PEG3*, b=*IG-DMR*, c=*GRB10*, d=*MEG8*, e=*ZNF*, f=*H19*, g=*GNAS* and h=*SNRPN*) were amplified in singleplex from the purified products of PCR2. The bands shown in the singleplex amplifications with the libraries prepared without a template are non-specific products. The singleplexes products were run in a 4% agarose gel electrophoresis at 90 volts for 1 hour and 20 minutes.

7.1.3.1.2. Clean-up optimisation

The study proceeded to optimise the beads:DNA ratios for nucleic acid clean-up after PCR1 and PCR2. From the results of the previous experiment, the primer concentration for PCR1 used was 0.20 µM and an annealing temperature of 56°C.

Six combinations between two ratios for purification after PCR1 (0.85:1 and 1:1) and three after PCR2 (0.70:1, 0.64:1 and 0.60:1) were used. As shown in Figure 57 with

replicate ‘a’, using a ratio of 0.85:1 after PCR1 recovered the highest concentration of libraries, regardless of the ratio used after PCR2. There was a higher degree of primer-dimer carryover with higher ratios, as expected.

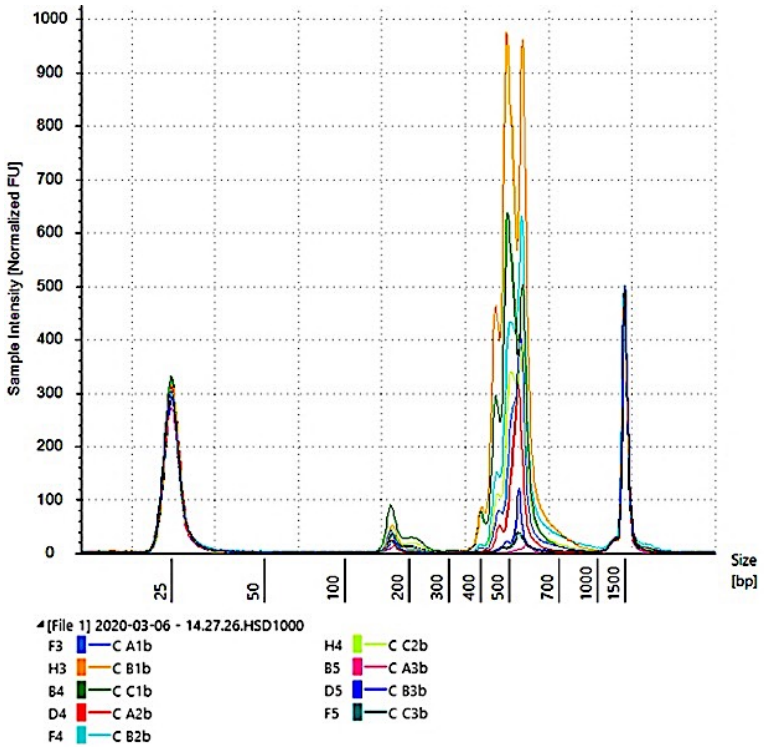


DNA	PCR primer concentration	PCR annealing temperature	Clean-up beads:DNA ratio
Control DNA (5 ng of completely unmethylated and bisulfite-converted gDNA)	PCR1 (equimolar): 0.20 µM	PCR1: 56°C	PCR1: **A:** 0.85:1. **B:** 1:1 PCR2: **1:** 0.70:1 **3:** 0.60:1 **2**: 0.65:1

Figure 57. First experiment for the clean-up optimisation after PCR1 and PCR2. All experiments were run in duplicates (a and b, b not shown) and two no-template controls were included (NC1 and NC2, not shown). ‘C’ stands for ‘cleaned-up after PCR2’.

The following experiment used nine different combinations between three ratios for purification after PCR1 (0.40:1, 0.50:1 and 0.60:1) and three after PCR2 (0.50:1, 0.40:1 and 0.30:1). Replicate ‘b’, in Figure 58, shows that the ratio combination that recovered

the highest concentration of libraries was 0.50:1 for both PCR1 and PCR2. Smaller ratios showed lower degree of primer-dimer carryover compared to the previous experiment.



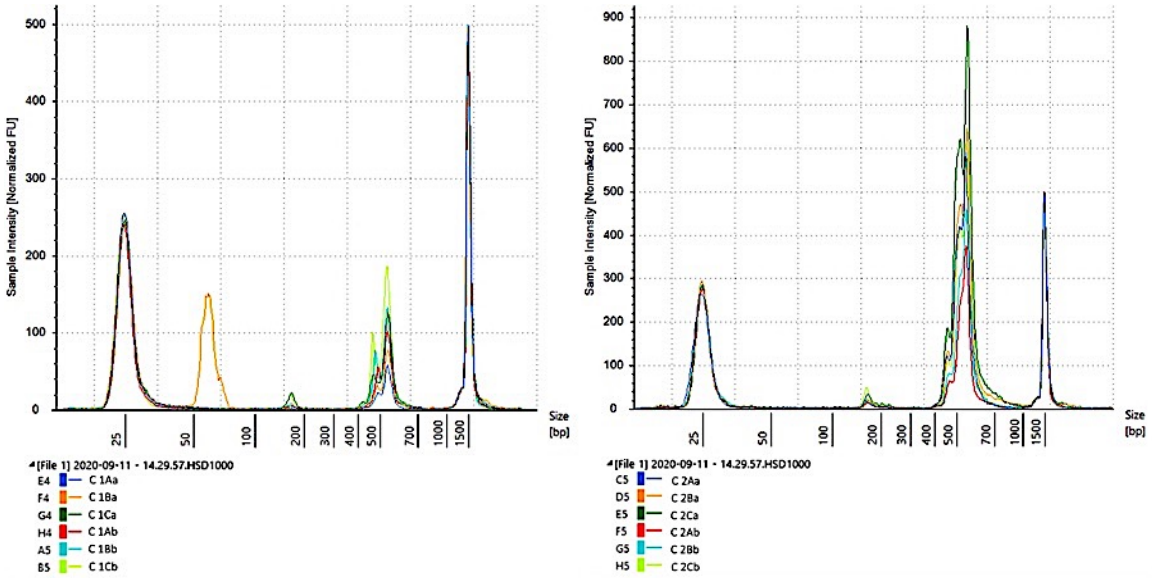
DNA	PCR primer concentration	PCR annealing temperature	Clean-up beads:DNA ratio
Control DNA (5 ng of completely unmethylated and bisulfite-converted gDNA)	PCR1 (equimolar): 0.20 µM	PCR1: 56°C	PCR1: **A:** 0.40:1. **C:** 0.60:1 **B:** 0.50:1 PCR2: **1:** 0.50:1 **3:** 0.30:1 **2**: 0.40:1

Figure 58. Second experiment for the clean-up optimisation after PCR1 and PCR2. All experiments were run in duplicates (a and b, a not shown) and two no-template controls were included (NC1 and NC2, not shown).

7.1.3.1.3. Number of cycles for PCR1 and PCR2 primer concentration

The next set of experiments used two different numbers of cycles for PCR1 (X15 and X20) and three primer concentrations for PCR2 (0.05 µM, 0.10 µM and 0.20 µM) producing six different combinations. Figure 59 shows that X20 cycles for PCR1 and a

primer concentration for PCR2 of 0.20 µM generated the highest libraries' concentrations.



DNA	PCR primer concentration	PCR annealing temperature	PCR number of cycles	Clean-up beads:DNA ratio
Control DNA (5 ng of completely unmethylated and bisulfite-converted gDNA)	PCR1 (equimolar): 0.20 µM PCR2: **A:** 0.05 µM **C:** 0.20 µM **B:** 0.10 µM	PCR1: 56°C	PCR1: **1:** X15 **2**: X20	PCR1**:** 0.50:1 PCR2: 0.50:1

Figure 59. Number of cycles for PCR1 and PCR2 primer concentration. All experiments were run in duplicates (a and b) and two no-template controls were included (NC1 and NC2, not shown).

7.1.3.2. Second phase: Amplification level analysis per target

This study proceeded to sequence the libraries generated using the optimised parameters from the first phase (Table 47).

Table 47. Optimised conditions for library preparation from the first phase of this study. The conditions not indicated in this table used the values from Table 46.

Primer concentration	PCR1 annealing temperature	PCR1 number of cycles	Clean-up beads:DNA ratio
PCR1: 0.20 µM PCR2: 0.20 µM	56°C	X20	PCR1**:** 0.50:1 PCR2: 0.50:1

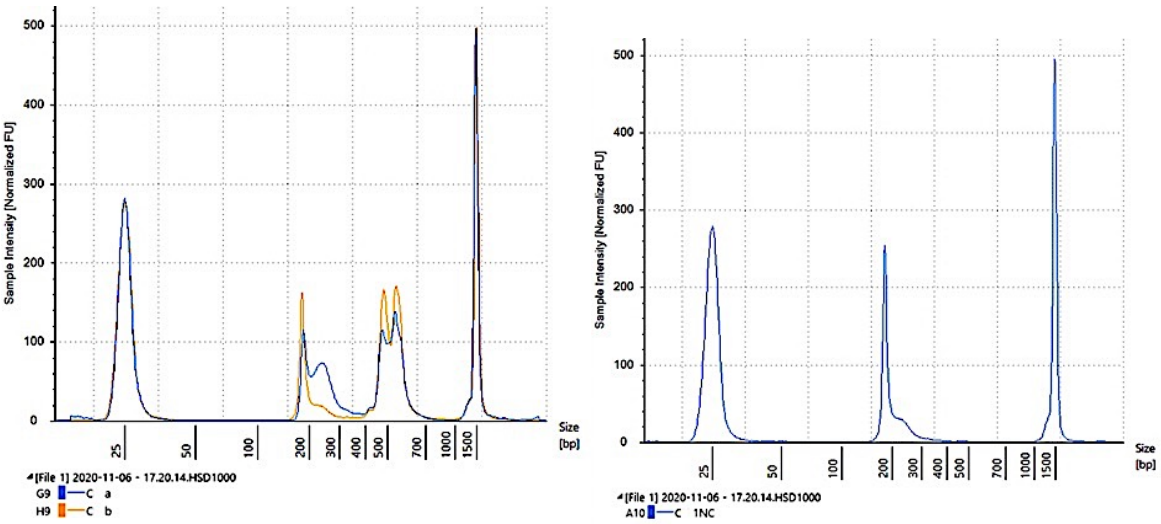
The resulting numbers of mapped reads per target are shown in Table 48. The amplification levels among DMRs varied from approximately 2.50 million reads mapped to 100. In fact, two out of the 46 DMRs, *PEG3* and *ZDBF2*, made up approximately 50% of the total reads mapped. This study proceeded to prepare a non-equimolar primer pool aiming to increase the amplification uniformity between targets during PCR1. Primer pairs that showed the highest number of reads mapped with the equimolar primer pool were included in the non-equimolar primer pool at lower concentrations.

Table 48. Reads mapped per target using an equimolar primer pool and the proportion of each target present in the non-equimolar primer pool for PCR1.

DMR	Number of reads mapped	Percentage of total reads mapped (%)	Proportion in non-equimolar primer pool	DMR	Number of reads mapped	Percentage of total reads mapped (%)	Proportion in non-equimolar primer pool
PEG3	2,464,573	27.37	4.85E-03	*IGF2 3'UTR*	29,866	0.33	1.94E-02
ZDBF2	2,066,006	22.94	9.71E-03	*WDR27*	22,356	0.25	2.43E-02
PEG3_UMI	648,861	7.21	1.46E-02	*KCNQ1OT1*	22,098	0.24	2.43E-02
INPP5F	556,535	6.18	1.46E-02	*IGF1R*	16,736	0.19	2.43E-02
ZNF597	519,956	5.77	1.46E-02	*GNAS-AS1*	14,134	0.16	2.43E-02
MEST	518,276	5.75	1.46E-02	*MAGEL2*	11,636	0.13	2.91E-02
FAM50B	326,401	3.62	1.46E-02	*VTRNA2-1*	11,220	0.12	2.91E-02
HTR5A	214,503	2.38	1.46E-02	*NHP2L1*	8,476	0.09	2.91E-02
GNAS-XL	185,460	2.06	1.46E-02	*GRB10*	7,688	0.08	2.91E-02
GNAS-NESP	165,895	1.84	1.46E-02	*RB1*	7,215	0.08	2.91E-02
SNRPN/SNURF	144,097	1.60	1.46E-02	*DIRAS3 Ex2*	6,797	0.07	2.91E-02
ZNF331 variant *7*	138,203	1.53	1.46E-02	*NAP1L5*	4,297	0.05	2.91E-02
DIRAS3	131,236	1.46	1.46E-02	*H19*	2,589	0.03	2.91E-02
L3MBTL1	108,780	1.21	1.46E-02	*ZNF597/NAA60*	2,491	0.03	2.91E-02
GNAS A/B	98,178	1.09	1.46E-02	*IGF2 DMR0*	2,079	0.02	2.91E-02
MEG8	86,706	0.96	1.94E-02	*MEG3*	1,519	0.02	2.91E-02
NDN	83,326	0.92	1.94E-02	*ZNF331* variants *4-6*	1,459	0.02	2.91E-02
PPIEL	79,871	0.89	1.94E-02	*PEG10*	1,150	0.01	2.91E-02
IG-DMR	79,071	0.88	1.94E-02	*PLAGL1*	928	0.01	2.91E-02
WRB	76,650	0.85	1.94E-02	*SNRPN* variant *4*	228	2.53E-03	2.91E-02
MKRN3	52,127	0.58	1.94E-02	*ERLIN2*	183	2.03E-03	2.91E-02
NNAT	45,725	0.51	1.94E-02	*GPR1-AS*	-	0	2.91E-02
MCTS2P	39,450	0.44	1.94E-02	**Total**	**9,005,031**	**100**	**1**

7.1.3.3. Third phase: Library preparation optimisation with a non-equimolar primer pool for PCR1 and index incorporation in primers for PCR2

The non-equimolar primer pool and the incorporation of indices were tested with 5 ng of control gDNA using the optimised conditions from Table 47. Figure 60 shows the purified libraries after PCR2. When comparing these results to Figure 59, there is a higher degree of primer-dimer carryover, likely due to the presence of a number primer pairs in the non-equimolar primer pool in very high concentrations.



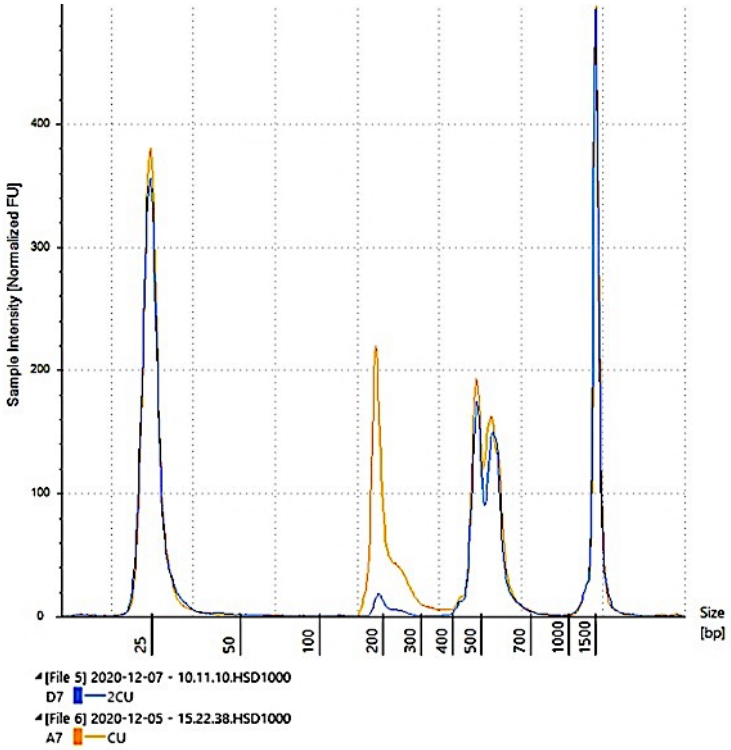
DNA	PCR primer concentration	PCR annealing temperature	PCR number of cycles	Clean-up beads:DNA ratio
Control DNA (5 ng of completely unmethylated and bisulfite-converted gDNA)	PCR1 (non-equimolar): 0.20 µM PCR2: 0.20 µM	PCR1: 56°C	PCR1: X20	PCR1: 0.50:1 PCR2: 0.50:1

Figure 60. Library preparation optimisation with a non-equimolar primer pool for PCR1 and index incorporation in primers for PCR2. All experiments were run in duplicates (a and b) and two no-template controls were included (NC1 and NC2, NC2 not shown).

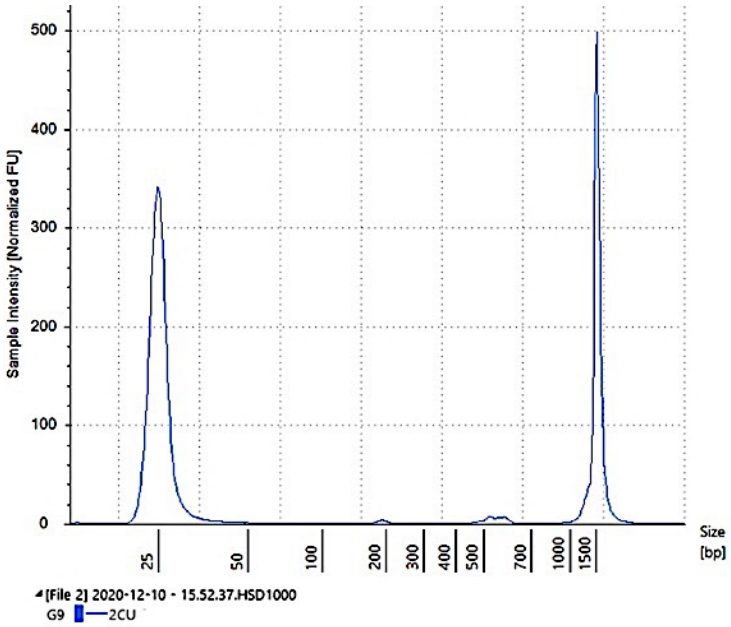
7.1.3.4. Fourth phase: Optimisation of gDNA extraction of clumps of cells and bisulfite conversion

This study tested two different bisulfite conversion kits on the isolated gDNA from clumps of buccal cells followed by library preparation. Both equimolar and non-equimolar primer pools for PCR1 and the primer pool for PCR2 were used at a concentration of 0.20 µM. Library preparations were run using X15 and X20 cycles for PCR1 and PCR2, respectively, followed by X20 and X20 cycles for both PCRs. Due to the high concentrations of primer-dimers produced when using single-cell amounts of DNA, a second clean-up step was performed after PCR2 using a ratio of 0.70:1. The bisulfite conversion with the EpiTect Bisulfite kit failed to produce libraries, but it was possible to visualise them using the Imprint Modification Kit (Figure 61 only shows results for the Imprint Modification Kit).

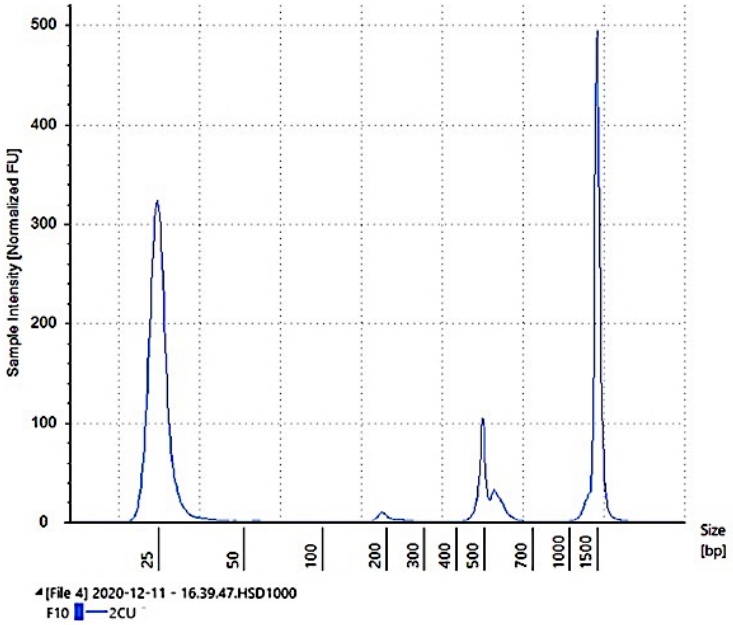
(A) Equimolar primer pool (PCR1), X20 (PCR1) and X20 (PCR2) cycles: Single vs. double clean-up.



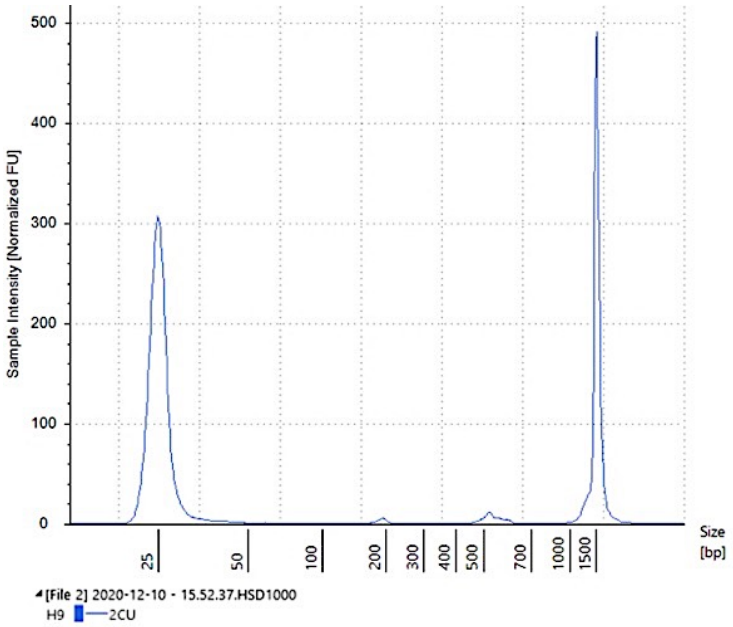
(B) Equimolar primer pool (PCR1), X15 (PCR1) and X20 (PCR2) cycles



(C) Non-equimolar primer pool (PCR1), X20 (PCR1) and X20 (PCR2) cycles



(D) Non-equimolar primer pool (PCR1), X20 (PCR1) and X15 (PCR2) cycles



DNA	Bisulfite conversion kit	PCR primer concentration	PCR annealing temperature	PCR number of cycles	Clean-up beads:DNA ratio
Clumps of eight to nine buccal cells	Imprint Modification Kit	PCR1 (equimolar and non-equimolar): 0.20 µM PCR2: 0.20 µM	PCR1: 56°C	PCR1 and PCR2: X15X20 PCR1 and PCR2: X20X20	PCR1: 0.50:1 PCR2: 0.50:1 and 0.70:1

Figure 61. Library preparation with equimolar and non-equimolar primer pool for PCR1 using clumps of eight to nine buccal cells. Two no-template controls were included (NC1 and NC2, not shown). 'C' stands for 'cleaned-up after PCR2' and 2C for 'second clean-up after PCR2'.

The sample intensity values in Figure 61 demonstrate that the libraries' concentration decreased sharply using X15 cycles for PCR1 for both equimolar and non-equimolar primer pools compared to X20 and that the non-equimolar primer pool produced lower concentration of libraries. Changes of the sample intensity curve shape indicate differences in amplification level patterns of targets between primer pools.

Table 49 sets out the library preparation conditions used for the following phase of this study. The aim was to use a low number of cycles for PCR1 to mitigate amplification bias and so, despite the results in the fourth phase, two cycle numbers for PCR1 were used for all library preparations (X15 and X20) to compare the sequencing results between them.

Table 49. Optimised conditions for library preparation from the fourth phase of this study. The conditions not indicated in this table used the values from Table 46.

Primer concentration (equimolar and non-equimolar)	PCR1 annealing temperature	PCR1 number of cycles	Clean-up beads:DNA ratio
PCR1: 0.20 µM PCR2: 0.20 µM	56°C	PCR1 and PCR2: X15X20 PCR1 and PCR2: X20X20	PCR1**:** 0.50:1 PCR2: 0.50:1 and 0.70:1

7.1.3.5. Fifth phase: Methylation level analysis of the targeted DMRs

Libraries produced from 100 ng of bisulfite-converted human male commercial gDNA (equimolar primer pool only) and bisulfite-converted clumps of buccal cells (equimolar and non-equimolar primer pools) with the Imprint Modification Kit were sequenced, mapped to the reference genome and analysed for average CpGs methylation percentage.

7.1.3.5.1. Methylation level analysis of CpGs using 100 ng of commercial gDNA

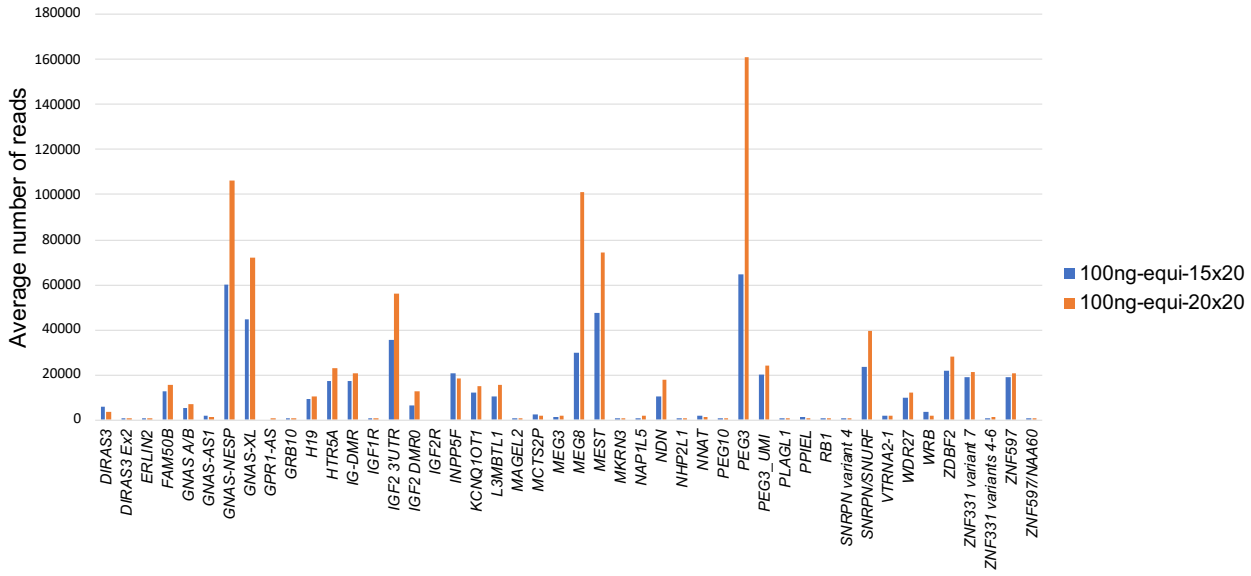


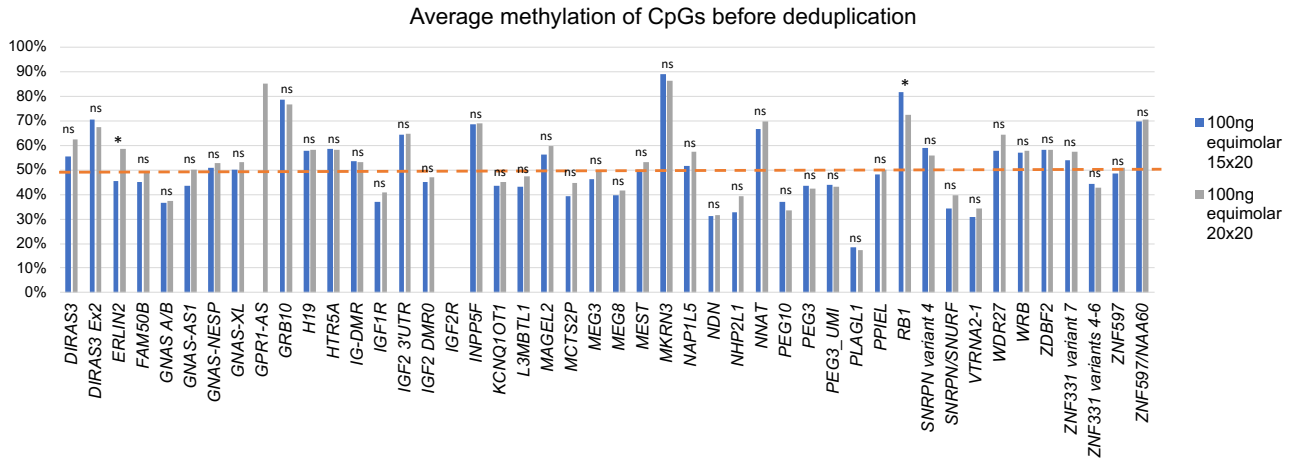
Figure 62. Average number of reads mapped to each CpGs per targeted DMR using 100ng of commercial gDNA. The standard deviation of the average number of reads mapped between all the DMR (excluding *IGF2R*) was 16,207 for X15X20 number of cycles and 33,586 for X20X20 number of cycles.

Using X15 cycles for PCR1 showed higher uniformity in the number of reads mapped between DMRs compared to X20 (Figure 62). Only *IGF2R* failed to map any reads with 100 ng of commercial gDNA.

Before read deduplication, the average methylation percentage of CpGs was maintained when the number of cycles for PCR1 was increased, except for *ERLIN2* and *RB1* ($p<0.05$). Even though methylation levels vary among DMRs, the observed trend approximates to 50% (ranging from 17% to 86%). Some outliers include *MKRN3* (86% and 89% average CpGs methylation for X15 and X20 PCR1 cycles, respectively), *RB1* (72% and 82%), and on the opposite extreme, *PLAGL1* (17% and 18%). *GPR1-AS* failed to produce any reads for X15 PCR1 cycles and produced only three for X20 Figure 63 (A).

In contrast to before deduplication, the methylation levels do not remain constant when the number of PCR1 cycles increase after deduplication (Figure 63(B)). Differences in CpGs methylation were significant ($p<0.05$) for 26 of the 46 targeted DMRs. Overall, average CpGs methylation highly fluctuated between DMRs and for many, these values differ appreciably from the expected 50% (ranging from 2% to 98%).

(A)



(B)

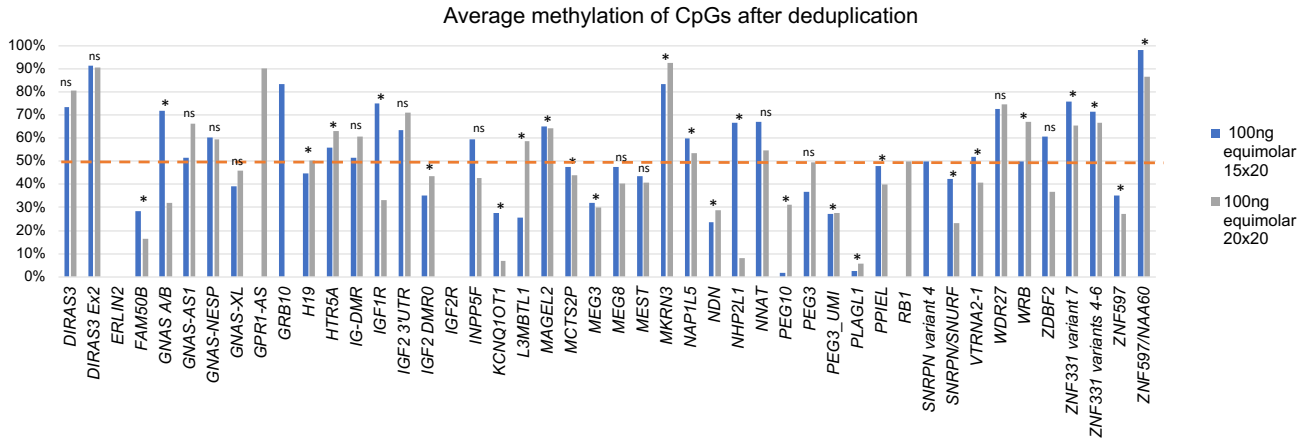


Figure 63. Average methylation percentage of CpGs for the 46 DMRs using 100 ng of commercial gDNA before and after deduplication. The standard deviation between the average methylation of CpGs per DMR was 14.34% for X15X20 number of cycles and 14% for X20X20 number of cycles before deduplication. After duplicate removal, the standard deviation was 21.72% and 22.76%, respectively. The dashed line indicates the expected methylation percentage (50%). Absent bars indicate a lack of data due to low read coverage. The statistical analysis was performed comparing individual CpGs methylation percentages using a Chi-square test or a Fisher's exact test depending on the number of CpGs in the DMR. '*' denotes a level of significance P <0.05; ns, not significant.

The average CpGs methylation percentages differ appreciably before and after deduplication for some DMRs. For instance, *ZNF597/NAA60* showed values of 70% and 71% for X15 and X20 PCR1 cycles, respectively, before deduplication, which increased to 98% and 87% after removing duplicated reads.

The experimental design initially included seven SNPs as controls. Seven additional heterozygous SNPs were identified in the commercial gDNA and the frequencies of the fourteen were analysed. SNPs that showed to be homozygous and SNPs that had low coverage (less than two reads aligned) were not included (Table 50). **Appendix D** contains information about the additional SNPs.

Table 50 compares the read coverage, the average CpGs methylation and heterozygous SNP allele frequency before and after deduplication. According to the bioinformatic methodology for read duplicate removal, the number of reads mapped decreased remarkably after deduplication. For these 10 DMRs, the maximum average number of reads observed after deduplication was 24 and the minimum was two. For instance, there was a 99.97% decrease in the average number of reads mapped for *PEG3*. Consequently, the coverage was strikingly reduced for both CpGs and SNPs. In the case of *MAGEL2* and *MKRN3*, the number of reads aligned to the SNP after deduplication was less than two making it impossible to obtain the frequency data. **Appendix D** shows the analysis results of all DMRs in the panel for both the commercial gDNA and clumps of cells.

Table 50. Comparison of DMRs average number of reads aligned, average CpGs methylation levels and heterozygous SNP corresponding allele frequency before and after deduplication for 100 ng of commercial gDNA. '-' denotes not enough coverage for analysis.

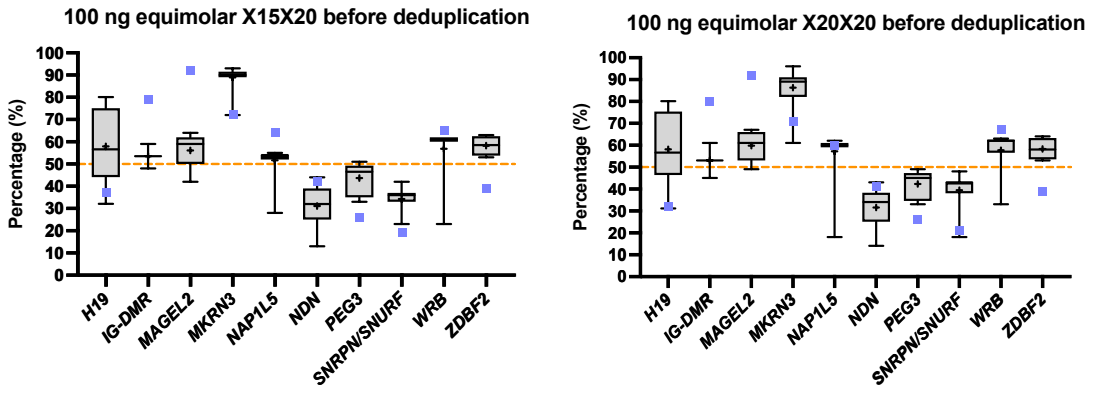
DMR	Before deduplication				After deduplication			
	100 ng equimolar primer pool							
	15x20 cycles		20x20 cycles		15x20 cycles		20x20 cycles	
	Average No. of reads	Average CpGs meth.	Average No. of reads	Average CpGs meth.	Average No. of reads	Average CpGs meth.	Average No. of reads	Average CpGs meth.
H19	9,503	58%	10,245	58%	8	45%	13	50%
IG-DMR	17,390	54%	20,830	53%	8	52%	14	61%
MAGEL2	296	56%	159	60%	5	65%	2	64%
MKRN3	994	89%	641	86%	6	83%	5	93%
NAP1L5	987	52%	1,694	57%	5	60%	6	54%
NDN	10,471	31%	18,027	32%	12	23%	24	29%
PEG3	64,739	44%	161,092	42%	19	37%	19	49%
SNRPN/SNURF	23,788	34%	39,525	39%	13	42%	22	23%
WRB	3,648	57%	2,004	58%	4	50%	9	67%
ZDBF2	21,894	58%	28,411	58%	12	61%	12	37%
DMR	**No. of reads**	**SNP Allele freq.**	**No. of reads**	**SNP Allele freq.**	**No. of reads**	**SNP Allele freq.**	**No. of reads**	**SNP Allele freq.**
H19	9,503	37%	10,245	32%	8	18%	13	19%
IG-DMR	17,390	79%	20,830	80%	8	60%	14	69%
MAGEL2	296	92%	159	92%	5	60%	2	-
MKRN3	994	72%	641	71%	6	-	5	-
NAP1L5	987	64%	1,694	60%	5	80%	6	100%
NDN	10,471	42%	18,027	41%	12	42%	24	31%
PEG3	64,739	26%	161,092	26%	19	10%	19	37%
SNRPN/SNURF	23,788	19%	39,525	21%	13	20%	22	4%
WRB	3,648	65%	2,004	67%	4	50%	9	56%
ZDBF2	21,894	39%	28,411	39%	12	86%	12	29%

The CpGs average methylation and the corresponding allele frequency before deduplication for the DMRs in Table 50 are plotted in Figure 64(A). Although both data sets follow the same pattern around the expected 50%, the difference in percentage between CpGs methylation mean and SNP allele frequency is notable. For instance, there is a difference of 36% and 32% between both data points for *MAGEL2* when the PCR1 cycles are X15 and X20, respectively. Although the low read coverage could be a factor, the difference of both data points for *IG-DMR* is 25% and the coverage for this locus is approximately 58 times higher than the coverage for *MAGEL2*.

After deduplication, the difference between the average CpGs methylation and SNP frequency was expected to decrease. Moreover, both averages were expected to be closer to 50%. However, Figure 64(B) shows that the SNP frequencies approximation to 50% did not improve after deduplication and that the difference in percentage between CpGs methylation mean and SNP allele frequency did not decrease for all DMRs. In fact, the difference increased for some instances such as *NAP1L5.*

In addition, unlike before deduplication, there was a change in SNP frequency for some DMRs from X15 to X20 cycles for PCR1, such as *ZDBF2* (consistent with the CpGs methylation results of Figure 63).

(A)



(B)

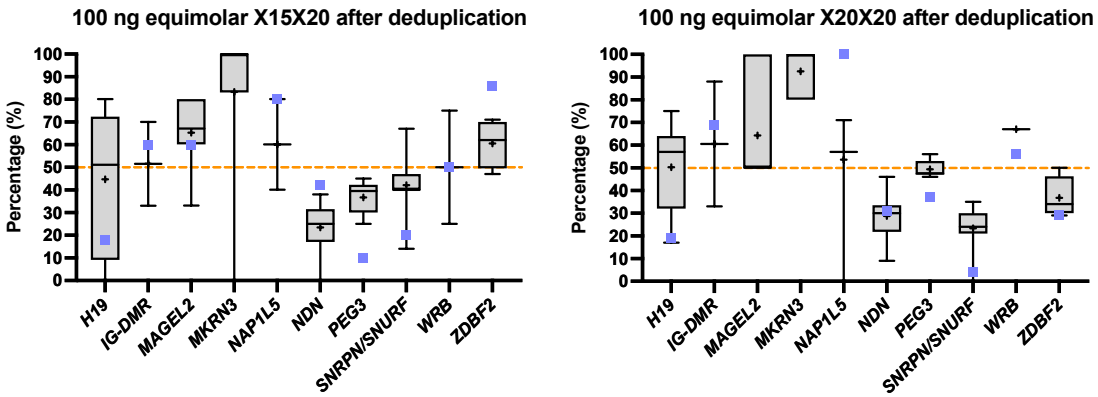


Figure 64. Comparison of DMRs average CpGs methylation levels and heterozygous SNP corresponding allele frequency before and after deduplication for 100 ng of commercial gDNA. The CpGs methylation distribution is presented with Box & Whiskers plots and the SNP frequency is indicated with the purple boxes. The orange dashed line indicates the expected methylation percentage (50%). Absent data points indicate a lack of information due to low read coverage. '+' denotes the CpGs methylation percentage mean in the plot.

7.1.3.5.2. Methylation level analysis of CpGs using clumps of buccal cells

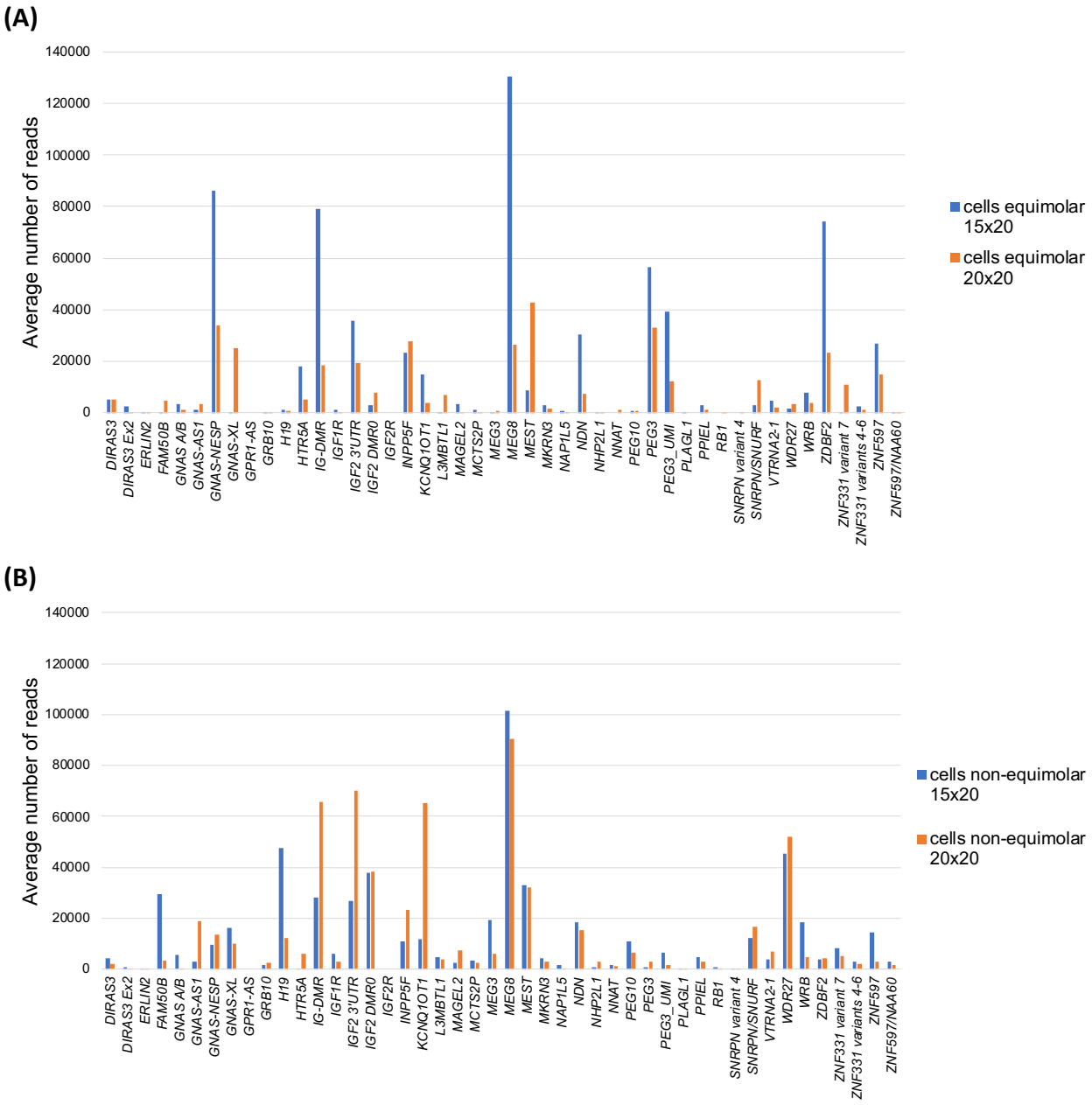


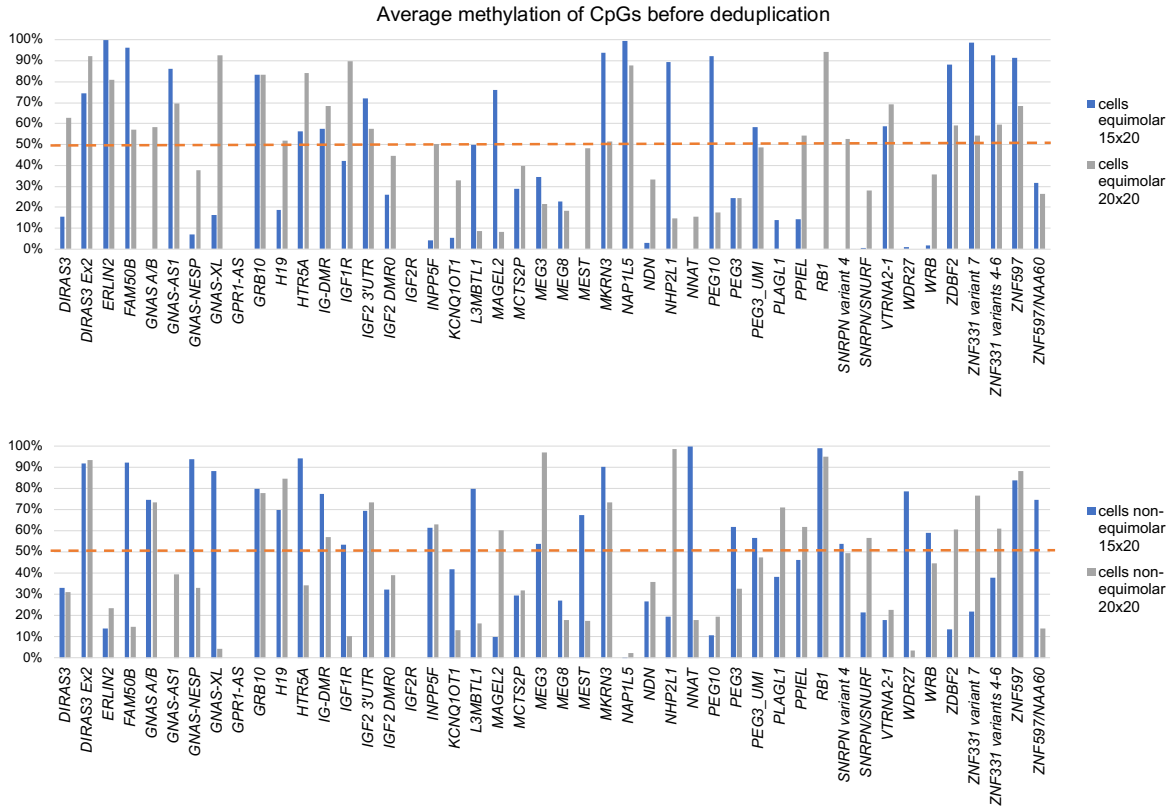
Figure 65. Average number of reads mapped to each CpGs per targeted DMR using clumps of buccal cells. **(A)** Equimolar primer pool. The standard deviation between the average number of reads mapped per targeted DMRs (excluding *IGF2R* and *GPR1-AS*) was 28,496 for X15 number of PCR1 cycles and 11,141 for X20 number of PCR1 cycles. **(B)** Non-equimolar primer pool. The standard deviation between the average number of reads mapped per targeted DMRs (excluding *IGF2R* and *GPR1-AS*) was 18,485 for X15 number of PCR1 cycles and 21,918 for X20 number of PCR1 cycles.

The average number of reads mapped to each CpGs per target was compared between libraries from clumps of cells produced from equimolar and non-equimolar

primer pools without deduplication (Figure 65). Only *IGF2R* and *GPR1-AS* failed to map reads. Clumps of cells amplified with the equimolar primer pool with X20 PCR1 cycles showed the highest uniformity in the number of reads mapped between DMRs since it produced the lowest standard deviation. Unlike with 100 ng of commercial gDNA, where the number of reads mapped increased when increasing PCR1 cycles from X15 to X20 for most targets (27 out of 45), increasing the number of cycles for PCR1 with clumps of cells increased the number of reads mapped for only 18 out of 44 targets using an equimolar primer pool and for 17 out of 44 using a non-equimolar primer pool.

Altogether, the plots from Figure 66 show that the data from single-cell levels of DNA distance remarkably from the expected 50% of CpGs methylation and heterozygous SNP frequency. In addition, the uniformity in the average CpGs methylation per DMR between X15 and X20 PCR1 cycles is lower compared to the results from 100 ng of commercial gDNA. Nevertheless, using a non-equimolar primer pool shows a better visual approximation to 50% and a better uniformity between different PCR1 number of cycles. For instance, with an equimolar primer pool, *WRB* shows close to 0% of methylated CpGs when X15 PCR1 cycles are used. With X20 cycles, the percentage increases to 36%. On the other hand, these percentages are 59% and 44%, respectively, with a non-equimolar primer pool. The same effect is observed for *RB1*. However, this does not apply across all samples. *NAP1L5* shows 99% and 88% of average CpGs methylation for X15 and X20 PCR1 cycles, respectively, using an equimolar primer pool and these values decrease to 0% and 2% with a non-equimolar primer pool.

(A)


Average methylation of CpGs before deduplication
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
DIRAS3
DIRAS3 Ex2
ERLIN2
FAM50B
GNAS A/B
GNAS-AS1
GNAS-NESP
GNAS-XL
GPR1-AS
GRB10
H19
HTR5A
IG-DMR
IGF1R
IGF2 3'UTR
IGF2 DMR0
IGF2R
INPP5F
KCNQ1OT1
L3MBTL1
MAGEL2
MCTS2P
MEG3
MEG8
MEST
MKRN3
NAP1L5
NDN
NHP2L1
NNAT
PEG10
PEG3
PEG3_UMI
PLAGL1
PPIEL
RB1
SNRPN variant 4
SNRPN/SNURF
VTRNA2-1
WDR27
WRB
ZDBF2
ZNF331 variant 7
ZNF331 variants 4-6
ZNF597
ZNF597/NAA60
cells equimolar 15x20
cells equimolar 20x20
cells non-equimolar 15x20
cells non-equimolar 20x20

(B)

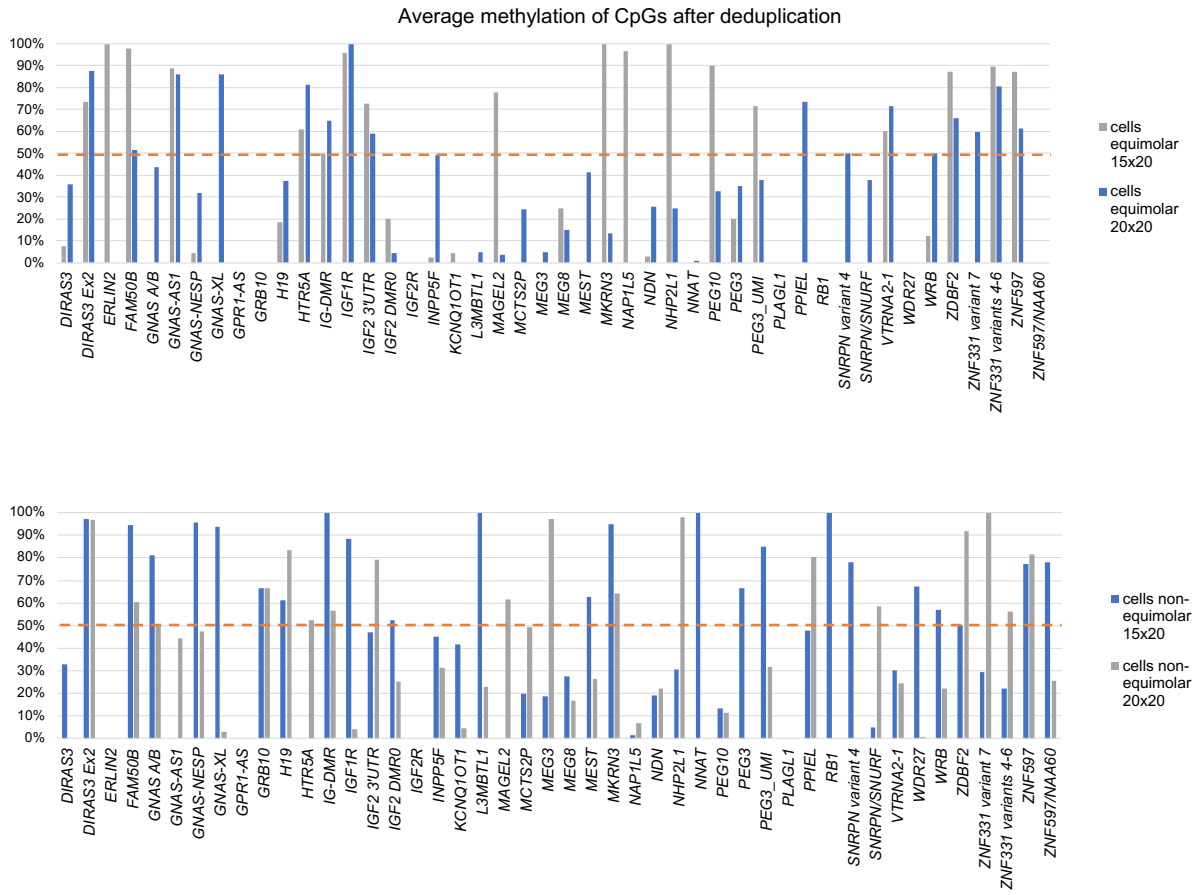


Figure 66. Average methylation percentage of CpG for the 46 DMRs using clumps of buccal cells. (A) Before deduplication, the standard deviation between the average methylation of CpGs per DMR with the equimolar primer pool was 36.40% for x15X20 number of cycles and 25.70% for X20X20 number of cycles. The standard deviation for the non-equimolar primer pool was 30.10 and 28.68%, respectively. **(B)** After duplicate removal, the standard deviation between the average methylation of CpGs per DMR with the equimolar primer pool was 40.69% for X15X20 number of cycles and 28.95% for X20X20 number of cycles. The standard deviation for the non-equimolar primer pool was 31.88 and 32.21%, respectively. The number of reads mapped for clumps of buccal cells was substantially smaller compared to commercial human gDNA impeding the statistical significance analyses between the differences in the average CpGs methylation percentage in this group.

The visual distancing from 50% in the methylation levels increases further after read deduplication, as seen with 100 ng of commercial gDNA. This is particularly evident for the equimolar primer pool. The methylation levels decrease abruptly for a significant percentage of DMRs, some reducing their percentages to or close to 0. **Appendix D** shows that the number of reads mapped decrease exceptionally, as seen with 100 ng of

gDNA. The striking reduction in the number of reads is true under all experimental conditions.

Figure 67 compares the average CpGs methylation percentage for seven DMRs with the heterozygous SNP frequency of the corresponding allele for the clumps of cells. There are a significant number of data points missing from the plots due to the lack of coverage after deduplication.

Figure 67(A) shows that before deduplication, values from the non-equimolar primer pool and X20 cycles for PCR1 approximate 50%. This is also true for after deduplication (Figure 67(B)), although the trend is less clear and some data is missing due to low coverage. Interestingly, the difference between the resulting CpGs methylation percentage mean and SNP frequency, when using X20 PCR1 cycles and a non-equimolar primer pool, decreases when compared to 100 ng of commercial gDNA. However, deduplication does not decrease the difference between the two values further for all DMRs.

The SNP frequency for *NAP1L5* shows the same lack of uniformity between number of PCR cycles and primer pool composition as the average CpGs methylation. According to **Appendix D**, the number of reads mapped for *NAP1L5* greatly varies between X15 and X20 PCR1 cycles (584 and 26, respectively) and equimolar and non-equimolar primer pools (1,626 and 248, respectively).

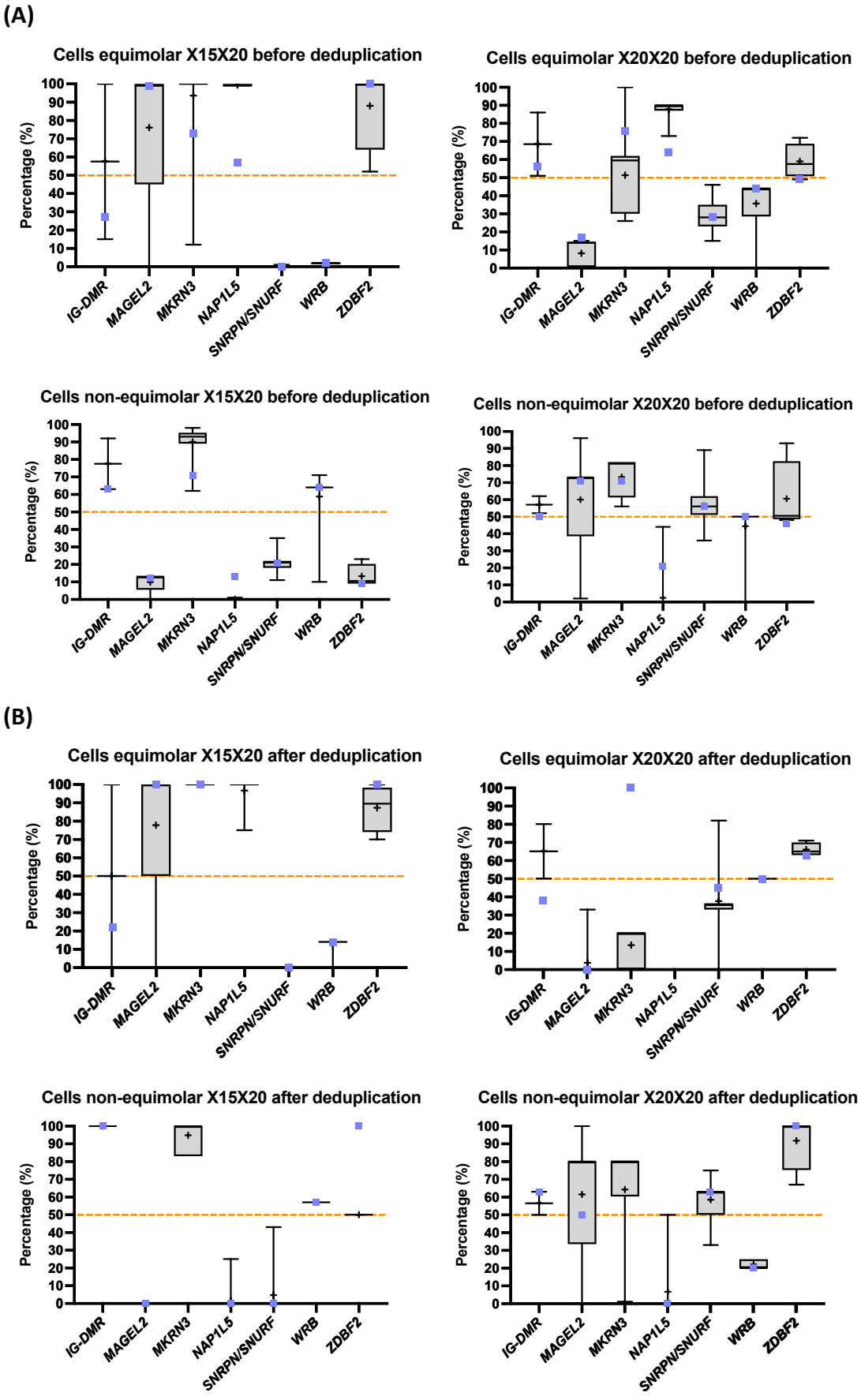


Figure 67. Comparison of DMRs average CpGs methylation levels and heterozygous SNP corresponding allele frequency before and after deduplication for clumps of cells.

7.1.4. Discussion

Epidemiological studies suggest a correlation between ART-conceived children and an increased incidence of rare imprinting disorders (Cortessis et al., 2018). However, robust methodologies and appropriate controls to study the methylation levels of imprinted genes are lacking. This study aimed to develop a targeted BS approach to studying the methylation levels of 46 DMRs. UMIs were used to remove duplicate reads produced by PCR that contribute to the distortion of the true biological frequency of alleles.

7.1.4.1. Amplification uniformity of the 46 targets after bisulfite conversion

PCR conditions were extensively optimised during the first stages of this study. Multiplexing PCR has become a convenient tool in both the clinical and research settings that has enhanced diagnostic capacity and saved time and resources. However, as opposed to conventional PCR, multiplex PCR poses a variety of difficulties, such as the preferential amplification of certain targets (Polz and Cavanaugh, 1998). In a BS application, the guanine and cytosine (GC) content of the DNA is greatly reduced after bisulfite treatment. On one hand, the primer specificity reduces with the resulting low-complexity DNA, while on the other hand, the likelihood of primer-dimer formation increases (Xi and Li, 2009). Therefore, the technical challenge of multiplex PCR increases, requiring diligence during primer and PCR design.

7.1.4.1.1. Individual primer concentration adjustment in the pool

Although the sequencing of 100 ng of commercial gDNA with an equimolar primer pool for PCR1 produced libraries for 45 out of the 46 targets, there was high variation between targets where the number of reads mapped from two DMRs constituted half of the total reads mapped to the reference genome. Aiming to mitigate this skew, a non-equimolar primer pool for PCR1 was used. However, unexpectedly, the uniformity in the number of mapped reads between DMRs did not increase using a non-equimolar primer pool. According to Markoulatos et al. (2002), the amount of primer in a pool for multiplex PCR can be optimised initially by employing equimolar primer concentrations and, if there is high variation in target amplification, the concentrations of primers that showed template underrepresentation can be increased empirically. Carlson et al. (2013) minimised template amplification bias from a pool targeting 56 loci with primer titration and demonstrated that increasing the concentration of individual primers also increased template representation for most instances that were initially underrepresented. At the same time, the concentration of primers with target overrepresentation was decreased. However, the authors also showed that for some primers, amplification did not improve and primers had to be redesigned. Further optimisation experiments for this protocol could include primer titration to modify primer concentration with a stringent method to replace empirical adjustment. This would also help to identify primers that need redesigning, although, due to the size of the DMRs panel, this would be time consuming.

7.1.4.1.2. PCR1 number of cycles

A key variable in the optimisation of the present methodology was the number of PCR cycles, particularly for PCR1. An increase in the number of PCR cycles increases the

overrepresentation of certain sequences, which can greatly affect the downstream DNA analysis for CpGs methylation percentages of DMRs (Olova et al., 2018). Furthermore, the present experimental design attaches the UMI to the target sequence during PCR1, making these amplicons prone to bias, which the UMIs will not be able to identify in subsequent bioinformatic analysis. The aim, therefore, was to maintain a low number of cycles for PCR1 without significantly compromising the template concentration for PCR2. However, a lower number of PCR1 cycles did not show higher uniformity among DMRs mapped reads using 100 ng of gDNA, which suggests that DMR overrepresentation was not exacerbated with a higher number of cycles. Nevertheless, targeted loci amplification uniformity is not informative regarding allele amplification or PCR duplicate bias.

7.1.4.1.3. DNA bisulfite treatment

Another important element that may have contributed to amplification bias was BS artefacts caused by incomplete cytosine conversion during bisulfite treatment (Olova et al., 2018). Bisulfite treatment can also be associated with degradation and loss of DNA. Bisulfite conversion efficiency can be assessed by designing PCR primers free of cytosines in addition to a bisulfite-converted DNA control (Hong and Shin, 2021). BS library preparations can be, therefore, spiked-in with, for instance, phage lambda bisulfite-converted DNA control. The primer pool would contain the cytosine-free primers designed to target a region of the phage lambda genome so that the bisulfite conversion efficiency could be assessed in each BS run.

Future library preparation optimisation experiments should also include iterations using the same sources of DNA (same lot number of commercial gDNA and clumps of

cells from the same individual) and the same PCR conditions for library preparation to study, with consistency, differences in amplification uniformity of DMRs.

7.1.4.2. Methylation analysis of DMRs

7.1.4.2.1. CpGs average methylation before read deduplication

The mean CpGs methylation approximated to the expected 50% using 100 ng of commercial gDNA. There were no significant differences between using X15 and X20 number of cycles for PCR1 for the vast majority of targets. Denomme et al. (2016) studied the average CpGs methylation percentage of *SNRPN*, *H19* and *KCNQ1OT1* in euploid oocyte donor blastocysts (used as control) employing a methodology that included bisulfite treatment, nested PCR and clonal sequencing. They found an average of CpGs methylation of 53.70%, 47.30% and 50.60%, respectively. The present study found an average of 57%, 58% and 44%, respectively. The slight differences between CpGs average methylation percentage can be due to differences in molecular analysis, the type of samples and epigenetic heterogeneity. Methylome sequencing studies in human preimplantation embryos have found dynamics in the CpGs methylation percentages of gDMRs during development. Embryos at the 2-cell stage to the morula stage present hypomethylation of *PEG10*. However, normal levels are established post-implantation (Zhu et al., 2018). This same pattern has been discovered in gDMRs of animal models, such as the pig (Ivanova et al., 2020).

There are limitations to comparing the overall results of DMRs CpGs methylation in this study with others. This is due to the differences in the panel size studied between

this chapter and the existing publications, which have mostly focused on *SNRPN, H19* and *KCNQ1OT1* due to their vital role in growth and development.

Contrary to commercial gDNA, there was a significant variation in the average CpGs methylation between PCR1 number of cycles in clumps of buccal cells. There was variation too between equimolar and non-equimolar primer pools. This is indicative of significant amplification bias. Very low starting template DNA in multiplex PCR can result in unbalanced amplification of certain loci or alleles (Markoulatos et al., 2002). This project increased the number of PCR1 cycles from X15 to X20 aiming to mitigate the problem of low DNA yield, which could have exacerbated any potential bias (Islam et al., 2014). However, the variation between CpGs methylation percentages within a DMR in clumps of cells improved with X20 PCR1 cycles, likely due to the increase in read coverage compared to X15. Although, the improvement was marginal. An additional control that could be included in further optimisation experiments is a higher input DNA, isolated, for instance, from venous blood from the same individual from which the buccal samples are obtained, to determine with accuracy the effect on of low DNA inputs in amplification bias.

7.1.4.2.2. Heterozygous SNPs frequency as a control for allele amplification bias

BS can result in significant bias that can hamper downstream analyses. This poses a significant challenge to the study of true biological methylation levels in asymmetrical methylation when methylated or unmethylated loci are preferentially amplified during PCR (Olova et al., 2018). An approach to overcome this challenge is the use of heterozygous SNPs as a control for allele amplification bias. Heterozygous SNPs are inherited in a parent-of-origin manner like genomic imprinting. Thus, both imprinted

gene methylation marks and heterozygous SNPs are inherited along the respective parental allele. Therefore, it is possible to detect the overrepresentation of one of the alleles after PCR with the heterozygous SNP frequency. In this study, heterozygous SNPs were studied in 10 DMRs. The CpGs methylation average and the heterozygous SNP frequency employing 100 ng of commercial gDNA remained consistent with different PCR1 number of cycles. For instance, the *WRB* gene is maternally imprinted, meaning the CpGs in the DMRs of the maternal allele are methylated. The average percentage methylation was 57% and the frequency of the corresponding allele (G/T) was 65% (T was 35%). The frequency of the G allele then confirms that the maternal allele is overrepresented. However, the average CpGs methylation is slightly lower, suggesting a slight biological DMR hypomethylation. *H19*, a paternally imprinted gene, on the other hand, showed an average CpG methylation percentage of 58% and the frequency of the corresponding allele (A/G) was 37% (G was 63%). Thus, the paternal allele is underrepresented and the observed hypermethylation is biologically true and the A allele frequency suggests the hypermethylation is higher than 58%.

Interestingly, clumps of buccal cells, when using X20 cycles for PCR1 and a non-equimolar primer pool, showed the lowest difference between average CpGs methylation percentage and heterozygous SNP frequency values. Thus, one can speculate that the CpGs methylation of these 10 DMRs is closer to 50%, compared to commercial gDNA. It is important to mention that both clumps of cells and commercial gDNA came from healthy individuals. Nevertheless, the results of clumps of cells also showed a high variation in the CpGs methylation within a DMR and low read coverage, thus, suggesting results must be interpreted with caution.

One major limitation in human imprinting studies is the high degree of natural epigenetic heterogeneity between individuals. In fact, epigenetic changes are more frequent than genetic changes and allele-specific methylation can be shaped by environment and stochastic factors (Potabattula et al., 2018). Epigenetic heterogeneity at imprinted loci has been observed in normal populations at the level of allele-specific expression of imprinted genes (Sakatani et al., 2001) and at the level of DNA methylation of DMRs (El Hajj et al., 2017). This variation makes the determination of 'normal' thresholds of methylation levels difficult. Nonetheless, the use of appropriate controls and large sample sizes can help in the interpretation of epigenetic data sets (Zacchini et al., 2019). Therefore, although intriguing, the results obtained from the clumps of cells using X20 cycles for PCR1 and a non-equimolar primer pool must be interpreted carefully, until additional controls are included, such as a higher DNA input to account for the amplification bias that results from low DNA yields.

7.1.4.2.3. Deduplication of PCR clones using unique molecular identifiers (UMIs)

Nowadays, library preparation protocols rely on PCR amplification to generate enough quantities of template DNA for successful sequencing. In addition, during PCR, NGS adapters are commonly incorporated. Recently, UMIs have been used to computationally identify read PCR duplicates in high-throughput sequencing experiments (Islam et al., 2014).

This study employed UMIs comprised of eight nucleotides. The UMI sequence was located between the TruSeq adapter and the targeted read and incorporated during PCR1 of library preparation. After read deduplication, the number of reads mapped to the reference genome decreased abruptly for all DMRs and under all experimental

conditions. The highest number of mapped reads observed after deduplication from all experimental conditions was 27, meaning that, for some of the DMRs that showed the highest number of reads before deduplication, the number of reads decreased approximately 3,000 times.

Singe-cell experiments that rely on increased PCR amplification cycles to increase libraries' concentrations are particularly prone to the production of PCR duplicate bias (Islam et al., 2014). In addition, amplicon sequencing, where the sequencing targets distinct loci as opposed to whole-genome sequencing, results in greater depth, therefore, making PCR duplication more acute (Smith et al., 2017). Another element of this study that further increases the vulnerability to PCR duplicate bias is the libraries with low sequence complexity and degraded nature, due to bisulfite conversion (Clement et al., 2018). Thus, it can be speculated that this study was highly prone to bias caused by PCR duplicates, especially clumps of cells, which could explain the high number of reads removed during deduplication. However, RNA-seq studies have found during their UMIs optimisation a fraction of duplicate reads of approximately 38% (Fu et al., 2018) and the fraction of duplicate reads in the present study far exceeds this number (for instance, a decrease of 99.97% in the average number of reads mapped for *PEG3*).

There was a remarkable drop in the coverage of both CpGs and SNPs for analysis because of extreme read removal after deduplication. Strikingly, in some instances, such as *MKRN3*, the heterozygous SNP frequency of the G allele increased from 76% before deduplication to 100% after deduplication in clumps of cells. From these results, it can be inferred that reads that were not real PCR duplicates were eliminated during computational deduplication.

After deduplication, the difference between the average CpGs methylation and SNP frequency was expected to decrease. Moreover, both averages were expected to be closer to 50%. Nonetheless, this was not the case for either the commercial gDNA or clumps of cells. All the above results suggest that the UMI design and duplicate removal execution were suboptimal. Two elements must be considered when using UMIs to overcome PCR bias. Firstly, the design of the UMI must include a sufficiently large sequence allowing for enough combinations to avoid UMI collision, which is the probability of two separate targets with the exact same sequence in the starting pool tagged with the same UMI that leads to the elimination of unique meaningful reads in downstream analysis (Fu et al., 2018). In this regard, the probability of UMI collision in amplicon sequencing is higher than in whole-genome sequencing because the number of reads that share the same sequence in genome-wide analysis is minimal. Many bioinformatical tools have been developed to process UMI information, but most of these computational tools are designed for whole-genome sequencing. Recently, however, new open-source tools have been developed that specialise in processing sequencing data from amplicon-based sequencing. These resources also aid in the UMI design, according to the study's experimental design, to determine the minimum UMI nucleotide sequence length to avoid UMI collisions. One is AmpUMI by Clement et al. (2018).

Secondly, UMIs are normally tagged to DNA fragments before PCR amplification (Kivioja et al., 2012). Nonetheless, the present study tagged the UMIs during the PCR1 of library preparation, due to the very low input DNA template (clumps of cells). As already mentioned, low DNA inputs are more prone to PCR duplicate bias, which the study could not account for in PCR1 with the use of UMIs. Future optimisation experiments can,

therefore, process UMIs information using whole-genome and amplicon sequencing UMI processing tools, including heterozygous SNPs as control for allele frequency, to compare performance in PCR duplicate removal. Unfortunately, the limitation of incorporating UMIs during PCR amplification, and not before, cannot be overcome due to the minute quantities of DNA found at single-cell levels.

7.1.4.3. Conclusions

Read duplicates are an important contributor to PCR amplification bias, which can negatively affect downstream DNA analysis when allele frequency is altered, leading to erroneous biological conclusions. One example of an approach that can be greatly affected is DNA methylation analysis with BS of CpGs located in DMRs. The use of heterozygous SNPs can help determine allele amplification bias. Nevertheless, there is high variation in the presence of heterozygous SNPs in populations and additional tools are needed to identify the bias produced by PCR duplicates.

This study aimed to develop a library preparation protocol for targeted BS for the analysis of 46 human DMRs in clumps of cells where UMIs were tagged to targets during the first of two PCRs for library preparation. Heterozygous SNPs, when present, were used as an allele frequency control. A high variation in the number of reads mapped to the reference genome was observed between the targeted DMRs and a non-equimolar pool did not improve uniformity. Decreasing the number of cycles in the first PCR from X20 to X15 also failed to show increased uniformity. There was an increased amplification bias in clumps of cells compared to 100 ng of gDNA, likely due to the minute amounts of template DNA and its degraded and low-complexity nature after bisulfite conversion. After computational deduplication using UMIs, the number of reads mapped to the

reference genome decreased abruptly, leading to poor coverage and a high variation in the average CpGs methylation values as well as heterozygous SNP frequencies. After duplicate removal, the difference between the heterozygous SNP frequencies and the average CpGs methylation did not decrease. Moreover, the approximation of these two values to 50% did not improve, providing evidence that the UMI design and execution were suboptimal.

Future experiments for further method optimisation should assess bisulfite conversion efficiency with strategies such as phage lambda bisulfite-converted DNA and perform primer titration to avoid target overrepresentation. In addition, increased iterations using the same experimental conditions can provide additional data to generate more significance in the number of reads mapped, average CpGs methylation percentage and heterozygous SNP frequency values to draw more accurate conclusions. Moreover, using the same source of clumps of cells to extract greater amounts of gDNA could also help to identify more accurately amplification bias caused by low DNA yields. Regarding the use of UMIs, future experiments can employ novel UMI information processing tools designed for amplicon NGS. These tools can also assist in designing UMI to avoid UMI collision due to a suboptimal sequence length.

The major limitation remains that low yields of DNA obtained from single-cell samples require an amplification step and that UMIs cannot be tagged before PCR, introducing bias that even an optimal UMI experimental design cannot account for. Nonetheless, this study's methodology can be optimised for approaches to analyse the methylation status of DMRs from samples with greater DNA yields, such as placental tissue and blood.

Ultimately, both preferential allele amplification and PCR duplication contribute to PCR amplification bias. Thus, ideally, future protocols will include heterozygous SNPs, when available, to account for the former and properly designed and executed UMIs to overcome the latter.

Chapter 8. Final conclusions

8.1.1. Thesis summary

8.1.1.1. Introductory remarks

Infertility is a prevalent disease that affects around one in 10 couples worldwide (Inhorn and Patrizio, 2015). The past 45 years have been characterised by an immense progress in ARTs. Pregnancies from previously untreatable cases are now possible because current ARTs allow not only for the in vitro handling of gametes, but also help to identify embryos for transfer with the potential to achieve implantation. The most well studied embryo viability factor is ploidy status. Transferring euploid embryos, using sensitive and validated PGT-A methods, has proven to improve sustained pregnancy rates and live births (Tiegs et al., 2021). However, IVF success rates remain low (HFEA, 2023).

Despite great advances in the field, one of the biggest challenges with ARTs is selecting the most competent embryo. Euploidy is an important predictive factor for embryo viability. Nonetheless, one third of chromosomally normal embryo transfers fail to achieve a live birth (Forman et al., 2013, Franasiak and Scott, 2017). In addition, PGT-A significantly adds to the cost of IVF treatment, which is already inaccessible to a big proportion of the world population (Asemota and Klatsky, 2015).

Progress in reproductive genetics has helped to understand infertility better, but much is still yet to be discovered about gametogenesis and the preimplantation stages, including the molecular mechanisms behind aneuploidy. Preimplantation development is

characterised by a complex interplay between genetic, epigenetic and metabolic factors. Thus, exploring new avenues can provide new biomarkers for embryo viability.

This thesis explored and developed strategies for the evaluation of embryo competency, focused on its cytogenetic and epigenetic status, by using existing and novel molecular methodologies for analysis.

8.1.1.2. Aneuploidy mechanisms in oocytes and preimplantation embryos

The mechanism of origin behind segmental aneuploidy is unrepaired DNA damage (DSBs) while whole chromosome aneuploidy arises from chromosome mis-segregation during cell division. Moreover, research has found that embryo numerical aneuploidy increases along with women's reproductive age (Franasiak et al., 2014), but there are no reports of a correlation between women's age and segmental chromosomal abnormalities (Fragouli et al., 2013). Evidence suggests that DSBs capabilities may be deficient during the first mitoses in the preimplantation embryo. This coincides with the stage prior the major EGA wave. Several genes are involved in DNA damage repair in somatic cells, including *ATM*. Studies have found that DDR and chromosome segregation surveillance pathways work jointly during oogenesis to prevent the generation of aneuploid oocytes (Collins et al., 2015, Charalambous et al., 2023). Interestingly, the expression of genes involved in the ATM-mediated DSBs repair pathway decreased in GVs from reproductively aged women (Oktay et al., 2015). However, direct evidence of an association between ATM dysfunction and impaired oocyte quality is lacking.

This part of the thesis aimed to produce experimental data to study the direct evidence of: (1) a decreased capacity of human cleavage stage embryos to repair DNA

damage through the expression of essential DDR genes and (2) a decreased competency of mice oocytes to produce euploid embryos when *ATM* is downregulated.

The expression study provided data that supports previous observations about the upregulation of DDR genes (*BRCA1, ATM* and *TP53*) in cleavage stage embryos, at the 4- to 10-cell stage, that are harbouring DNA damage (Wells et al., 2005b). This study found instances of such upregulation in embryos at the 5- and 6-cell stage, suggesting that the embryo is capable of responding to DSBs very early into development. These findings are substantiated by recent findings of *ATM* activation gene transcripts of embryonic origin in 1- to 2-cell stage embryos (Asami et al., 2022). Thus, although likely sub optimally, the embryo may activate its own DDR pathways either before or at the beginning of the major EGA wave. The aneuploidy analysis of parthenotes produced from *ATM* knocked-down GV's, along with the existing literature (Marangos and Carroll, 2012, Bakhoum et al., 2014, Mannini et al., 2015), could not establish an association between ATM-pathway dysfunction and bad oocyte quality, as reflected in impaired mechanisms to ensure faithful meiotic chromosome segregation and DSBs repair in mice GV-stage oocyte onwards.

8.1.1.3. Non-invasive methods for embryo genetic testing

PGT-A routinely requires a biopsy of TE cells. Thus, this procedure is considered invasive and needs sophisticated personnel and equipment to avoid significant harm to the embryo. Similarly, prenatal genetic testing, performed in the clinic to test for aneuploidies, involves CVS or amniocentesis. Therefore, it is also considered an invasive procedure. Although NIPT is becoming a routine screening test, non-invasive methods for aneuploidy testing in IVF embryos are still being investigated. BF was initially explored.

Nonetheless, reported ploidy concordance with corresponding TE biopsies and whole embryos is highly variable (Tobler et al., 2015, Capalbo et al., 2018, Magli et al., 2019). Reported concordance rates of SCM, on the other hand, tend to be higher but discrepancies also exist (Xu et al., 2016, Vera-Rodriguez et al., 2018, Huang et al., 2019). This is thought to be partly due to varying levels of embryo manipulation prior to sample collection.

This part of the thesis aimed to: (1) assess the accuracy of a truly non-invasive PGT protocol for aneuploidies using SCM, (2) develop a PCR-based NIPT protocol for the detection of the Y chromosome and (3) develop a PCR-based non-invasive sex selection tool for horse BF.

From this study it can be concluded that, compared to previous reports, interventions facilitate the release of embryonic DNA into the media, as reflected in the decrease in overall ploidy and karyotype concordance rates between TE biopsy and SCM (Xu et al., 2016, Liu et al., 2017, Huang et al., 2019). Yet, collecting SCM at later stages of development increases sample informativity. These findings are supported by other studies where embryo manipulation was also minimised (Vera-Rodriguez et al., 2018). So far, there are limited accessible tools to account for female DNA contamination in SCM accurately and there is still much to understand about how to interpret SCM niPGT results. Therefore, the accuracy of niPGT is currently inadequate to justify its routine clinical use.

It has been suggested that the inadequacy of BF for niPGT is due to the bad quality and low quantity of embryonic DNA (Leaver and Wells, 2020). Nonetheless, this thesis failed to support the hypothesis that bigger blastocyst sizes provide better sample amplifiability from increased accumulation of nuclear DNA. Although, studies with bigger

sample sizes and at different stages of horse blastocyst development are needed to confirm this. This thesis obtained promising results about the potential for a highly sensitive PCR-based Y chromosome NIPT that would be able to detect placental cfDNA at week-10 gestations and, potentially, earlier.

8.1.1.4. Epigenetic avenues for embryo genetic testing

The preimplantation stage of embryo development is characterised by a tight dynamic control of gene expression programs where the epigenome is particularly vulnerable (Monk et al., 2019). Thus, studies have analysed the presence of methylation perturbations in imprinted genes, mainly *H19*, *SNRPN*, *KCNQ1OT1,* and *PEG3,* in IVF embryos (Fauque et al., 2007, Market-Velker et al., 2010b, Kindsfather et al., 2019). Nonetheless, there is high variation in the reported DNA methylation levels of imprinted loci, suggesting the need for robust methods for CpGs methylation analysis, in addition to appropriate controls. Heterozygous SNPs can be used as controls to detect amplification bias that distort the true abundance of molecules in biological samples. Nonetheless, their presence varies amongst individuals and additional tools are needed to identify the bias produced by PCR duplicates.

This section aimed to explore the potential for UMIs technology to directly count molecules in a targeted BS approach to simultaneously detect methylation perturbations in 46 DMRs in human preimplantation embryos.

To our knowledge, this is the first time a panel of almost 50 DMRs was analysed in parallel and where the UMIs technology was tested in CpGs methylation level analysis. One major limitation of this approach is the unfeasibility of introducing the UMIs in DNA

from a single cell before PCR, since the identifiers cannot account for the bias produced in the amplification where they were attached to the molecules. This is particularly problematic because the results of the heterozygous SNPs frequency also show that the degree of amplification bias is higher in lower DNA inputs.

8.1.2. Study limitations

Individual limitations for each chapter are addressed in the corresponding discussion. Here, generalised limitations of the methodologies and approaches employed are briefly discussed in the context of ongoing research.

Progress in the field of reproductive genetics has produced sensitive molecular methodologies for analysis and diagnostics. Nonetheless, mosaicism PGT-A result management in the clinic has recently been restated due to reports of technical variation in the assay from artefacts introduced by WGA and NGS (Capalbo et al., 2017). High-level mosaicism can negatively affect the pregnancy outcome (Viotti et al., 2021, Babariya et al., 2023). Highly sensitive methodologies that allow for the distinction of true biological mosaicism are not only needed in the clinic, but also in research. As discussed in **Chapter 3**, identifying true mosaicism could have provided pivotal information to identify the origin of aneuploidy. Albeit, to identify mosaicism several embryonic cells would ideally be included, an impediment in cleavage stage embryo analysis. Moreover, **Chapter 5** showcases a developing approach to identify embryonic aneuploidy in SCM where this thesis and other studies have suggested current NGS technologies may not be sensitive enough to detect mosaicism (Vera-Rodriguez et al., 2018).

Sample availability is a major restriction in reproductive genetics research. Regarding meiotic aneuploidy, the origin and mechanisms behind it can only be properly

studied if oocytes or embryos are paired with PBs, as observed in **Chapter 4**. However, collection is laborious and, thus, not always available. Furthermore, as discussed in **Chapter 5**, in some instances, a large cohort of whole embryos is needed to analyse karyotype and ploidy concordance rates for novel PGT validation studies (Capalbo et al., 2017), due to the high frequency of mosaicism observed in preimplantation embryos. Yet, in practice, access to whole embryos is rarely feasible. What is more, when analysing the DDR capacity of preimplantation embryos in **Chapter 3**, ideal references would include embryos without signs of DNA damage, which further restricts sample availability.

8.1.3. Future research

Much is still yet to be understood about DNA repair mechanisms in preimplantation embryos. Asami et al. (2022) argue that embryo gene expression begins at the one-cell stage, and their work is supported by previously published data (Piras et al., 2014). Their findings have remarkable implications in the field. Nonetheless, results produced from single-cell RNA-seq must be followed up with, for instance, RNAi studies targeting genes of interest. As discussed earlier, access to human embryos is extremely rare, and so more tractable animal models can be employed instead.

Embryo viability biomarkers different from ploidy status have been explored, including epigenetic regulation in the preimplantation stages. Epidemiological studies have consistently indicated a correlation between ARTs and an increased incidence of rare imprinting disorders (Vermeiden and Bernardus, 2013, Cortessis et al., 2018). These findings have prompted the emergence of molecular studies to find a direct effect of ARTs on methylation levels of imprinted loci. Nonetheless, it is technically challenging to analyse a large cohort of characterised DMRs in single-cell samples. Proper controls are

lacking and there is high degree of natural epigenetic heterogeneity between individuals (Sakatani et al., 2001). This thesis showed that using UMIs to manage amplification bias in BS may not be feasible with single-cell samples. Nonetheless, future research could experiment with placental tissue, where more samples are available, to obtain better gDNA yields and attach UMIs before amplification for BS library preparation. In addition, methylation perturbations findings must be corroborated with gene expression studies.

References

Adamson, G. D., de Mouzon, J., Chambers, G. M., Zegers-Hochschild, F., Mansour, R., Ishihara, O., Banker, M. and Dyer, S. (2018) 'International Committee for Monitoring Assisted Reproductive Technology: world report on assisted reproductive technology, 2011', *Fertility and sterility,* 110(6), pp. 1067-1080.

Adiga, S. K., Toyoshima, M., Shiraishi, K., Shimura, T., Takeda, J., Taga, M., Nagai, H., Kumar, P. and Niwa, O. (2007) 'p21 provides stage specific DNA damage control to preimplantation embryos', *Oncogene,* 26(42), pp. 6141-6149.

Aguilar, J., Luzuriaga, I., Chavero, V. P., Leaniz, J., Casale, P., Carboneti, L. and Marino, V. (2012) 'Interruption of gestation and reuse of recipient mares in a commercial embryo transfer program', *Journal of Equine Veterinary Science,* 7(32), pp. 412-413.

Aguilera, A. and García-Muse, T. (2013) 'Causes of genome instability', *Annual review of genetics,* 47, pp. 1-32.

Alberry, M., Maddocks, D., Jones, M., Abdel Hadi, M., Abdel-Fattah, S., Avent, N. and Soothill, P. (2007) 'Free fetal DNA in maternal plasma in anembryonic pregnancies: confirmation that the origin is the trophoblast', *Prenatal Diagnosis: Published in Affiliation With the International Society for Prenatal Diagnosis,* 27(5), pp. 415-418.

Ariga, H., Ohto, H., Busch, M. P., Imamura, S., Watson, R., Reed, W. and Lee, T. H. (2001) 'Kinetics of fetal cellular and cell-free DNA in the maternal circulation during and after pregnancy: implications for noninvasive prenatal diagnosis', *Transfusion,* 41(12), pp. 1524-1530.

Asami, M., Lam, B. Y., Ma, M. K., Rainbow, K., Braun, S., VerMilyea, M. D., Yeo, G. S. and Perry, A. C. (2022) 'Human embryonic genome activation initiates at the one-cell stage', *Cell Stem Cell,* 29(2), pp. 209-216. e4.

Asemota, O. A. and Klatsky, P. 'Access to infertility care in the developing world: the family promotion gap'. *Seminars in reproductive medicine*: Thieme Medical Publishers, 017-022.

Ata, B., Capuzzo, M., Turkgeldi, E., Yildiz, S. and La Marca, A. (2021) 'Progestins for pituitary suppression during ovarian stimulation for ART: a comprehensive and systematic review including meta-analyses', *Human Reproduction Update,* 27(1), pp. 48-66.

Aurich, C. and Schneider, J. (2014) 'Sex determination in horses—Current status and future perspectives', *Animal Reproduction Science,* 146(1-2), pp. 34-41.

Baart, E. B., Martini, E., van den Berg, I., Macklon, N., Galjaard, R. H., Fauser, B. and Van Opstal, D. (2006) 'Preimplantation genetic screening reveals a high incidence of aneuploidy and mosaicism in embryos from young women undergoing IVF', *Human reproduction,* 21(1), pp. 223-233.

Babariya, D., Fragouli, E., Alfarawati, S., Spath, K. and Wells, D. (2017) 'The incidence and origin of segmental aneuploidy in human oocytes and preimplantation embryos', *Human Reproduction,* 32(12), pp. 2549-2560.

Babariya, D., Gill, P., Tao, X., Zhan, Y., Figliuzzi, M., Werner, M., Jalas, C. and Capalbo, A. (2023) 'O-290 Clinical utility of putative mosaicism detected using concurrent

copy-number and genotyping PGT method: outcomes from multisite, prospective, non-selection study including 9828 single embryo transfer cycles', *Human Reproduction,* 38(Supplement_1), pp. dead093. 354.

Bailey, J. L. (2010) 'Factors regulating sperm capacitation', *Systems biology in reproductive medicine,* 56(5), pp. 334-348.

Bakhoum, S. F., Kabeche, L., Compton, D. A., Powell, S. N. and Bastians, H. (2017) 'Mitotic DNA damage response: at the crossroads of structural and numerical cancer chromosome instabilities', *Trends in cancer,* 3(3), pp. 225-234.

Bakhoum, S. F., Kabeche, L., Murnane, J. P., Zaki, B. I. and Compton, D. A. (2014) 'DNA-damage response during mitosis induces whole-chromosome missegregation', *Cancer discovery,* 4(11), pp. 1281-1289.

Bartolomei, M. S. and Ferguson-Smith, A. C. (2011) 'Mammalian genomic imprinting', *Cold Spring Harbor perspectives in biology,* 3(7), pp. a002592.

Baudat, F., Imai, Y. and De Massy, B. (2013) 'Meiotic recombination in mammals: localization and regulation', *Nature Reviews Genetics,* 14(11), pp. 794-806.

Baumann, P. and West, S. C. (1998) 'Role of the human RAD51 protein in homologous recombination and double-stranded-break repair', *Trends in biochemical sciences,* 23(7), pp. 247-251.

Berntsen, S., Laivuori, H., la Cour Freiesleben, N., Loft, A., Söderström-Anttila, V., B Oldereid, N., Romundstad, L. B., Magnusson, Å., Petzold, M. and Bergh, C. (2021) 'A systematic review and meta-analysis on the association between ICSI and chromosome abnormalities', *Human Reproduction Update,* 27(5), pp. 801-847.

Berrada, H., Soriano, J., Pico, Y. and Manes, J. (2006) 'Quantification of Listeria monocytogenes in salads by real time quantitative PCR', *International Journal of Food Microbiology,* 107(2), pp. 202-206.

Betteridge, K., Eaglesome, M., Mitchell, D., Flood, P. and Beriault, R. (1982) 'Development of horse embryos up to twenty two days after ovulation: observations on fresh specimens', *Journal of anatomy,* 135(Pt 1), pp. 191.

Betts, D. and Madan, P. (2008) 'Permanent embryo arrest: molecular and cellular concepts', *MHR: Basic science of reproductive medicine,* 14(8), pp. 445-453.

Bhattacharya, S., Hamilton, M., Shaaban, M., Khalaf, Y., Seddler, M., Ghobara, T., Braude, P., Kennedy, R., Rutherford, A. and Hartshorne, G. (2001) 'Conventional in-vitro fertilisation versus intracytoplasmic sperm injection for the treatment of non-male-factor infertility: a randomised controlled trial', *The lancet,* 357(9274), pp. 2075-2079.

Bohrer, R. C., Dicks, N., Gutierrez, K., Duggavathi, R. and Bordignon, V. (2018) 'Double-strand DNA breaks are mainly repaired by the homologous recombination pathway in early developing swine embryos', *The FASEB Journal,* 32(4), pp. 1818-1829.

Bolton, H., Graham, S. J., Van der Aa, N., Kumar, P., Theunis, K., Fernandez Gallardo, E., Voet, T. and Zernicka-Goetz, M. (2016) 'Mouse model of chromosome mosaicism reveals lineage-specific depletion of aneuploid cells and normal developmental potential', *Nature communications,* 7(1), pp. 11165.

Bonduelle, M., Van Assche, E., Joris, H., Keymolen, K., Devroey, P., Van Steirteghem, A. and Liebaers, I. (2002) 'Prenatal testing in ICSI pregnancies: incidence of

chromosomal anomalies in 1586 karyotypes and relation to sperm parameters', *Human reproduction,* 17(10), pp. 2600-2614.

Bos-Mikich, A., Wood, M., Candy, C. and Whittingham, D. (1995) 'Cytogenetical analysis and developmental potential of vitrified mouse oocytes', *Biology of reproduction,* 53(4), pp. 780-785.

Bosch, E., Broer, S., Griesinger, G., Grynberg, M., Humaidan, P., Kolibianakis, E., Kunicki, M., La Marca, A. and Lainas, G. (2020) 'ESHRE guideline: ovarian stimulation for IVF/ICSI', *Human reproduction open,* 2020(2), pp. hoaa009.

Bowman-Smart, H., Gyngell, C., Mand, C., Amor, D. J., Delatycki, M. B. and Savulescu, J. (2021) 'Non-Invasive Prenatal Testing for "Non-Medical" Traits: Ensuring Consistency in Ethical Decision-Making', *The American Journal of Bioethics*, pp. 1-18.

Bowman-Smart, H., Savulescu, J., Gyngell, C., Mand, C. and Delatycki, M. B. (2020) 'Sex selection and non-invasive prenatal testing: A review of current practices, evidence, and ethical issues', *Prenatal Diagnosis,* 40(4), pp. 398-407.

Braekeleer, M. D. and Dao, T.-N. (1990) 'Cytogenetic studies in couples experiencing repeated pregnancy losses', *Human Reproduction,* 5(5), pp. 519-528.

Braude, P., Bolton, V. and Moore, S. (1988) 'Human gene expression first occurs between the four-and eight-cell stages of preimplantation development', *Nature,* 332(6163), pp. 459-461.

Brunschwig, H., Levi, L., Ben-David, E., Williams, R. W., Yakir, B. and Shifman, S. (2012) 'Fine-scale maps of recombination rates and hotspots in the mouse genome', *Genetics,* 191(3), pp. 757-764.

Buehler, E., Chen, Y.-C. and Martin, S. (2012) 'C911: A bench-level control for sequence specific siRNA off-target effects', *PLoS One,* 7(12), pp. e51942.

Burma, S., Chen, B. P. and Chen, D. J. (2006) 'Role of non-homologous end joining (NHEJ) in maintaining genomic integrity', *DNA repair,* 5(9-10), pp. 1042-1048.

Canick, J. A., Palomaki, G. E., Kloza, E. M., Lambert-Messerlian, G. M. and Haddow, J. E. (2013) 'The impact of maternal plasma DNA fetal fraction on next generation sequencing tests for common fetal aneuploidies', *Prenatal diagnosis,* 33(7), pp. 667-674.

Capalbo, A., Rienzi, L., Cimadomo, D., Maggiulli, R., Elliott, T., Wright, G., Nagy, Z. P. and Ubaldi, F. M. (2014) 'Correlation between standard blastocyst morphology, euploidy and implantation: an observational study in two centers involving 956 screened blastocysts', *Hum Reprod,* 29(6), pp. 1173-81.

Capalbo, A., Romanelli, V., Patassini, C., Poli, M., Girardi, L., Giancani, A., Stoppa, M., Cimadomo, D., Ubaldi, F. M. and Rienzi, L. (2018) 'Diagnostic efficacy of blastocoel fluid and spent media as sources of DNA for preimplantation genetic testing in standard clinical conditions', *Fertility and sterility,* 110(5), pp. 870-879. e5.

Capalbo, A., Ubaldi, F. M., Rienzi, L., Scott, R. and Treff, N. (2017) 'Detecting mosaicism in trophectoderm biopsies: current challenges and future possibilities', *Human reproduction,* 32(3), pp. 492-498.

Carlson, C. S., Emerson, R. O., Sherwood, A. M., Desmarais, C., Chung, M.-W., Parsons, J. M., Steen, M. S., LaMadrid-Herrmannsfeldt, M. A., Williamson, D. W. and

Livingston, R. J. (2013) 'Using synthetic templates to design an unbiased multiplex PCR assay', *Nature communications,* 4(1), pp. 2680.

Carson, S. A. and Kallen, A. N. (2021) 'Diagnosis and management of infertility: a review', *Jama,* 326(1), pp. 65-76.

Carvalho, F., Moutou, C., Dimitriadou, E., Dreesen, J., Giménez, C., Goossens, V., Kakourou, G., Vermeulen, N. and Zuccarello, D. (2020) 'ESHRE PGT Consortium good practice recommendations for the detection of monogenic disorders', *Human reproduction open,* 2020(3), pp. hoaa018.

Charalambous, C., Webster, A. and Schuh, M. (2023) 'Aneuploidy in mammalian oocytes and the impact of maternal ageing', *Nature Reviews Molecular Cell Biology,* 24(1), pp. 27-44.

Chen, J.-J., Silver, D., Cantor, S., Livingston, D. M. and Scully, R. (1999) 'BRCA1, BRCA2, and Rad51 operate in a common DNA damage response pathway', *Cancer research,* 59(7 Supplement), pp. 1752s-1756s.

Chen, Y., Li, H. and Meng, L. (2013) 'Prenatal sex selection and missing girls in China: Evidence from the diffusion of diagnostic ultrasound', *Journal of Human Resources,* 48(1), pp. 36-70.

Cheng, J.-M., Li, J., Tang, J.-X., Hao, X.-X., Wang, Z.-P., Sun, T.-C., Wang, X.-X., Zhang, Y., Chen, S.-R. and Liu, Y.-X. (2017) 'Merotelic kinetochore attachment in oocyte meiosis II causes sister chromatids segregation errors in aged mice', *Cell Cycle,* 16(15), pp. 1404-1413.

Cheung, H. W., Wong, K. S., Lin, V. Y., Wan, T. S. and Ho, E. N. (2021) 'A duplex qPCR assay for human erythropoietin (EPO) transgene to control gene doping in horses', *Drug testing and analysis,* 13(1), pp. 113-121.

Choi, Y., Gustafson-Seabury, A., Velez, I., Hartman, D., Bliss, S., Riera, F., Roldán, J., Chowdhary, B. and Hinrichs, K. (2010) 'Viability of equine embryos after puncture of the capsule and biopsy for preimplantation genetic diagnosis', *Reproduction,* 140(6), pp. 893.

Chronopoulou, E. and Harper, J. C. (2015) 'IVF culture media: past, present and future', *Human reproduction update,* 21(1), pp. 39-55.

Clement, K., Farouni, R., Bauer, D. E. and Pinello, L. (2018) 'AmpUMI: design and analysis of unique molecular identifiers for deep amplicon sequencing', *Bioinformatics,* 34(13), pp. i202-i210.

Colchen, S., Battut, I., Fieni, F., Tainturier, D., Siliart, B. and Bruyas, J. (2000) 'Quantitative histological analysis of equine embryos at exactly 156 and 168 h after ovulation', *Journal of Reproduction and fertility. Supplement,* (56), pp. 527-537.

Coll, L., Parriego, M., Boada, M., Devesa, M., Arroyo, G., Rodríguez, I., Coroleu, B., Vidal, F. and Veiga, A. (2018) 'Transition from blastomere to trophectoderm biopsy: comparing two preimplantation genetic testing for aneuploidies strategies', *Zygote,* 26(3), pp. 191-198.

Collins, J. K., Lane, S. I., Merriman, J. A. and Jones, K. T. (2015) 'DNA damage induces a meiotic arrest in mouse oocytes mediated by the spindle assembly checkpoint', *Nature communications,* 6(1), pp. 8553.

Colls, P., Goodall, N., Zheng, X. and Munné, S. (2009) 'Increased efficiency of preimplantation genetic diagnosis for aneuploidy by testing 12 chromosomes', *Reproductive biomedicine online,* 19(4), pp. 532-538.

Cortessis, V. K., Azadian, M., Buxbaum, J., Sanogo, F., Song, A. Y., Sriprasert, I., Wei, P. C., Yu, J., Chung, K. and Siegmund, K. D. (2018) 'Comprehensive meta-analysis reveals association between multiple imprinting disorders and conception by assisted reproductive technology', *Journal of assisted reproduction and genetics,* 35, pp. 943-952.

Cox, G. F., Burger, J., Lip, V., Mau, U. A., Sperling, K., Wu, B. L. and Horsthemke, B. (2002) 'Intracytoplasmic sperm injection may increase the risk of imprinting defects', *Am J Hum Genet,* 71(1), pp. 162-4.

Craciunas, L., Gallos, I., Chu, J., Bourne, T., Quenby, S., Brosens, J. J. and Coomarasamy, A. (2019) 'Conventional and modern markers of endometrial receptivity: a systematic review and meta-analysis', *Human reproduction update,* 25(2), pp. 202-223.

Crișan, M., Damian, A., Morar, I., Páll, E., Pește an, C. and Groza, I. Ș. (2016) 'Equine embryo sexing and ultrasonographic foetal sexing–interests and applicability', *Anatomia, histologia, embryologia,* 45(5), pp. 329-337.

D'Alessandro, A., Federica, G., Palini, S., Bulletti, C. and Zolla, L. (2012) 'A mass spectrometry-based targeted metabolomics strategy of human blastocoele fluid: a promising tool in fertility research', *Molecular BioSystems,* 8(4), pp. 953-958.

Danylevska, A., Kovacovicova, K., Awadova, T. and Anger, M. (2014) 'The frequency of precocious segregation of sister chromatids in mouse female meiosis I is affected by genetic background', *Chromosome research,* 22, pp. 365-373.

De Rycke, M. and Berckmoes, V. (2020) 'Preimplantation genetic testing for monogenic disorders', *Genes,* 11(8), pp. 871.

DeBaun, M. R., Niemitz, E. L. and Feinberg, A. P. (2003) 'Association of in vitro fertilization with Beckwith-Wiedemann syndrome and epigenetic alterations of LIT1 and H19', *The American Journal of Human Genetics,* 72(1), pp. 156-160.

Debrock, S., Melotte, C., Spiessens, C., Peeraer, K., Vanneste, E., Meeuwis, L., Meuleman, C., Frijns, J.-P., Vermeesch, J. R. and D'Hooghe, T. M. (2010) 'Preimplantation genetic screening for aneuploidy of embryos after in vitro fertilization in women aged at least 35 years: a prospective randomized trial', *Fertility and sterility,* 93(2), pp. 364-373.

Delhanty, J. D., Griffin, D. K., Handyside, A. H., Harper, J., Atkinson, G. H., Pieters, M. H. and Winston, R. M. (1993) 'Detection of aneuploidy and chromosomal mosaicism in human embryos during preimplantation sex determination by fluorescent in situ hybridisation,(FISH)', *Human molecular genetics,* 2(8), pp. 1183-1185.

Denomme, M. M., McCallie, B. R., Parks, J. C., Schoolcraft, W. B. and Katz-Jaffe, M. G. (2016) 'Epigenetic dysregulation observed in monosomy blastocysts further compromises developmental potential', *PLoS One,* 11(6), pp. e0156980.

Derive, N., Landmann, C., Montembault, E., Claverie, M.-C., Pierre-Elies, P., Goutte-Gattat, D., Founounou, N., McCusker, D. and Royou, A. (2015) 'Bub3–BubR1-dependent sequestration of Cdc20Fizzy at DNA breaks facilitates the correct segregation of broken chromosomes', *Journal of Cell Biology,* 211(3), pp. 517-532.

Di Giacomo, M., Barchi, M., Baudat, F., Edelmann, W., Keeney, S. and Jasin, M. (2005) 'Distinct DNA-damage-dependent and-independent responses drive the loss of oocytes in recombination-defective mouse mutants', *Proceedings of the National Academy of Sciences,* 102(3), pp. 737-742.

Diefenbach, R. J., Lee, J. H., Kefford, R. F. and Rizos, H. (2018) 'Evaluation of commercial kits for purification of circulating free DNA', *Cancer genetics,* 228, pp. 21-27.

Dimitriadou, E., Van der Aa, N., Cheng, J., Voet, T. and Vermeesch, J. R. (2014) 'Single cell segmental aneuploidy detection is compromised by S phase', *Molecular cytogenetics,* 7, pp. 1-10.

Doornbos, M. E., Maas, S. M., McDonnell, J., Vermeiden, J. P. and Hennekam, R. C. (2007) 'Infertility, assisted reproduction technologies and imprinting disturbances: a Dutch study', *Human Reproduction,* 22(9), pp. 2476-2480.

Duncan, F. E., Chiang, T., Schultz, R. M. and Lampson, M. A. (2009) 'Evidence that a defective spindle assembly checkpoint is not the primary cause of maternal age-associated aneuploidy in mouse eggs', *Biology of reproduction,* 81(4), pp. 768-776.

Duncan, F. E. and Gerton, J. L. (2018) 'Mammalian oogenesis and female reproductive aging', *Aging (Albany NY),* 10(2), pp. 162.

Duncan, F. E., Hornick, J. E., Lampson, M. A., Schultz, R. M., Shea, L. D. and Woodruff, T. K. (2012) 'Chromosome cohesion decreases in human eggs with advanced maternal age', *Aging cell,* 11(6), pp. 1121-1124.

Eggenschwiler, J., Ludwig, T., Fisher, P., Leighton, P. A., Tilghman, S. M. and Efstratiadis, A. (1997) 'Mouse mutant embryos overexpressing IGF-II exhibit phenotypic features of the Beckwith–Wiedemann and Simpson–Golabi–Behmel syndromes', *Genes & development,* 11(23), pp. 3128-3142.

El Hajj, N., Haertle, L., Dittrich, M., Denk, S., Lehnen, H., Hahn, T., Schorsch, M. and Haaf, T. (2017) 'DNA methylation signatures in cord blood of ICSI children', *Human Reproduction,* 32(8), pp. 1761-1769.

Eldridge-Panuska, W., Di Brienza, V. C., Seidel Jr, G., Squires, E. and Carnevale, E. (2005) 'Establishment of pregnancies after serial dilution or direct transfer by vitrified equine embryos', *Theriogenology,* 63(5), pp. 1308-1319.

Elnifro, E. M., Ashshi, A. M., Cooper, R. J. and Klapper, P. E. (2000) 'Multiplex PCR: optimization and application in diagnostic virology', *Clinical microbiology reviews,* 13(4), pp. 559-570.

Fauque, P., Jouannet, P., Lesaffre, C., Ripoche, M.-A., Dandolo, L., Vaiman, D. and Jammes, H. (2007) 'Assisted Reproductive Technology affects developmental kinetics, H19 Imprinting Control Region methylation and H19 gene expression in individual mouse embryos', *BMC developmental biology,* 7, pp. 1-19.

Filippo, J. S., Sung, P. and Klein, H. (2008) 'Mechanism of Eukaryotic Homologous Recombination', *Annual Review of Biochemistry,* 77(1), pp. 229-257.

Findikli, N., Kahraman, S., Kumtepe, Y., Donmez, E., Benkhalifa, M., Biricik, A., Sertyel, S., Berkil, H. and Oncu, N. (2004) 'Assessment of DNA fragmentation and aneuploidy on poor quality human embryos', *Reproductive BioMedicine Online,* 8(2), pp. 196-206.

Fiorentino, F., Biricik, A., Bono, S., Spizzichino, L., Cotroneo, E., Cottone, G., Kokocinski, F. and Michel, C.-E. (2014) 'Development and validation of a next-generation

sequencing–based protocol for 24-chromosome aneuploidy screening of embryos', *Fertility and sterility,* 101(5), pp. 1375-1382. e2.

Fogarty, N. M. E., McCarthy, A., Snijders, K. E., Powell, B. E., Kubikova, N., Blakeley, P., Lea, R., Elder, K., Wamaitha, S. E., Kim, D., Maciulyte, V., Kleinjung, J., Kim, J.-S., Wells, D., Vallier, L., Bertero, A., Turner, J. M. A. and Niakan, K. K. (2017) 'Genome editing reveals a role for OCT4 in human embryogenesis', *Nature,* 550(7674), pp. 67-73.

Forman, E. J., Hong, K. H., Ferry, K. M., Tao, X., Taylor, D., Levy, B., Treff, N. R. and Scott Jr, R. T. (2013) 'In vitro fertilization with single euploid blastocyst transfer: a randomized controlled trial', *Fertility and sterility,* 100(1), pp. 100-107. e1.

Fouks, Y. and Yogev, Y. (2022) 'Twinning in ART: Single embryo transfer policy', *Best Practice & Research Clinical Obstetrics & Gynaecology*.

Fragouli, E., Alfarawati, S., Daphnis, D. D., Goodall, N.-n., Mania, A., Griffiths, T., Gordon, A. and Wells, D. (2011) 'Cytogenetic analysis of human blastocysts with the use of FISH, CGH and aCGH: scientific data and technical evaluation', *Human Reproduction,* 26(2), pp. 480-490.

Fragouli, E., Alfarawati, S., Spath, K., Babariya, D., Tarozzi, N., Borini, A. and Wells, D. (2017) 'Analysis of implantation and ongoing pregnancy rates following the transfer of mosaic diploid–aneuploid blastocysts', *Human Genetics,* 136(7), pp. 805-819.

Fragouli, E., Alfarawati, S., Spath, K., Jaroudi, S., Sarasa, J., Enciso, M. and Wells, D. (2013) 'The origin and impact of embryonic aneuploidy', *Human genetics,* 132, pp. 1001-1013.

Fragouli, E., Munne, S. and Wells, D. (2019) 'The cytogenetic constitution of human blastocysts: insights from comprehensive chromosome screening strategies', *Human reproduction update,* 25(1), pp. 15-33.

Franasiak, J. M., Forman, E. J., Hong, K. H., Werner, M. D., Upham, K. M., Treff, N. R. and Scott Jr, R. T. (2014) 'The nature of aneuploidy with increasing age of the female partner: a review of 15,169 consecutive trophectoderm biopsies evaluated with comprehensive chromosomal screening', *Fertility and sterility,* 101(3), pp. 656-663. e1.

Franasiak, J. M. and Scott, R. T. (2017) 'Contribution of immunology to implantation failure of euploid embryos', *Fertility and sterility,* 107(6), pp. 1279-1283.

França, A., Freitas, A. I., Henriques, A. F. and Cerca, N. (2012) 'Optimizing a qPCR gene expression quantification assay for S. epidermidis biofilms: a comparison between commercial kits and a customized protocol', *PloS one,* 7(5), pp. e37480.

Fu, Y., Wu, P.-H., Beane, T., Zamore, P. D. and Weng, Z. (2018) 'Elimination of PCR duplicates in RNA-seq and small RNA-seq using unique molecular identifiers', *Bmc Genomics,* 19, pp. 1-14.

Gao, Y., Sun, Y., Frank, K. M., Dikkes, P., Fujiwara, Y., Seidl, K. J., Sekiguchi, J. M., Rathbun, G. A., Swat, W. and Wang, J. (1998) 'A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis', *Cell,* 95(7), pp. 891-902.

Garcia, V., Gray, S., Allison, R. M., Cooper, T. J. and Neale, M. J. (2015) 'Tel1ATM-mediated interference suppresses clustered meiotic double-strand-break formation', *Nature,* 520(7545), pp. 114-118.

Gardner, D. K. (1998) 'Changes in requirements and utilization of nutrients during mammalian preimplantation embryo development and their significance in embryo culture', *Theriogenology,* 49(1), pp. 83-102.

Gardner, D. K., Lane, M., Calderon, I. and Leeton, J. (1996) 'Environment of the preimplantation human embryo in vivo: metabolite analysis of oviduct and uterine fluids and metabolism of cumulus cells', *Fertility and sterility,* 65(2), pp. 349-353.

Gardner, D. K. and Schoolcraft, W. B. (1999) 'Culture and transfer of human blastocysts', *Current Opinion in Obstetrics and Gynecology,* 11(3), pp. 307-311.

Girardi, L., Serdarogullari, M., Patassini, C., Poli, M., Fabiani, M., Caroselli, S., Coban, O., Findikli, N., Boynukalin, F. K. and Bahceci, M. (2020) 'Incidence, origin, and predictive model for the detection and clinical management of segmental aneuploidies in human embryos', *The American Journal of Human Genetics,* 106(4), pp. 525-534.

Glujovsky, D., Pesce, R., Sueldo, C., Retamar, A. M. Q., Hart, R. J. and Ciapponi, A. (2020) 'Endometrial preparation for women undergoing embryo transfer with frozen embryos or embryos derived from donor oocytes', *Cochrane database of systematic reviews,* (10).

Gray, S. and Cohen, P. E. (2016) 'Control of meiotic crossovers: from double-strand break formation to designation', *Annual review of genetics,* 50, pp. 175-210.

Greally, J. M. (2018) 'A user's guide to the ambiguous word'epigenetics'', *Nature Reviews Molecular Cell Biology,* 19(4), pp. 207-208.

Griffin, D. K. and Ogur, C. (2023) 'PGT-SR: A Comprehensive Overview and a Requiem for the Interchromosomal Effect', *DNA,* 3(1), pp. 41-64.

Groff, A. F., Resetkova, N., DiDomenico, F., Sakkas, D., Penzias, A., Rinn, J. L. and Eggan, K. (2019) 'RNA-seq as a tool for evaluating human embryo competence', *Genome research,* 29(10), pp. 1705-1718.

Gruhn, J. R., Zielinska, A. P., Shukla, V., Blanshard, R., Capalbo, A., Cimadomo, D., Nikiforov, D., Chan, A. C.-H., Newnham, L. J. and Vogel, I. (2019) 'Chromosome errors in human eggs shape natural fertility over reproductive life span', *Science,* 365(6460), pp. 1466-1469.

Gyanchandani, R., Kvam, E., Heller, R., Finehout, E., Smith, N., Kota, K., Nelson, J. R., Griffin, W., Puhalla, S. and Brufsky, A. M. (2018) 'Whole genome amplification of cell-free DNA enables detection of circulating tumor DNA mutations from fingerstick capillary blood', *Scientific reports,* 8(1), pp. 1-12.

Haering, C. H., Farcas, A.-M., Arumugam, P., Metson, J. and Nasmyth, K. (2008) 'The cohesin ring concatenates sister DNA molecules', *Nature,* 454(7202), pp. 297-301.

Hamidi, T., Singh, A. K. and Chen, T. (2015) 'Genetic alterations of DNA methylation machinery in human diseases', *Epigenomics,* 7(2), pp. 247-265.

Hammond, E. R., McGillivray, B. C., Wicker, S. M., Peek, J. C., Shelling, A. N., Stone, P., Chamley, L. W. and Cree, L. M. (2017) 'Characterizing nuclear and mitochondrial DNA in spent embryo culture media: genetic contamination identified', *Fertility and sterility,* 107(1), pp. 220-228. e5.

Handyside, A. H., Kontogianni, E. H., Hardy, K. and Winston, R. M. (1990) 'Pregnancies from biopsied human preimplantation embryos sexed by Y-specific DNA amplification', *Nature,* 344(6268), pp. 768-770.

Hardarson, T., Caisander, G., Sjögren, A., Hanson, C., Hamberger, L. and Lundin, K. (2003) 'A morphological and chromosomal study of blastocysts developing from morphologically suboptimal human pre-embryos compared with control blastocysts', *Human reproduction,* 18(2), pp. 399-407.

Hardy, K. (1999) 'Apoptosis in the human embryo', *Reviews of Reproduction,* 4, pp. 125-134.

Hardy, K., Stark, J. and Winston, R. M. (2003) 'Maintenance of the inner cell mass in human blastocysts from fragmented embryos', *Biology of Reproduction,* 68(4), pp. 1165-1169.

Harper, J. C., de Die-Smulders, C., Goossens, V., Harton, G., Moutou, C., Repping, S., Scriven, P., SenGupta, S., Traeger-Synodinos, J. and Van Rij, M. (2008) 'ESHRE PGD consortium data collection VII: cycles from January to December 2004 with pregnancy follow-up to October 2005', *Human Reproduction,* 23(4), pp. 741-755.

Hart, R. J. (2016) 'Physiological aspects of female fertility: role of the environment, modern lifestyle, and genetics', *Physiological reviews,* 96(3), pp. 873-909.

Hassold, T., Hall, H. and Hunt, P. (2007) 'The origin of human aneuploidy: where we have been, where we are going', *Human molecular genetics,* 16(R2), pp. R203-R208.

He, D.-J., Wang, L., Zhang, Z.-B., Guo, K., Li, J.-Z., He, X.-C., Cui, Q.-H. and Zheng, P. (2018) 'Maternal gene Ooep may participate in homologous recombination-mediated DNA double-strand break repair in mouse oocytes', *Zoological research,* 39(6), pp. 387.

Herrera, C., Morikawa, M., Bello, M., Von Meyeren, M., Centeno, J. E., Dufourq, P., Martinez, M. and Llorente, J. (2014) 'Setting up equine embryo gender determination by preimplantation genetic diagnosis in a commercial embryo transfer program', *Theriogenology,* 81(5), pp. 758-763.

Herrera, C., Morikawa, M., Castex, C. B., Pinto, M. R., Ortega, N., Fanti, T., Garaguso, R., Franco, M., Castañares, M. and Castañeira, C. (2015) 'Blastocele fluid from in vitro–and in vivo–produced equine embryos contains nuclear DNA', *Theriogenology,* 83(3), pp. 415-420.

HFEA 2019. Fertility treatment 2019: Trends and figures. Human Fertilisation and Embryology Authority.

HFEA (2023) 'Fertility treatment 2021: preliminary trends and figures'.

Hinrichs, K. (2018) 'Assisted reproductive techniques in mares', *Reproduction in Domestic Animals,* 53, pp. 4-13.

Hinrichs, K. and Choi, Y.-H. (2016) 'Micromanipulation of equine blastocysts to allow vitrification', *Reproduction, Fertility and Development,* 28(8), pp. 1092-1096.

Hong, S. R. and Shin, K.-J. (2021) 'Bisulfite-converted DNA quantity evaluation: a multiplex quantitative real-time PCR system for evaluation of bisulfite conversion', *Frontiers in Genetics,* 12, pp. 618955.

Hou, Y., Fan, W., Yan, L., Li, R., Lian, Y., Huang, J., Li, J., Xu, L., Tang, F. and Xie, X. S. (2013) 'Genome analyses of single human oocytes', *Cell,* 155(7), pp. 1492-1506.

Hou, Y., Wu, K., Shi, X., Li, F., Song, L., Wu, H., Dean, M., Li, G., Tsang, S. and Jiang, R. (2015) 'Comparison of variations detection between whole-genome amplification methods used in single-cell resequencing', *Gigascience,* 4(1), pp. s13742-015-0068-3.

Huang, J. Y., Chen, H. Y., Park, J. Y. S., Tan, S. L. and Chian, R.-C. (2008) 'Comparison of spindle and chromosome configuration in in vitro-and in vivo-matured mouse oocytes after vitrification', *Fertility and Sterility,* 90(4), pp. 1424-1432.

Huang, L., Bogale, B., Tang, Y., Lu, S., Xie, X. S. and Racowsky, C. (2019) 'Noninvasive preimplantation genetic testing for aneuploidy in spent medium may be more reliable than trophectoderm biopsy', *Proceedings of the National Academy of Sciences,* 116(28), pp. 14105-14112.

Huang, L., Ma, F., Chapman, A., Lu, S. and Xie, X. S. (2015) 'Single-cell whole-genome amplification and sequencing: methodology and applications', *Annual review of genomics and human genetics,* 16, pp. 79-102.

Huhtinen, M., Peippo, J. and Bredbacka, P. (1997) 'Successful transfer of biopsied equine embryos', *Theriogenology,* 48(3), pp. 361-367.

Hui, L. and Bianchi, D. W. (2020) 'Fetal fraction and noninvasive prenatal testing: what clinicians need to know', *Prenatal Diagnosis,* 40(2), pp. 155-163.

Hustedt, N. and Durocher, D. (2017) 'The control of DNA repair by the cell cycle', *Nature cell biology,* 19(1), pp. 1-9.

Ibala-Romdhane, S., Al-Khtib, M., Khoueiry, R., Blachère, T., Guérin, J.-F. and Lefèvre, A. (2011) 'Analysis of H19 methylation in control and abnormal human embryos, sperm and oocytes', *European journal of human genetics,* 19(11), pp. 1138-1143.

Inhorn, M. C. and Patrizio, P. (2015) 'Infertility around the globe: new thinking on gender, reproductive technologies and global movements in the 21st century', *Human reproduction update,* 21(4), pp. 411-426.

Islam, S., Zeisel, A., Joost, S., La Manno, G., Zajac, P., Kasper, M., Lönnerberg, P. and Linnarsson, S. (2014) 'Quantitative single-cell RNA-seq with unique molecular identifiers', *Nature methods,* 11(2), pp. 163-166.

Ivanova, E., Canovas, S., Garcia-Martínez, S., Romar, R., Lopes, J. S., Rizos, D., Sanchez-Calabuig, M. J., Krueger, F., Andrews, S. and Perez-Sanz, F. (2020) 'DNA methylation changes during preimplantation development reveal inter-species differences and reprogramming events at imprinted genes', *Clinical epigenetics,* 12, pp. 1-18.

Jackson, A. L., Bartz, S. R., Schelter, J., Kobayashi, S. V., Burchard, J., Mao, M., Li, B., Cavet, G. and Linsley, P. S. (2003) 'Expression profiling reveals off-target gene regulation by RNAi', *Nature biotechnology,* 21(6), pp. 635-637.

Jackson, A. L. and Linsley, P. S. (2004) 'Noise amidst the silence: off-target effects of siRNAs?', *TRENDS in Genetics,* 20(11), pp. 521-524.

Jaiswal, S. K., Oh, J. J. and DePamphilis, M. L. (2020) 'Cell cycle arrest and apoptosis are not dependent on p53 prior to p53-dependent embryonic stem cell differentiation', *Stem Cells,* 38(9), pp. 1091-1106.

Jaiswal, S. K., Raj, S. and DePamphilis, M. L. (2021) 'Developmental acquisition of p53 functions', *Genes,* 12(11), pp. 1675.

Jakobsson, A.-H. (2018) 'Is it necessary to collapse blastocysts before vitrification?'.

Janečka, J. E., Davis, B. W., Ghosh, S., Paria, N., Das, P. J., Orlando, L., Schubert, M., Nielsen, M. K., Stout, T. A. E., Brashear, W., Li, G., Johnson, C. D., Metz, R. P., Zadjali, A. M. A., Love, C. C., Varner, D. D., Bellott, D. W., Murphy, W. J., Chowdhary, B. P. and Raudsepp, T. (2018) 'Horse Y chromosome assembly displays unique evolutionary features and putative stallion fertility genes', *Nature Communications,* 9(1), pp. 2945.

Jin, Z.-L., Shen, X.-H., Shuang, L., Kwon, J.-w., Seong, M.-J. and Kim, N.-H. (2019) 'Inhibition of DNA repair protein RAD51 affects porcine preimplantation embryo development', *Reproduction,* 157(3), pp. 223-234.

Jorgez, C. J., Dang, D. D., Simpson, J. L., Lewis, D. E. and Bischoff, F. Z. (2006) 'Quantity versus quality: optimal methods for cell-free DNA isolation from plasma of pregnant women', *Genetics in Medicine,* 8(10), pp. 615-619.

Jukam, D., Shariati, S. A. M. and Skotheim, J. M. (2017) 'Zygotic genome activation in vertebrates', *Developmental cell,* 42(4), pp. 316-332.

Kakourou, G., Jaroudi, S., Tulay, P., Heath, C., Serhal, P., Harper, J. C. and SenGupta, S. B. (2013) 'Investigation of gene expression profiles before and after embryonic genome activation and assessment of functional pathways at the human metaphase II oocyte and blastocyst stage', *Fertility and sterility,* 99(3), pp. 803-814. e23.

Kamphuis, E. I., Bhattacharya, S., Van Der Veen, F., Mol, B. W. and Templeton, A. (2014) 'Are we overusing IVF?', *Bmj,* 348.

Kaufman, M. (1982) 'The chromosome complement of single-pronuclear haploid mouse embryos following activation by ethanol treatment'.

Khokhlova, E. V., Fesenko, Z. S., Sopova, J. V. and Leonova, E. I. (2020) 'Features of DNA repair in the early stages of mammalian embryonic development', *Genes,* 11(10), pp. 1138.

Khoueiry, R., Ibala-Romdhane, S., Al-Khtib, M., Blachère, T., Lornage, J. and Lefèvre, A. (2013) 'Abnormal methylation of KCNQ1OT1 and differential methylation of H19 imprinting control regions in human ICSI embryos', *Zygote,* 21(2), pp. 129-138.

Kim, E. M. and Burke, D. J. (2008) 'DNA damage activates the SAC in an ATM/ATR-dependent manner, independently of the kinetochore', *PLoS genetics,* 4(2), pp. e1000015.

Kim, S., Peterson, S. E., Jasin, M. and Keeney, S. 'Mechanisms of germ line genome instability'. *Seminars in Cell & Developmental Biology*: Elsevier, 177-187.

Kim, S.-T., Xu, B. and Kastan, M. B. (2002) 'Involvement of the cohesin protein, Smc1, in Atm-dependent and independent responses to DNA damage', *Genes & development,* 16(5), pp. 560-570.

Kindsfather, A. J., Czekalski, M. A., Pressimone, C. A., Erisman, M. P. and Mann, M. R. (2019) 'Perturbations in imprinted methylation from assisted reproductive technologies but not advanced maternal age in mouse preimplantation embryos', *Clinical epigenetics,* 11, pp. 1-14.

Kirsch, C., Weickmann, S., Schmidt, B. and Fleischhacker, M. (2008) 'An improved method for the isolation of free-circulating plasma DNA and cell-free DNA from other body fluids', *Annals of the New York Academy of Sciences,* 1137(1), pp. 135-139.

Kitajima, T. S., Ohsugi, M. and Ellenberg, J. (2011) 'Complete kinetochore tracking reveals error-prone homologous chromosome biorientation in mammalian oocytes', *Cell,* 146(4), pp. 568-581.

Kivioja, T., Vähärautio, A., Karlsson, K., Bonke, M., Enge, M., Linnarsson, S. and Taipale, J. (2012) 'Counting absolute numbers of molecules using unique molecular identifiers', *Nature methods,* 9(1), pp. 72-74.

Kleckner, N. (1996) 'Meiosis: how could it work?', *Proceedings of the National Academy of Sciences,* 93(16), pp. 8167-8174.

Kolialexi, A., Tounta, G., Apostolou, P., Vrettou, C., Papantoniou, N., Kanavakis, E., Antsaklis, A. and Mavrou, A. (2012) 'Early non-invasive detection of fetal Y chromosome sequences in maternal plasma using multiplex PCR', *European Journal of Obstetrics & Gynecology and Reproductive Biology,* 161(1), pp. 34-37.

Konstantinidis, M., Milligan, K., Berkeley, A., Kennedy, J., Maxson, W., Racowsky, C., Wells, D. and Munne, S. (2016) 'Use of single nucleotide polymorphism (SNP) arrays and next generation sequencing (NGS) to study the incidence, type and origin of aneuploidy in the human preimplantation embryo', *Fertility and Sterility,* 106(3), pp. e22-e23.

Krupina, K., Goginashvili, A. and Cleveland, D. W. (2021) 'Causes and consequences of micronuclei', *Current opinion in cell biology,* 70, pp. 91-99.

Krüssel, J., Huang, H., Simon, C., Behr, B., Pape, A., Wen, Y., Bielfeld, P. and Polan, M. (1998) 'Single blastomeres within human preimplantation embryos express different amounts of messenger ribonucleic acid for beta-actin and interleukin-1 receptor type I', *The Journal of clinical endocrinology and metabolism,* 83(3), pp. 953-959.

Kubicek, D., Hornak, M., Horak, J., Navratil, R., Tauwinklova, G., Rubes, J. and Vesela, K. (2019) 'Incidence and origin of meiotic whole and segmental chromosomal aneuploidies detected by karyomapping', *Reproductive biomedicine online,* 38(3), pp. 330-339.

Kuznyetsov, V., Madjunkova, S., Antes, R., Abramov, R., Motamedi, G., Ibarrientos, Z. and Librach, C. (2018) 'Evaluation of a novel non-invasive preimplantation genetic screening approach', *Plos one,* 13(5), pp. e0197262.

Kypri, E., Ioannides, M., Achilleos, A., Koumbaris, G., Patsalis, P. and Stumm, M. (2022) 'Non-invasive prenatal screening tests–update 2022', *LaboratoriumsMedizin,* 46(4), pp. 311-320.

Lahn, B. T. and Page, D. C. (1997) 'Functional coherence of the human Y chromosome', *Science,* 278(5338), pp. 675-680.

Lam, I. and Keeney, S. (2015) 'Mechanism and regulation of meiotic recombination initiation', *Cold Spring Harbor perspectives in biology,* 7(1), pp. a016634.

Lane, M., Zander-Fox, D., Hamilton, H., Jasper, M., Hodgson, B., Fraser, M. and Bell, F. (2017a) 'Ability to detect aneuploidy from cell free DNA collected from media is dependent on the stage of development of the embryo', *Fertility and Sterility,* 108(3), pp. e61.

Lane, S. I., Chang, H.-Y., Jennings, P. C. and Jones, K. T. (2010) 'The Aurora kinase inhibitor ZM447439 accelerates first meiosis in mouse oocytes by overriding the spindle assembly checkpoint', *Reproduction,* 140(4), pp. 521-530.

Lane, S. I., Morgan, S. L., Wu, T., Collins, J. K., Merriman, J. A., ElInati, E., Turner, J. M. and Jones, K. T. (2017b) 'DNA damage induces a kinetochore-based ATM/ATR-independent SAC arrest unique to the first meiotic division in mouse oocytes', *Development,* 144(19), pp. 3475-3486.

Lane, S. I., Yun, Y. and Jones, K. T. (2012) 'Timing of anaphase-promoting complex activation in mouse oocytes is predicted by microtubule-kinetochore attachment but not by bivalent alignment or tension', *Development,* 139(11), pp. 1947-1955.

Lara-Gonzalez, P., Pines, J. and Desai, A. 'Spindle assembly checkpoint activation and silencing at kinetochores'. *Seminars in cell & developmental biology*: Elsevier, 86-98.

Lascombes, F. and Pashen, R. (2001) 'Results from embryo freezing and post ovulation breeding in a commercial embryo transfer programme', *EQUINE EMBRYO TRANSFER*, pp. 95.

Lazaraviciute, G., Kauser, M., Bhattacharya, S., Haggarty, P. and Bhattacharya, S. (2014) 'A systematic review and meta-analysis of DNA methylation levels and imprinting disorders in children conceived by IVF/ICSI compared with children conceived spontaneously', *Human reproduction update,* 20(6), pp. 840-852.

Leaver, M. and Wells, D. (2020) 'Non-invasive preimplantation genetic testing (niPGT): the next revolution in reproductive genetics?', *Human reproduction update,* 26(1), pp. 16-42.

Lee, H., Park, C., Na, W., Park, K. H. and Shin, S. (2020) 'Precision cell-free DNA extraction for liquid biopsy by integrated microfluidics', *NPJ precision oncology,* 4(1), pp. 3.

Leemans, B., Gadella, B. M., Stout, T. A., De Schauwer, C., Nelis, H., Hoogewijs, M. and Van Soom, A. (2016) 'Why doesn't conventional IVF work in the horse?: the equine oviduct as a microenvironment for capacitation/fertilization', *Reproduction,* 152(6), pp. 233-245.

Leighton, P. A., Ingram, R. S., Eggenschwiler, J., Efstratiadis, A. and Tilghman, S. M. (1995) 'Disruption of imprinting caused by deletion of the H19 gene region in mice', *Nature,* 375(6526), pp. 34-39.

Lightfoot, D. A., Kouznetsova, A., Mahdy, E., Wilbertz, J. and Höög, C. (2006) 'The fate of mosaic aneuploid embryos during mouse development', *Developmental biology,* 289(2), pp. 384-394.

Lim, D.-S. and Hasty, P. (1996) 'A mutation in mouse rad51 results in an early embryonic lethal that is suppressed by a mutation in p53', *Molecular and cellular biology,* 16(12), pp. 7133-7143.

Liu, D. and Lampson, M. A. (2009) 'Regulation of kinetochore–microtubule attachments by Aurora B kinase', *Biochemical Society Transactions,* 37(5), pp. 976-980.

Liu, W., Liu, J., Du, H., Ling, J., Sun, X. and Chen, D. (2017) 'Non-invasive pre-implantation aneuploidy screening and diagnosis of beta thalassemia IVSII654 mutation using spent embryo culture medium', *Annals of medicine,* 49(4), pp. 319-328.

Lo, Y. D., Corbetta, N., Chamberlain, P. F., Rai, V., Sargent, I. L., Redman, C. W. and Wainscoat, J. S. (1997) 'Presence of fetal DNA in maternal plasma and serum', *The lancet,* 350(9076), pp. 485-487.

Lo, Y. D., Tein, M. S., Lau, T. K., Haines, C. J., Leung, T. N., Poon, P. M., Wainscoat, J. S., Johnson, P. J., Chang, A. M. and Hjelm, N. M. (1998) 'Quantitative analysis of fetal

DNA in maternal plasma and serum: implications for noninvasive prenatal diagnosis', *The American Journal of Human Genetics,* 62(4), pp. 768-775.

Lo, Y. M. D., Chan, K. C. A., Sun, H., Chen, E. Z., Jiang, P., Lun, F. M. F., Zheng, Y. W., Leung, T. Y., Lau, T. K., Cantor, C. R. and Chiu, R. W. K. (2010) 'Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus', *Science Translational Medicine,* 2(61), pp. 61ra91-61ra91.

Lorenz, T. C. (2012) 'Polymerase chain reaction: basic protocol plus troubleshooting and optimization strategies', *JoVE (Journal of Visualized Experiments),* (63), pp. e3998.

Luo, J., Wang, S., Zhang, S., He, Y., Li, S., Han, J., Xu, M. and Deng, G. (2022) 'Performance of ImproGene Cell-Free DNA Tubes for Stabilization and Analysis of cfDNA in Blood Samples', *Fetal and Pediatric Pathology,* 41(5), pp. 771-780.

MacLennan, M., Crichton, J. H., Playfoot, C. J. and Adams, I. R. 'Oocyte development, meiosis and aneuploidy'. *Seminars in cell & developmental biology*: Elsevier, 68-76.

Magli, M. C., Albanese, C., Crippa, A., Tabanelli, C., Ferraretti, A. P. and Gianaroli, L. (2019) 'Deoxyribonucleic acid detection in blastocoelic fluid: a new predictor of embryo ploidy and viable pregnancy', *Fertility and sterility,* 111(1), pp. 77-85.

Mahdavi, S., Karami, F. and Sabbaghi, S. (2019) 'Non-invasive prenatal diagnosis of foetal gender through maternal circulation in first trimester of pregnancy', *Journal of Obstetrics and Gynaecology,* 39(8), pp. 1071-1074.

Manejwala, F. M., Cragoe Jr, E. J. and Schultz, R. M. (1989) 'Blastocoel expansion in the preimplantation mouse embryo: role of extracellular sodium and chloride and possible apical routes of their entry', *Developmental biology,* 133(1), pp. 210-220.

Mani, S. and Mainigi, M. 'Embryo culture conditions and the epigenome'. *Seminars in reproductive medicine*: Thieme Medical Publishers, 211-220.

Mannini, L., Cucco, F., Quarantotti, V., Amato, C., Tinti, M., Tana, L., Frattini, A., Delia, D., Krantz, I. D. and Jessberger, R. (2015) 'SMC1B is present in mammalian somatic cells and interacts with mitotic cohesin proteins', *Scientific reports,* 5(1), pp. 1-11.

Mantikou, E., Wong, K. M., Repping, S. and Mastenbroek, S. (2012) 'Molecular origin of mitotic aneuploidies in preimplantation embryos', *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease,* 1822(12), pp. 1921-1930.

Mantikou, E., Youssef, M. A., van Wely, M., van der Veen, F., Al-Inany, H. G., Repping, S. and Mastenbroek, S. (2013) 'Embryo culture media and IVF/ICSI success rates: a systematic review', *Human reproduction update,* 19(3), pp. 210-220.

Marangos, P. and Carroll, J. (2012) 'Oocytes progress beyond prophase in the presence of DNA damage', *Current Biology,* 22(11), pp. 989-994.

Marangos, P., Stevense, M., Niaka, K., Lagoudaki, M., Nabti, I., Jessberger, R. and Carroll, J. (2015) 'DNA damage-induced metaphase I arrest is mediated by the spindle assembly checkpoint and maternal age', *Nature communications,* 6(1), pp. 8706.

Market-Velker, B., Fernandes, A. and Mann, M. (2010a) 'Side-by-side comparison of five commercial media systems in a mouse model: suboptimal in vitro culture interferes with imprint maintenance', *Biology of reproduction,* 83(6), pp. 938-950.

Market-Velker, B. A., Fernandes, A. D. and Mann, M. R. (2010b) 'Side-by-side comparison of five commercial media systems in a mouse model: suboptimal in vitro culture interferes with imprint maintenance', *Biol Reprod,* 83(6), pp. 938-50.

Markoulatos, P., Siafakas, N. and Moncany, M. (2002) 'Multiplex polymerase chain reaction: a practical approach', *Journal of clinical laboratory analysis,* 16(1), pp. 47-51.

Mateusen, B., Van Soom, A., Maes, D. G., Donnay, I., Duchateau, L. and Lequarre, A.-S. (2005) 'Porcine embryo development and fragmentation and their relation to apoptotic markers: a cinematographic and confocal laser scanning microscopic study', *Reproduction,* 129(4).

Matsuoka, S., Ballif, B. A., Smogorzewska, A., McDonald III, E. R., Hurov, K. E., Luo, J., Bakalarski, C. E., Zhao, Z., Solimini, N. and Lerenthal, Y. (2007) 'ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage', *science,* 316(5828), pp. 1160-1166.

McCue, P. M. and Squires, E. L. (2015) *Equine embryo transfer.* CRC Press.

McGrath, J. and Solter, D. (1983) 'Nuclear transplantation in the mouse embryo by microsurgery and cell fusion', *Science,* 220(4603), pp. 1300-1302.

McGrath, J. and Solter, D. (1984) 'Completion of mouse embryogenesis requires both the maternal and paternal genomes', *Cell,* 37(1), pp. 179-183.

Melzak, K. A., Sherwood, C. S., Turner, R. F. and Haynes, C. A. (1996) 'Driving forces for DNA adsorption to silica in perchlorate solutions', *Journal of colloid and interface science,* 181(2), pp. 635-644.

Menezo, Y. J., Russo, G., Tosti, E., Mouatassim, S. E. and Benkhalifa, M. (2007) 'Expression profile of genes coding for DNA repair in human oocytes using pangenomic microarrays, with a special focus on ROS linked decays', *Journal of assisted reproduction and genetics,* 24, pp. 513-520.

Merriman, J. A., Jennings, P. C., McLaughlin, E. A. and Jones, K. T. (2012) 'Effect of aging on superovulation efficiency, aneuploidy rates, and sister chromatid cohesion in mice aged up to 15 months', *Biology of reproduction,* 86(2), pp. 49, 1-6.

Mestdagh, P., Feys, T., Bernard, N., Guenther, S., Chen, C., Speleman, F. and Vandesompele, J. (2008) 'High-throughput stem-loop RT-qPCR miRNA expression profiling using minute amounts of input RNA', *Nucleic acids research,* 36(21), pp. e143-e143.

Mikwar, M., MacFarlane, A. J. and Marchetti, F. (2020) 'Mechanisms of oocyte aneuploidy associated with advanced maternal age', *Mutation Research/Reviews in Mutation Research,* 785, pp. 108320.

Miura, K., Higashijima, A., Shimada, T., Miura, S., Yamasaki, K., Abe, S., Jo, O., Kinoshita, A., Yoshida, A. and Yoshimura, S. (2011) 'Clinical application of fetal sex determination using cell-free fetal DNA in pregnant carriers of X-linked genetic disorders', *Journal of human genetics,* 56(4), pp. 296-299.

Mohn, F. and Schübeler, D. (2009) 'Genetics and epigenetics: stability and plasticity during cellular differentiation', *Trends in Genetics,* 25(3), pp. 129-136.

Monk, D., Mackay, D. J., Eggermann, T., Maher, E. R. and Riccio, A. (2019) 'Genomic imprinting disorders: lessons on how genome, epigenome and environment interact', *Nature Reviews Genetics,* 20(4), pp. 235-248.

Mouliere, F., Robert, B., Arnau Peyrotte, E., Del Rio, M., Ychou, M., Molina, F., Gongora, C. and Thierry, A. R. (2011) 'High fragmentation characterizes tumour-derived circulating DNA', *PloS one,* 6(9), pp. e23418.

Moynahan, M. E. (2002) 'The cancer connection: BRCA1 and BRCA2 tumor suppression in mice and humans', *Oncogene,* 21(58), pp. 8994-9007.

Mukaida, T., Wada, S., Takahashi, K., Pedro, P., An, T. and Kasai, M. (1998) 'Vitrification of human embryos based on the assessment of suitable conditions for 8-cell mouse embryos', *Human reproduction,* 13(10), pp. 2874-2879.

Munné, S., Lee, A., Rosenwaks, Z., Grifo, J. and Cohen, J. (1993) 'Fertilization and early embryology: Diagnosis of major chromosome aneuploidies in human preimplantation embryos', *Human reproduction,* 8(12), pp. 2185-2191.

Musacchio, A. (2015) 'The molecular biology of spindle assembly checkpoint signaling dynamics', *Current biology,* 25(20), pp. R1002-R1018.

Nagata, S. (2000) 'Apoptotic DNA fragmentation', *Experimental cell research,* 256(1), pp. 12-18.

Nakagawa, S. and FitzHarris, G. (2017) 'Intrinsically defective microtubule dynamics contribute to age-related chromosome segregation errors in mouse oocyte meiosis-I', *Current Biology,* 27(7), pp. 1040-1047.

Nicholls, R. D. and Knepper, J. L. (2001) 'Genome organization, function, and imprinting in Prader-Willi and Angelman syndromes', *Annual review of genomics and human genetics,* 2(1), pp. 153-175.

Nishimura, F., Uno, N., Chiang, P.-C., Kaku, N., Morinaga, Y., Hasegawa, H. and Yanagihara, K. (2019) 'The effect of in vitro hemolysis on measurement of cell-free DNA', *Journal of Applied Laboratory Medicine,* 4(2), pp. 235-240.

Nuffield Council on Bioethics (2017) *Non-invasive Prenatal Testing: Ethical Issues: Short Guide.* Nuffield Council on Bioethics.

O'Driscoll, M., Cerosaletti, K. M., Girard, P.-M., Dai, Y., Stumm, M., Kysela, B., Hirsch, B., Gennery, A., Palmer, S. E. and Seidel, J. (2001) 'DNA ligase IV mutations identified in patients exhibiting developmental delay and immunodeficiency', *Molecular cell,* 8(6), pp. 1175-1185.

O'neill, G., Rolfe, L. and Kaufman, M. (1991) 'Developmental potential and chromosome constitution of strontium-induced mouse parthenogenones', *Molecular Reproduction and Development,* 30(3), pp. 214-219.

O'Connor, C. (2008) 'Meiosis, genetic recombination, and sexual reproduction', *Nature education,* 1(1), pp. 174.

Oktay, K., Turan, V., Titus, S., Stobezki, R. and Liu, L. (2015) 'BRCA mutations, DNA repair deficiency, and ovarian aging', *Biology of Reproduction,* 93(3), pp. 67, 1-10.

Olooto, W. E., Amballi, A. A. and Banjo, T. A. (2012) 'A review of Female Infertility; important etiological factors and management', *J Microbiol Biotech Res,* 2(3), pp. 379-385.

Olova, N., Krueger, F., Andrews, S., Oxley, D., Berrens, R. V., Branco, M. R. and Reik, W. (2018) 'Comparison of whole-genome bisulfite sequencing library preparation strategies identifies sources of biases affecting DNA methylation data', *Genome biology,* 19(1), pp. 1-19.

Pal, N., Huang, Y., Madson, D., Kuster, C., Meng, X., Halbur, P. and Opriessnig, T. (2008) 'Development and validation of a duplex real-time PCR assay for the simultaneous detection and quantification of porcine circovirus type 2 and an internal control on porcine semen samples', *Journal of virological methods,* 149(2), pp. 217-225.

Palazzo, L., Della Monica, R., Visconti, R., Costanzo, V. and Grieco, D. (2014) 'ATM controls proper mitotic spindle structure', *Cell Cycle,* 13(7), pp. 1091-1100.

Palermo, G. D., Schlegel, P. N., Colombero, L. T., Zaninovic, N., Moy, F. and Rosenwaks, Z. (1996) 'Aggressive sperm immobilization prior to intracytoplasmic sperm injection with immature spermatozoa improves fertilization and pregnancy rates', *Human Reproduction,* 11(5), pp. 1023-1029.

Palini, S., Galluzzi, L., De Stefani, S., Bianchi, M., Wells, D., Magnani, M. and Bulletti, C. (2013) 'Genomic DNA in human blastocoele fluid', *Reproductive biomedicine online,* 26(6), pp. 603-610.

Palmerola, K. L., Amrane, S., De Los Angeles, A., Xu, S., Wang, N., de Pinho, J., Zuccaro, M. V., Taglialatela, A., Massey, D. J. and Turocy, J. (2022) 'Replication stress impairs chromosome segregation and preimplantation development in human embryos', *Cell,* 185(16), pp. 2988-3007. e20.

Papathanasiou, S., Markoulaki, S., Blaine, L. J., Leibowitz, M. L., Zhang, C.-Z., Jaenisch, R. and Pellman, D. (2021) 'Whole chromosome loss and genomic instability in mouse embryos after CRISPR-Cas9 genome editing', *Nature communications,* 12(1), pp. 5855.

Pappas, J. J., Toulouse, A. and Bradley, W. E. C. (2013) 'The bisulfite genomic sequencing protocol'.

Pauerova, T., Radonova, L., Kovacovicova, K., Novakova, L., Skultety, M. and Anger, M. (2020) 'Aneuploidy during the onset of mouse embryo development', *Reproduction,* 160(5), pp. 773-782.

Paull, T. T. (2015) 'Mechanisms of ATM activation', *Annual review of biochemistry,* 84, pp. 711-738.

Paulsen, M. and Ferguson-Smith, A. C. (2001) 'DNA methylation in genomic imprinting, development, and disease', *The Journal of pathology,* 195(1), pp. 97-110.

Peters, J. (2014) 'The role of genomic imprinting in biology and disease: an expanding view', *Nature Reviews Genetics,* 15(8), pp. 517-530.

Petit, F. M., Frydman, N., Benkhalifa, M., Le Du, A., Aboura, A., Fanchin, R., Frydman, R. and Tachdjian, G. (2005) 'Could sperm aneuploidy rate determination be used as a predictive test before intracytoplasmic sperm injection?', *Journal of andrology,* 26(2), pp. 235-241.

Petrini, J. H. (1999) 'The Mammalian Mre11-Rad50-Nbs1 Protein Complex: Integration of Functions in the Cellular DNA–Damage Response', *The American Journal of Human Genetics,* 64(5), pp. 1264-1269.

Petropoulos, S., Edsgärd, D., Reinius, B., Deng, Q., Panula, S. P., Codeluppi, S., Reyes, A. P., Linnarsson, S., Sandberg, R. and Lanner, F. (2016) 'Single-cell RNA-seq reveals lineage and X chromosome dynamics in human preimplantation embryos', *Cell,* 165(4), pp. 1012-1026.

Piras, V., Tomita, M. and Selvarajoo, K. (2014) 'Transcriptome-wide variability in single embryonic development cells', *Scientific reports,* 4(1), pp. 7137.

Polz, M. F. and Cavanaugh, C. M. (1998) 'Bias in Template-to-Product Ratios in Multitemplate PCR', *Applied and Environmental Microbiology,* 64(10), pp. 3724-3730.

Porcu, E., Fabbri, R., Seracchioli, R., Ciotti, P. M., Magrini, O. and Flamigni, C. (1997) 'Birth of a healthy female after intracytoplasmic sperm injection of cryopreserved human oocytes', *Fertility and sterility,* 68(4), pp. 724-726.

Potabattula, R., Dittrich, M., Böck, J., Haertle, L., Müller, T., Hahn, T., Schorsch, M., Hajj, N. E. and Haaf, T. (2018) 'Allele-specific methylation of imprinted genes in fetal cord blood is influenced by cis-acting genetic variants and parental factors', *Epigenomics,* 10(10), pp. 1315-1326.

Powell, S. N. and Kachnic, L. A. (2003) 'Roles of BRCA1 and BRCA2 in homologous recombination, DNA replication fidelity and the cellular response to ionizing radiation', *Oncogene,* 22(37), pp. 5784-5791.

Qi, S.-T., Liang, L.-F., Xian, Y.-X., Liu, J.-Q. and Wang, W. (2014) 'Arrested human embryos are more likely to have abnormal chromosomes than developing embryos from women of advanced maternal age', *Journal of ovarian research,* 7(1), pp. 1-9.

Qiu, Y. and Liu, C. (2019) 'Quantitative detection of cell-free fetal DNA in peripheral blood of pregnant women during early pregnancy', *Clinical and Experimental Obstetrics & Gynecology,* 46(4), pp. 611-614.

Raudsepp, T., Finno, C. J., Bellone, R. R. and Petersen, J. L. (2019) 'Ten years of the horse reference genome: insights into equine biology, domestication and population dynamics in the post-genome era', *Animal genetics,* 50(6), pp. 569-597.

Reinders, J. and Paszkowski, J. (2010) 'Bisulfite methylation profiling of large genomes', *Epigenomics,* 2(2), pp. 209-220.

Rienzi, L., Gracia, C., Maggiulli, R., LaBarbera, A. R., Kaser, D. J., Ubaldi, F. M., Vanderpoel, S. and Racowsky, C. (2017) 'Oocyte, embryo and blastocyst cryopreservation in ART: systematic review and meta-analysis comparing slow-freezing versus vitrification to produce evidence for the development of global guidance', *Human reproduction update,* 23(2), pp. 139-155.

Rienzi, L. F., Iussig, B., Dovere, L., Fabozzi, G., Cimadomo, D. and Ubaldi, F. M. 'Perspectives in gamete and embryo cryopreservation'. *Seminars in Reproductive Medicine*: Thieme Medical Publishers, 253-264.

Rijnders, R., Van Der Luijt, R., Peters, E., Goeree, J., Van Der Schoot, C., Ploos Van Amstel, J. and Christiaens, G. (2003) 'Earliest gestational age for fetal sexing in cell-free maternal plasma', *Prenatal Diagnosis: Published in Affiliation With the International Society for Prenatal Diagnosis,* 23(13), pp. 1042-1044.

Rinaldi, V. D., Bolcun-Filas, E., Kogo, H., Kurahashi, H. and Schimenti, J. C. (2017) 'The DNA damage checkpoint eliminates mouse oocytes with chromosome synapsis failure', *Molecular cell,* 67(6), pp. 1026-1036. e2.

Rochette, N. C., Rivera-Colón, A. G., Walsh, J., Sanger, T. J., Campbell-Staton, S. C. and Catchen, J. M. (2023) 'On the causes, consequences, and avoidance of PCR duplicates: Towards a theory of library complexity', *Molecular ecology resources*.

Rødgaard, T., Heegaard, P. M. and Callesen, H. (2015) 'Non-invasive assessment of in-vitro embryo quality to improve transfer success', *Reproductive biomedicine online,* 31(5), pp. 585-592.

Rosner, B. (2011) 'Fundamentals of Biostatistics (The 7th edition)', *Boston, MA: Brooks/Cole*.

Rubio, C., Navarro-Sánchez, L., García-Pascual, C. M., Ocali, O., Cimadomo, D., Venier, W., Barroso, G., Kopcow, L., Bahçeci, M. and Kulmann, M. I. R. (2020) 'Multicenter prospective study of concordance between embryonic cell-free DNA and trophectoderm biopsies from 1301 human blastocysts', *American journal of obstetrics and gynecology,* 223(5), pp. 751. e1-751. e13.

Rubio, C., Rienzi, L., Navarro-Sánchez, L., Cimadomo, D., García-Pascual, C. M., Albricci, L., Soscia, D., Valbuena, D., Capalbo, A. and Ubaldi, F. (2019a) 'Embryonic cell-free DNA versus trophectoderm biopsy for aneuploidy testing: concordance rate and clinical implications', *Fertility and sterility,* 112(3), pp. 510-519.

Rubio, C., Rodrigo, L., Garcia-Pascual, C., Peinado, V., Campos-Galindo, I., Garcia-Herrero, S. and Simón, C. (2019b) 'Clinical application of embryo aneuploidy testing by next-generation sequencing', *Biology of reproduction,* 101(6), pp. 1083-1090.

Rubio, C., Rodrigo, L., Mir, P., Mateu, E., Peinado, V., Milán, M., Al-Asmar, N., Campos-Galindo, I., Garcia, S. and Simón, C. (2013) 'Use of array comparative genomic hybridization (array-CGH) for embryo assessment: clinical results', *Fertility and sterility,* 99(4), pp. 1044-1048.

Rubio, C., Simon, C., Vidal, F., Rodrigo, L., Pehlivan, T., Remohi, J. and Pellicer, A. (2003) 'Chromosomal abnormalities and embryo development in recurrent miscarriage couples', *Human Reproduction,* 18(1), pp. 182-188.

Sabina, J. and Leamon, J. H. (2015) 'Bias in whole genome amplification: causes and considerations', *Whole genome amplification: Methods and protocols*, pp. 15-41.

Sakatani, T., Wei, M., Katoh, M., Okita, C., Wada, D., Mitsuya, K., Meguro, M., Ikeguchi, M., Ito, H. and Tycko, B. (2001) 'Epigenetic heterogeneity at imprinted loci in normal populations', *Biochemical and Biophysical Research Communications,* 283(5), pp. 1124-1130.

Sansinena, M. (2020) 'Assisted reproductive biotechnologies in the horse', *Reproductive Technologies in Animals*: Elsevier, pp. 13-30.

Schulte, M. M., Tsai, J.-h. and Moley, K. H. (2015) 'Obesity and PCOS: the effect of metabolic derangements on endometrial receptivity at the time of implantation', *Reproductive Sciences,* 22(1), pp. 6-14.

Scott, R. T., Upham, K. M., Forman, E. J., Hong, K. H., Scott, K. L., Taylor, D., Tao, X. and Treff, N. R. (2013) 'Blastocyst biopsy with comprehensive chromosome screening and fresh embryo transfer significantly increases in vitro fertilization implantation and delivery rates: a randomized controlled trial', *Fertility and sterility,* 100(3), pp. 697-703.

Shen, K. C., Heng, H., Wang, Y., Lu, S., Liu, G., Deng, C.-X., Brooks, S. and Wang, Y. A. (2005) 'ATM and p21 cooperate to suppress aneuploidy and subsequent tumor development', *Cancer research,* 65(19), pp. 8747-8753.

Shi, X., Zhang, Z., Cram, D. S. and Liu, C. (2015) 'Feasibility of noninvasive prenatal testing for common fetal aneuploidies in an early gestational window', *Clinica Chimica Acta,* 439, pp. 24-28.

Shomper, M., Lappa, C. and FitzHarris, G. (2014) 'Kinetochore microtubule establishment is defective in oocytes from aged mice', *Cell cycle,* 13(7), pp. 1171-1179.

Simon, C., Gutiérrez, A., Vidal, A., de los Santos, M. J., Tarín, J. J., Remohí, J. and Pellicer, A. (1994) 'Outcome of patients with endometriosis in assisted reproduction:

results from in-vitro fertilization and oocyte donation', *Human reproduction,* 9(4), pp. 725-729.

Smith, T., Heger, A. and Sudbery, I. (2017) 'UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy', *Genome research,* 27(3), pp. 491-499.

Somfai, T., Haraguchi, S., Dang-Nguyen, T. Q., Kaneko, H. and Kikuchi, K. (2023) 'Vitrification of porcine immature oocytes and zygotes results in different levels of DNA damage which reflects developmental competence to the blastocyst stage', *Plos one,* 18(3), pp. e0282959.

Soto, M., Raaijmakers, J. A., Bakker, B., Spierings, D. C., Lansdorp, P. M., Foijer, F. and Medema, R. H. (2017) 'p53 prohibits propagation of chromosome segregation errors that produce structural aneuploidies', *Cell reports,* 19(12), pp. 2423-2431.

Spits, C., De Rycke, M., Verpoest, W., Lissens, W., Van Steirteghem, A., Liebaers, I. and Sermon, K. (2006) 'Preimplantation genetic diagnosis for Marfan syndrome', *Fertility and sterility,* 86(2), pp. 310-320.

Squires, E. (2020) 'Current reproductive technologies impacting equine embryo production', *Journal of equine veterinary science,* 89, pp. 102981.

Staessen, C., Platteau, P., Van Assche, E., Michiels, A., Tournaye, H., Camus, M., Devroey, P., Liebaers, I. and Van Steirteghem, A. (2004) 'Comparison of blastocyst transfer with or without preimplantation genetic diagnosis for aneuploidy screening in couples with advanced maternal age: a prospective randomized controlled trial', *Human Reproduction,* 19(12), pp. 2849-2858.

Stigliani, S., Anserini, P., Venturini, P. and Scaruffi, P. (2013) 'Mitochondrial DNA content in embryo culture medium is significantly associated with human embryo fragmentation', *Human reproduction,* 28(10), pp. 2652-2660.

Stigliani, S., Persico, L., Lagazio, C., Anserini, P., Venturini, P. and Scaruffi, P. (2014) 'Mitochondrial DNA in Day 3 embryo culture medium is a novel, non-invasive biomarker of blastocyst potential and implantation outcome', *Molecular human reproduction,* 20(12), pp. 1238-1246.

Stout, T. A. (2012) 'Cryopreservation of equine embryos: current state-of-the-art', *Reproduction in Domestic Animals,* 47, pp. 84-89.

Stukenberg, P. T. and Burke, D. J. (2015) 'Connecting the microtubule attachment status of each kinetochore to cell cycle arrest through the spindle assembly checkpoint', *Chromosoma,* 124, pp. 463-480.

Sullivan, A. E., Silver, R. M., LaCoursiere, D. Y., Porter, T. F. and Branch, D. W. (2004) 'Recurrent fetal aneuploidy and recurrent miscarriage', *Obstetrics & Gynecology,* 104(4), pp. 784-788.

Sun, S.-C., Wang, Z.-B., Xu, Y.-N., Lee, S.-E., Cui, X.-S. and Kim, N.-H. (2011) 'Arp2/3 complex regulates asymmetric division and cytokinesis in mouse oocytes', *PloS one,* 6(4), pp. e18392.

Sunderam, S., Kissin, D. M., Zhang, Y., Jewett, A., Boulet, S. L., Warner, L., Kroelinger, C. D. and Barfield, W. D. (2022) 'Assisted reproductive technology surveillance—United States, 2018', *MMWR Surveillance Summaries,* 71(4), pp. 1.

Swain, J. E. (2019) 'Controversies in ART: can the IVF laboratory influence preimplantation embryo aneuploidy?', *Reproductive biomedicine online,* 39(4), pp. 599-607.

Tabor, A. and Alfirevic, Z. (2010) 'Update on procedure-related risks for prenatal diagnosis techniques', *Fetal diagnosis and therapy,* 27(1), pp. 1-7.

Tan, M., van Tol, H. T., Mokry, M., Stout, T. A. and Roelen, B. A. (2020) 'Microinjection induces changes in the transcriptome of bovine oocytes', *Scientific Reports,* 10(1), pp. 11211.

Tan, S. C. and Yiap, B. C. (2009) 'DNA, RNA, and protein extraction: the past and the present', *Journal of Biomedicine and Biotechnology,* 2009.

Taylor, S., Wakem, M., Dijkman, G., Alsarraj, M. and Nguyen, M. (2010) 'A practical approach to RT-qPCR—publishing data that conform to the MIQE guidelines', *Methods,* 50(4), pp. S1-S5.

Tercyak, K. P., Johnson, S. B., Roberts, S. F. and Cruz, A. C. (2001) 'Psychological response to prenatal genetic counseling and amniocentesis', *Patient education and counseling,* 43(1), pp. 73-84.

Tiegs, A. W., Tao, X., Zhan, Y., Whitehead, C., Kim, J., Hanson, B., Osman, E., Kim, T. J., Patounakis, G., Gutmann, J., Castelbaum, A., Seli, E., Jalas, C. and T. Scott, R. J. (2021) 'A multicenter, prospective, blinded, nonselection study evaluating the predictive value of an aneuploid diagnosis using a targeted next-generation sequencing–based preimplantation genetic testing for aneuploidy assay and impact of biopsy', *Fertility and Sterility,* 115(3), pp. 627-637.

Titus, S., Li, F., Stobezki, R., Akula, K., Unsal, E., Jeong, K., Dickler, M., Robson, M., Moy, F. and Goswami, S. (2013) 'Impairment of BRCA1-related DNA double-strand break repair leads to ovarian aging in mice and humans', *Science translational medicine,* 5(172), pp. 172ra21-172ra21.

Tjoa, M. L., Cindrova-Davies, T., Spasic-Boskovic, O., Bianchi, D. W. and Burton, G. J. (2006) 'Trophoblastic oxidative stress and the release of cell-free feto-placental DNA', *The American journal of pathology,* 169(2), pp. 400-404.

Tobler, K. J., Zhao, Y., Ross, R., Benner, A. T., Xu, X., Du, L., Broman, K., Thrift, K., Brezina, P. R. and Kearns, W. G. (2015) 'Blastocoel fluid from differentiated blastocysts harbors embryonic genomic material capable of a whole-genome deoxyribonucleic acid amplification and comprehensive chromosome microarray analysis', *Fertility and sterility,* 104(2), pp. 418-425.

Traversa, M., Carey, L. and Leigh, D. (2010) 'A molecular strategy for routine preimplantation genetic diagnosis in both reciprocal and Robertsonian translocation carriers', *MHR: Basic science of reproductive medicine,* 16(5), pp. 329-337.

Treff, N. R., Krisher, R. L., Tao, X., Garnsey, H., Bohrer, C., Silva, E., Landis, J., Taylor, D., Scott, R. T. and Woodruff, T. K. (2016) 'Next generation sequencing-based comprehensive chromosome screening in mouse polar bodies, oocytes, and embryos', *Biology of Reproduction,* 94(4), pp. 76, 1-11.

Treff, N. R., Tao, X., Ferry, K. M., Su, J., Taylor, D. and Scott Jr, R. T. (2012) 'Development and validation of an accurate quantitative real-time polymerase chain reaction–based assay for human blastocyst comprehensive chromosomal aneuploidy screening', *Fertility and sterility,* 97(4), pp. 819-824. e2.

Tšuiko, O., Zhigalina, D. I., Jatsenko, T., Skryabin, N. A., Kanbekova, O. R., Artyukhova, V. G., Svetlakov, A. V., Teearu, K., Trošin, A. and Salumets, A. (2018) 'Karyotype of

the blastocoel fluid demonstrates low concordance with both trophectoderm and inner cell mass', *Fertility and Sterility,* 109(6), pp. 1127-1134. e1.

Turan, V. and Oktay, K. (2020) 'BRCA-related ATM-mediated DNA double-strand break repair and ovarian aging', *Human reproduction update,* 26(1), pp. 43-57.

Tveden-Nyborg, P., Alexopoulos, N. I., Cooney, M. A., French, A. J., Tecirlioglu, R. T., Holland, M. K., Thomsen, P. and D'Cruz, N. T. (2008) 'Analysis of the expression of putatively imprinted genes in bovine peri-implantation embryos', *Theriogenology,* 70(7), pp. 1119-1128.

Tyler, C. and Edman, J. C. (2004) 'Down syndrome, Turner syndrome, and Klinefelter syndrome: primary care throughout the life span', *Primary Care: Clinics in Office Practice,* 31(3), pp. 627-648.

Uetake, Y. and Sluder, G. (2010) 'Prolonged prometaphase blocks daughter cell proliferation despite normal completion of mitosis', *Current Biology,* 20(18), pp. 1666-1671.

van den Berg, M. M., van Maarle, M. C., van Wely, M. and Goddijn, M. (2012) 'Genetics of early miscarriage', *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease,* 1822(12), pp. 1951-1959.

Vanderwall, D. K. and Woods, G. L. (2007) 'Embryo transfer and newer assisted reproductive techniques for horses', *Current therapy in large animal theriogenology*, pp. 211-218.

Vera-Rodriguez, M., Chavez, S. L., Rubio, C., Pera, R. A. R. and Simon, C. (2015) 'Prediction model for aneuploidy in early human embryo development revealed by single-cell analysis', *Nature communications,* 6(1), pp. 7601.

Vera-Rodriguez, M., Diez-Juan, A., Jimenez-Almazan, J., Martinez, S., Navarro, R., Peinado, V., Mercader, A., Meseguer, M., Blesa, D. and Moreno, I. (2018) 'Origin and composition of cell-free DNA in spent medium from human embryo culture during preimplantation development', *Human reproduction,* 33(4), pp. 745-756.

Vera-Rodríguez, M., Michel, C.-E., Mercader, A., Bladon, A. J., Rodrigo, L., Kokocinski, F., Mateu, E., Al-Asmar, N., Blesa, D. and Simón, C. (2016) 'Distribution patterns of segmental aneuploidies in human blastocysts identified by next-generation sequencing', *Fertility and Sterility,* 105(4), pp. 1047-1055. e2.

Verlinsky, Y., Ginsberg, N., Lifchez, A., Valle, J., Moise, J. and Strom, C. M. (1990) 'Analysis of the first polar body: preconception genetic diagnosis', *Human reproduction,* 5(7), pp. 826-829.

Vermeiden, J. P. and Bernardus, R. E. (2013) 'Are imprinting disorders more prevalent after human in vitro fertilization or intracytoplasmic sperm injection?', *Fertility and sterility,* 99(3), pp. 642-651.

Vermeulen, J., Derveaux, S., Lefever, S., De Smet, E., De Preter, K., Yigit, N., De Paepe, A., Pattyn, F., Speleman, F. and Vandesompele, J. (2009) 'RNA pre-amplification enables large-scale RT-qPCR gene-expression studies on limiting sample amounts', *BMC research notes,* 2, pp. 1-9.

Vidal, F., Gimenez, C., Rubio, C., Simon, C., Pellicer, A., Santalo, J. and Egozcue, J. (1998) 'FISH preimplantation diagnosis of chromosome aneuploidy in recurrent pregnancy wastage', *Journal of Assisted Reproduction and Genetics,* 15, pp. 310-313.

Viotti, M., Victor, A. R., Barnes, F. L., Zouves, C. G., Besser, A. G., Grifo, J. A., Cheng, E.-H., Lee, M.-S., Horcajadas, J. A. and Corti, L. (2021) 'Using outcome data from one thousand mosaic embryo transfers to formulate an embryo ranking system for clinical use', *Fertility and Sterility,* 115(5), pp. 1212-1224.

Voet, T., Kumar, P., Van Loo, P., Cooke, S. L., Marshall, J., Lin, M.-L., Zamani Esteki, M., Van der Aa, N., Mateiu, L. and McBride, D. J. (2013) 'Single-cell paired-end genome sequencing reveals structural variation per cell cycle', *Nucleic acids research,* 41(12), pp. 6119-6138.

Wagner, A., Blackstone, N., Cartwright, P., Dick, M., Misof, B., Snow, P., Wagner, G. P., Bartels, J., Murtha, M. and Pendleton, J. (1994) 'Surveys of gene families using polymerase chain reaction: PCR selection and PCR drift', *Systematic Biology,* 43(2), pp. 250-261.

Wald, N. J., Lau, K. W., Bestwick, J. P., Old, R. W., Huttly, W. J. and Cheng, R. (2018) 'Specifying a gold standard for the validation of fetal fraction estimation in prenatal screening', *Clinical Chemistry,* 64(9), pp. 1394-1399.

Wallach, E. E., Goshen, R., Ben-Rafael, Z., Gonik, B., Lustig, O., Tannos, V., de-Groot, N. and Hochberg, A. A. (1994) 'The role of genomic imprinting in implantation', *Fertility and sterility,* 62(5), pp. 903-910.

Wang, X. and Pepling, M. E. (2021) 'Regulation of meiotic prophase one in mammalian oocytes', *Frontiers in Cell and Developmental Biology,* 9, pp. 667306.

Wang, Y., Li, R., Yang, R., Zheng, D., Zeng, L., Lian, Y., Zhu, Y., Zhao, J., Liang, X. and Li, W. (2024) 'Intracytoplasmic sperm injection versus conventional in-vitro fertilisation for couples with infertility with non-severe male factor: a multicentre, open-label, randomised controlled trial', *The Lancet,* 403(10430), pp. 924-934.

Watson, A. J. and Kidder, G. M. (1988) 'Immunofluorescence assessment of the timing of appearance and cellular distribution of Na/K-ATPase during mouse embryogenesis', *Developmental biology,* 126(1), pp. 80-90.

Wells, D., Bermudez, M., Steuerwald, N., Thornhill, A., Walker, D., Malter, H., Delhanty, J. and Cohen, J. (2005a) 'Expression of genes regulating chromosome segregation, the cell cycle and apoptosis during human preimplantation development', *Human reproduction,* 20(5), pp. 1339-1348.

Wells, D., Bermúdez, M. G., Steuerwald, N., Malter, H. E., Thornhill, A. R. and Cohen, J. (2005b) 'Association of abnormal morphology and altered gene expression in human preimplantation embryos', *Fertility and sterility,* 84(2), pp. 343-355.

White, C. R., Denomme, M. M., Tekpetey, F. R., Feyles, V., Power, S. G. and Mann, M. R. (2015) 'High frequency of imprinted methylation errors in human preimplantation embryos', *Scientific reports,* 5(1), pp. 17311.

Willems, E., Mateizel, I., Kemp, C., Cauffman, G., Sermon, K. and Leyns, L. (2004) 'Selection of reference genes in mouse embryos and in differentiating human and mouse ES cells', *International Journal of Developmental Biology,* 50(7), pp. 627-635.

Williams, A. B. and Schumacher, B. (2016) 'p53 in the DNA-Damage-Repair Process', *Cold Spring Harbor perspectives in medicine,* 6(5), pp. a026070.

Wreczycka, K., Gosdschan, A., Yusuf, D., Grüning, B., Assenov, Y. and Akalin, A. (2017) 'Strategies for analyzing bisulfite sequencing data', *Journal of biotechnology,* 261, pp. 105-115.

Xi, Y. and Li, W. (2009) 'BSMAP: whole genome bisulfite sequence MAPping program', *BMC Bioinformatics,* 10(1), pp. 232.

Xiao, M., Zhang, S., Liu, Z., Mo, Y., Wang, H., Zhao, X., Yang, X., Boohaker, R. J., Chen, Y. and Han, Y. (2022) 'Dual-functional significance of ATM-mediated phosphorylation of spindle assembly checkpoint component Bub3 in mitosis and the DNA damage response', *Journal of Biological Chemistry,* 298(3).

Xu, J., Fang, R., Chen, L., Chen, D., Xiao, J.-P., Yang, W., Wang, H., Song, X., Ma, T. and Bo, S. (2016) 'Noninvasive chromosome screening of human embryos by genome sequencing of embryo culture medium for in vitro fertilization', *Proceedings of the National Academy of Sciences,* 113(42), pp. 11907-11912.

Xue, L., Cai, J.-Y., Ma, J., Huang, Z., Guo, M.-X., Fu, L.-Z., Shi, Y.-B. and Li, W.-X. (2013) 'Global expression profiling reveals genetic programs underlying the developmental divergence between mouse and human embryogenesis', *BMC genomics,* 14, pp. 1-17.

Yamaguchi, Y., Tayama, C., Tomikawa, J., Akaishi, R., Kamura, H., Matsuoka, K., Wake, N., Minakami, H., Kato, K. and Yamada, T. (2019) 'Placenta-specific epimutation at H19-DMR among common pregnancy complications: its frequency and effect on the expression patterns of H19 and IGF2', *Clinical Epigenetics,* 11(1), pp. 1-13.

Yang, C., Tang, X., Guo, X., Niikura, Y., Kitagawa, K., Cui, K., Wong, Stephen T. C., Fu, L. and Xu, B. (2011) 'Aurora-B mediated ATM serine 1403 phosphorylation is required for mitotic ATM activation and the spindle checkpoint', *Molecular Cell,* 44(4), pp. 597-608.

Yang, M., Rito, T., Naftaly, J., Hu, J., Albertini, D. F., Barad, D. H., Brivanlou, A. H. and Gleicher, N. (2020) 'Self-correction of mosaicism in human self-organizing gastruloids as potential explanation for normal births after transfer of chromosomal-abnormal embryos', *Fertility and Sterility,* 114(3), pp. e14-e15.

Yin, B., Zhang, H., Xie, J., Wei, Y., Zhang, C. and Meng, L. (2021) 'Validation of preimplantation genetic tests for aneuploidy (PGT-A) with DNA from spent culture media (SCM): concordance assessment and implication', *Reproductive Biology and Endocrinology,* 19(1), pp. 41.

Zacchini, F., Sampino, S., Stankiewicz, A. M., Haaf, T. and Ptak, G. E. (2019) 'Assessing the epigenetic risks of assisted reproductive technologies: A way forward', *International Journal of Developmental Biology,* 63(3-4-5), pp. 217-222.

Zeng, Y. and Chen, T. (2019) 'DNA methylation reprogramming during mammalian development', *Genes,* 10(4), pp. 257.

Zhang, H., Wang, J., Qin, Y., Zhang, C., Wang, B. and Wang, Y. (2023) 'Anticipatory Governance of Noninvasive Prenatal Testing for “Non-Medical” Traits: Lessons from Regulation of Medically Assisted Reproduction', *The American Journal of Bioethics,* 23(3), pp. 45-47.

Zhang, Y., Li, N., Wang, L., Sun, H., Ma, M., Wang, H., Xu, X., Zhang, W., Liu, Y. and Cram, D. S. (2016) 'Molecular analysis of DNA in blastocoele fluid using next-generation sequencing', *Journal of assisted reproduction and genetics,* 33, pp. 637-645.

Zhou, L.-q. and Dean, J. (2015) 'Reprogramming the genome to totipotency in mouse embryos', *Trends in cell biology,* 25(2), pp. 82-91.

Zhu, P., Guo, H., Ren, Y., Hou, Y., Dong, J., Li, R., Lian, Y., Fan, X., Hu, B. and Gao, Y. (2018) 'Single-cell DNA methylome sequencing of human preimplantation embryos', *Nature genetics,* 50(1), pp. 12-19.

Zielinska, A. P., Bellou, E., Sharma, N., Frombach, A.-S., Seres, K. B., Gruhn, J. R., Blayney, M., Eckel, H., Moltrecht, R. and Elder, K. (2019) 'Meiotic kinetochores fragment into multiple lobes upon cohesin loss in aging eggs', *Current Biology,* 29(22), pp. 3749-3765. e7.

Zimmermann, B., El-Sheikhah, A., Nicolaides, K., HOLzGREVE, W. and Hahn, S. (2005) 'Optimized real-time quantitative PCR measurement of male fetal DNA in maternal plasma', *Clinical chemistry,* 51(9), pp. 1598-1604.

Zuccaro, M. V., Xu, J., Mitchell, C., Marin, D., Zimmerman, R., Rana, B., Weinstein, E., King, R. T., Palmerola, K. L. and Smith, M. E. (2020) 'Allele-specific chromosome removal after Cas9 cleavage in human embryos', *Cell,* 183(6), pp. 1650-1664. e15.

Appendices

Appendix A

Table 51. Description of blastocysts used in the study of the expression of nine genes with roles in DNA repair. '*' denotes unknown since this information was not provided by the clinic.

Sample ID	Embryo	Day	Karyotype (routine PGT-A)
2.3E	Blastocyst	6	48,XY,+4,+8
2.4E	Blastocyst	5	45,*,-13,-21q21.1q22.3
3.1E	Blastocyst	6	48,XY,+11,+15
3.2E	Blastocyst	5	45,X0,-X
3.3E	Blastocyst	6	45,XY,-10
3.4E	Blastocyst	6	45,*,+7q21.11q36.3,-22
4.1E	Blastocyst	6	45,XY,-13
4.2E	Blastocyst	6	45,XY,-16
4.3E	Blastocyst	6	45,*,+7q21.2q36.3,-16
4.4E	Blastocyst	6	47,*,+8p23.3q13.3-8q21.11q24.3,+16
5.1E	Blastocyst	6	47,XY,+22
5.2E	Blastocyst	5	47,*,+10p12.31q26.3,+21
5.3E	Blastocyst	6	45,*,-16q21q24.2,-15
5.4E	Blastocyst	6	45,XX,-15
6.1 E	Blastocyst	5	47,*,+19,-20q13.2q13.33
6.2E	Blastocyst	6	47,XY,+7
6.3E	Blastocyst	6	48,XX,+9,+19
6.4E	Blastocyst	6	46,XX,-4,+21
7.1E	Blastocyst	6	47,XY,+5
7.2E	Blastocyst	6	45,*,+1p36.32p33,-3
7.3E	Blastocyst	6	45,XX,-11
7.4E	Blastocyst	6	45,*,+8q21.2q24.3,-22
8.1E	Blastocyst	6	47,*,-1q41q44,+9
8.2E	Blastocyst	6	47,XY,+18
8.3E	Blastocyst	6	47,*,+4p16.3q32.3,-4q32.3q35.2,+9
8.4E	Blastocyst	7	45,*,-4p16.3p14,-19
9.1E	Blastocyst	6	47,*,+8q23.1q24.3,+19
9.2E	Blastocyst	5	47,*,+ 6q15q27,+14
9.3E	Blastocyst	6	47,*,-6q22.31q23.1,-6q23.3q27,+12p13.2p11.23,+22
9.4E	Blastocyst	6	47,XY,+8
10.1E	Blastocyst	6	47,*,+11,-18p11.32q12.3,18q12.3q23
10.2E	Blastocyst	6	47,*,+6q21q27,+19
10.3E	Blastocyst	6	45,*,-14,-Xq21.1q28
10.4E	Blastocyst	5	47,*,-13q14.12q14.2,+Y
11.1E	Blastocyst	6	45,XY,-8

11.2E	Blastocyst	5	47,*,-1q21.1q44,+9
11.3E	Blastocyst	7	47,*,-18q22.3q23,+20

Table 52. Description of cleavage stage embryos used in the study of the expression of nine genes with roles in DNA repair. The aneuploidy analysis employed was not able to detect the location of chromosomal breakpoints where there was a segmental loss or gain. Thus, in these results, the segmental abnormality is only indicated as a '(seg)' next to the chromosome where it was detected '*' denotes unknown since this information was not provided by the clinic. CE, cleavage stage.

Sample ID	Embryo	Micro-nucleation	Number of cells	Karyotype (VeriSeq)
1E	CE	No	3	*No result*
2E	CE	Yes	6	*No result*
3E	CE	Yes	8	45,0Y,-1,-2,+3,+13(seg),-X
4E	CE	Yes	5	*No result*
5E	CE	No	3	*No result*
6E	CE	Yes	5	*No result*
7E	CE	No	6	*No result*
8E	CE	Yes	5	52,0Y,+5,+7,-7(seg),-8,-9,+10,+11,+12,+13,+14,+15,-19,+20,-21(seg),+22,-X
9E	CE	Yes	6	45,XY,-17
10E	CE	No	8	*No result*
11E	CE	No	8	*No result*
12E	CE	No	8	48,XX,+1,+2,-3,-9,-11,-12(seg)+13,-13(seg),-15,+16,+18,+20,+21,-22
13E	CE	Yes	8	49,0YY,+2,-4,+5,+10,-18,+19,+20,-X,+Y
14E	CE	Yes	8	*No result*
15E	CE	No	8	*No result*
16E	CE	No	8	*No result*
17E	CE	No	8	*No result*
18E	CE	No	8	52,00,+1,+3(seg),+4,+5,+11,+12,-13,+14,-16,+17,-17(seg),-18,+20,+22,-X,-Y
20E	CE	No	3	*No result*
21E	CE	No	6	*No result*
22E	CE	No	6	44,0X,-6,+8,-8(seg),-9(seg)+10,+11,-14,-15,+16,+18,-18(seg),-19,-20,-21,-X
23E	CE	Yes	7	43,XX,+1,+2,-3,-4(seg),-5,+6,+8,-8(seg),-11,-12,-13(seg),-14,+16(seg),-18,-21,+21(seg)
25E	CE	No	8	*No result*
26E	CE	No	6	*No result*
27E	CE	No	7	*No result*
28E	CE	Yes	8	*No result*
29E	CE	No	7	*No result*
30E	CE	Yes	5	47,XX,-3(seg)+13,-13(seg),+15(seg)
31E	CE	No	6	42,XX,+1,-2,+4,+5,-6,-7,+8,-8(seg),-9,+10,-10(seg),-11,-13,-14,-16,-18,-19,-20(seg),+22
32E	CE	No	8	46,0YY,-3,+15,-X,+Y
33E	CE	Yes	4	*No result*

34E	CE	No	4	*No result*
35E	CE	No	6	*No result*
36E	CE	No	8	*No result*
37E	CE	Yes	8	*No result*
39E	CE	*	*	*No result*
40E	CE	*	*	*No result*
41E	CE	*	*	*No result*
42E	CE	*	*	*No result*
43E	CE	*	*	*No result*
44E	CE	*	*	*No result*

Appendix B

Table 53. Aneuploidy analysis results of parthenotes and three PBs. The aneuploidy analysis employed was not able to detect the location of chromosomal breakpoints where there was a segmental loss or gain. Thus, in these results, the segmental abnormality is only indicated as a '(seg)' next to the chromosome where it was detected.

Sample ID	Sample	Injection	NGS results conclusions
A1	Single parthenote	siATM	20,X
A2	Single parthenote	siATM	20,X
A3	Single parthenote	siATM	20,X
A4	Single parthenote	siATM	20,X
A5	Single parthenote	siATM	20,X
A6	Single parthenote	siATM	19,X,-5,+X(seg)
A7	Single parthenote	siATM	20,X,+1,-7
A8	Single parthenote	siATM	24,X,+5,+6,+10,+16
A9	Single parthenote	siATM	20,X
A10	Single parthenote	siATM	19,X,+1(seg),-14
A11	Single parthenote	siATM	24,X,+1,+3,+5,+17
A12	Single parthenote	siATM	20,X,-5(seg),-8(seg),+10(seg)
A13	Single parthenote	siATM	20,X,-2(seg),-3(seg),-5(seg) ,+9,-19
A14	Single parthenote	siATM	19,X,-11
A15	Single parthenote	siATM	20,X,-12(seg)
A16	Single parthenote	siATM	20,X,+2,-6,+10,-14
A17	Single parthenote	siATM	20,X,-10(seg)
A18	Single parthenote	siATM	20,X
A19	Single parthenote	siATM	20,X,+6(seg)
A20	Single parthenote	siATM	21,X,+18
A21	Single parthenote	siATM	20,X
A22	Single parthenote	siATM	20,X
A23	Single parthenote	siATM	20,X
A24	Single parthenote	siATM	19,X,-11
A25	Single parthenote	siATM	20,X,-9,+15
A26	Single parthenote	siATM	20,X
A27	Single parthenote	siATM	19,X,-10
A28	Single parthenote	siATM	20,X
C1	Single parthenote	Uninjected	*No result*
C2	Single parthenote	Uninjected	20,X,-12(seg),+15(seg)
C3	Single parthenote	Uninjected	20,X
C4	Single parthenote	Uninjected	20,X
C5	Single parthenote	Uninjected	20,X
C6	Single parthenote	Uninjected	20,X
C7	Single parthenote	Uninjected	20,X
C8	Single parthenote	Uninjected	15,XX,-3,-5,-7,-10,-11,-12,+X
C9	Single parthenote	Uninjected	20,X
C10	Single parthenote	Uninjected	20,X

C11	Single parthenote	Uninjected	*No result*
C12	Single parthenote	Uninjected	20,X
C13	One polar body	Uninjected	17,X,-1,+6(seg),-9(seg),-12,-18(seg)
N1	Single parthenote	Scrambled siRNA	20,X,+15(seg)
N2	Single parthenote	Scrambled siRNA	20,X
N3	Single parthenote	Scrambled siRNA	20,X
N4	Single parthenote	Scrambled siRNA	20,X
N5	Single parthenote	Scrambled siRNA	21,X,+14
N6	Single parthenote	Scrambled siRNA	20,X
N7	Single parthenote	Scrambled siRNA	20,X
N8	Single parthenote	Scrambled siRNA	20,X
N9	Single parthenote	Scrambled siRNA	22,XX,-5(seg),+6,-11(seg),+12,+15(seg),+X
N10	Single parthenote	Scrambled siRNA	20,X
N11	Single parthenote	Scrambled siRNA	18,0,-5(seg),-6,-X,
N12	Single parthenote	Scrambled siRNA	20,X
N13	Single parthenote	Scrambled siRNA	19,X,-6
N14	Single parthenote	Scrambled siRNA	20,X
N15	Single parthenote	Scrambled siRNA	21,XX,+X
N16	Single parthenote	Scrambled siRNA	20,X
N17	Single parthenote	Scrambled siRNA	21,X,-1,+6,+8
N18	Single parthenote	Scrambled siRNA	22,X,+2,+15
N19	Single parthenote	Scrambled siRNA	20,X
N20	Single parthenote	Scrambled siRNA	20,X
N21	Single parthenote	Scrambled siRNA	20,X
N22	Single parthenote	Scrambled siRNA	18,0,-13,-X
N23	Single parthenote	Scrambled siRNA	18,X,-1,-7
N24	Single parthenote	Scrambled siRNA	20,X,-3(seg),+13(seg)
N25	One polar body	Scrambled siRNA	20,X
N26	One polar body	Scrambled siRNA	20,X,-9(seg)

Appendix C

Table 54. Description of the SCM samples included for NGS.

Sample ID	Scientist	Processing days	Day of SCM collection	Clinic	PGT-A karyotype of corresponding TE biopsy	Sex	PGT-A classification
I1	C	0	5	Clinic 1	45,XX,-11p15.5p15.1,-11q14.1q25(mos),-22	Female	Aneuploid
I10	C	1	6	Clinic 1	47,XX,+17(mos)	Female	Mosaic
I11	C	1	6	Clinic 1	49,XY,-6q11.1q27(mos),+15,+17,+22	Male	Aneuploid
I13	C	1	5	Clinic 1	49,XX,+12,+16,+22	Female	Aneuploid
I14	C	1	5	Clinic 1	46,XY	Male	Euploid
I15	C	1	5	Clinic 1	45,XX,-3	Female	Aneuploid
I17	C	0	6	Clinic 1	*No result*	*N/A*	*No result*
I18	C	0	5	Clinic 1	46,XY	Male	Euploid
I19	C	1	5	Clinic 1	44,XY,-7(mos),-13(mos)	Male	Mosaic
I2	C	1	5	Clinic 1	45,XY,+7q21.11q36.3(mos),-16	Male	Aneuploid
I20	C	1	5	Clinic 1	45,XX,-14	Female	Aneuploid
I21	C	1	6	Clinic 1	46,XX	Female	Euploid
I22	C	1	6	Clinic 1	50,XY,+Y(mos),+4(mos),+6(mos),+12(mos),-15,+21(mos)	Male	Aneuploid
I23	C	1	6	Clinic 1	46,XX,-14q21.2q32.33	Female	Aneuploid
I24	C	1	6	Clinic 1	46,XY	Male	Euploid
I25	C	1	6	Clinic 1	47,XX,+15	Female	Aneuploid
I26	C	0	5	Clinic 1	46,XX	Female	Euploid
I27	C	0	5	Clinic 1	46,XX,-16q12.1q24.2(mos)	Female	Mosaic
I28	C	0	5	Clinic 1	46,XX	Female	Euploid
I29	C	0	6	Clinic 1	46,XX	Female	Euploid
I30	C	0	6	Clinic 1	47,XY,-5q13.2q35.3(mos),+22	Male	Aneuploid

I31	C	0	6	Clinic 1	46,XX	Female	Euploid
I7	C	0	6	Clinic 1	47,XX,+8	Female	Aneuploid
A02	L	>7	6	Clinic 1	46,XY,-3p11.1q29(mos)	Male	Mosaic
B01	L	>7	6	Clinic 1	45,XX,-21	Female	Aneuploid
B03	L	1	6	Clinic 1	46, XX	Female	Euploid
C03	L	1	6	Clinic 1	46, XX	Female	Euploid
D01	L	>7	6	Clinic 1	46,XY	Male	Euploid
D03	L	1	6	Clinic 1	47, XX,+22	Female	Aneuploid
E01	L	>7	6	Clinic 1	47,XX,+8	Female	Aneuploid
E02	L	>7	6	Clinic 1	46,XY	Male	Euploid
E03	L	1	5	Clinic 1	46, XY	Male	Euploid
F02	L	>7	5	Clinic 1	46,XY	Male	Euploid
H02	L	1	6	Clinic 1	45,XX,-20	Female	Aneuploid
B10	C	3	5	Clinic 2	46,XX	Female	Euploid
B11	C	0	6	Clinic 2	46,XX,+9,-17	Female	Aneuploid
B13	C	0	5	Clinic 2	46, XX	Female	Euploid
B16	C	0	5	Clinic 2	46,XY	Male	Euploid
B2	C	3	5	Clinic 2	50,XX,+10,+11,+13,+17	Female	Aneuploid
B20	C	0	5	Clinic 2	47,XY,+22	Male	Aneuploid
B21	C	0	5	Clinic 2	46,XX	Female	Euploid
B24	C	0	6	Clinic 2	43,XX,-4(mos),-21,-22(mos)	Female	Aneuploid
B25	C	0	6	Clinic 2	48,XX,+12,+22	Female	Aneuploid
B26	C	0	5	Clinic 2	46,XX,+1,-6q14.1,+9q22.1q34.3(mos),+13q31.3q34(mos),-21	Female	Aneuploid
B27	C	0	5	Clinic 2	46,XX,-6q14.1	Female	Aneuploid
B28	C	0	6	Clinic 2	46,XY,-6q14.1	Male	Aneuploid
B29	C	0	5	Clinic 2	46,XX,+9(mos)	Female	Mosaic

B3	C	3	5	Clinic 2	48,XX,+7,+22	Female	Aneuploid
B30	C	0	5	Clinic 2	46,XX	Female	Euploid
B32	C	0	6	Clinic 2	46,XX	Female	Euploid
B33	C	0	5	Clinic 2	*No result*	*N/A*	*No result*
B34	C	3	6	Clinic 2	40,XY,-2(mos),-4(mos),-5(mos),-7(mos),-11,-18(mos),-19(mos),+21	Male	Aneuploid
B35	C	3	6	Clinic 2	*Unamplified*	*N/A*	*No result*
B36	C	3	6	Clinic 2	47,XY,+16	Male	Aneuploid
B37	C	3	6	Clinic 2	46,XY	Male	Euploid
B4	C	3	5	Clinic 2	47,XY,+18	Male	Aneuploid
B40	C	3	5	Clinic 2	46,XX	Female	Euploid
B41	C	3	6	Clinic 2	45,XY,-19(mos)	Male	Mosaic
B42	C	1	6	Clinic 2	46,XY	Male	Euploid
B43	C	3	5	Clinic 2	46,XY,+16,-19(mos)	Male	Aneuploid
B44	C	3	5	Clinic 2	47,XY,-16q12.1q24.2,+21	Male	Aneuploid
B48	C	3	5	Clinic 2	45,XY,-13	Male	Aneuploid
B6	C	3	5	Clinic 2	46,XY,-4q(mos),-20p11.21q13.33(mos)	Male	Mosaic
B7	C	3	5	Clinic 2	46,XY	Male	Euploid
B8	C	3	5	Clinic 2	46,XY,+19,-22	Male	Aneuploid
A04	L	1	5	Clinic 2	46,XX	Female	Euploid
A05	L	1	5	Clinic 2	45,XX,-22	Female	Aneuploid
A06	L	1	5	Clinic 2	*No result*	*N/A*	*No result*
B02	L	>7	5	Clinic 2	46,XY	Male	Euploid
B04	L	1	5	Clinic 2	46,XX	Female	Euploid
B05	L	1	6	Clinic 2	44,XX,-10,-20	Female	Aneuploid
B06	L	1	5	Clinic 2	46,XY	Male	Euploid
C02	L	>7	5	Clinic 2	46,XY	Male	Euploid

C04	L	1	5	Clinic 2	46,XX	Female	Euploid
C05	L	1	5	Clinic 2	*No result*	*N/A*	*No result*
C06	L	1	5	Clinic 2	46,XX	Female	Euploid
D02	L	>7	5	Clinic 2	46,XY,-1q(mos)	Male	Mosaic
D04	L	1	5	Clinic 2	46,XY	Male	Euploid
D06	L	1	5	Clinic 2	46,XY	Male	Euploid
E04	L	1	5	Clinic 2	46,XY,+11p15.5p13(mos)	Male	Mosaic
E06	L	1	6	Clinic 2	46,XY	Male	Euploid
F01	L	>7	5	Clinic 2	46,XX	Female	Euploid
F04	L	1	6	Clinic 2	46,XX,-Xq26.2q28(mos)	Female	Mosaic
F05	L	1	5	Clinic 2	46,XX	Female	Euploid
F06	L	1	5	Clinic 2	46,XX	Female	Euploid
G04	L	1	6	Clinic 2	45,XX,-4p16.3p15.33,-12	Female	Aneuploid
G05	L	1	6	Clinic 2	46,XX	Female	Euploid
G06	L	1	5	Clinic 2	46,XY	Male	Euploid
H03	L	1	6	Clinic 2	46,XY,-5q13.3q23.1(mos),-5q23.1q35.3	Male	Aneuploid
H04	L	1	5	Clinic 2	*XX,+15,-Xq(mos)	Female	Aneuploid
H05	L	1	5	Clinic 2	46,XX,+16,-21	Female	Aneuploid
H06	L	1	5	Clinic 2	45,XX,-22	Female	Aneuploid
S1	C	0	6	Clinic 3	47,XY,+22	Male	Aneuploid
S2	C	0	6	Clinic 3	*No result*	*N/A*	*No result*
S3	C	0	6	Clinic 3	46,XY	Male	Euploid
S4	C	4	5	Clinic 3	46,XY	Male	Euploid
S5	C	4	5	Clinic 3	46,XX	Female	Euploid
S6	C	2	6	Clinic 3	45,XY,-21	Male	Aneuploid
S7	C	2	6	Clinic 3	47,XX,+16	Female	Aneuploid
9P (46,XX control)	C	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*

D05 (47,XX,+18 control)	L	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*
F03 (46,XY,-10p14p12.1 control)	L	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*

Table 55. Libraries' concentrations and NGS quality scores for SCM samples and one control. In this study, mosaicism was disregarded from PGT-A TE biopsy classification due to the inability of the niPGT to detect mosaicism in SCM.

Sample ID	Processing days	Day of SCM collection	Clinic	PGT-A classification	Conc. after WGA (ng/µl)	NGS quality Score
I1	1	5	Clinic 1	Mosaic/Euploid	8.45	0.35
I10	1	5	Clinic 1	Aneuploid	7.53	0.35
I11	1	5	Clinic 1	Aneuploid	8.4	0.48
I13	1	6	Clinic 1	Euploid	25.3	0.21
I14	1	6	Clinic 1	Aneuploid	24.6	0.21
I15	1	6	Clinic 1	Aneuploid	30.9	0.13
I17	1	6	Clinic 1	Euploid	23	0.19
I18	1	6	Clinic 1	Aneuploid	22.6	0.18
I19	0	5	Clinic 1	Euploid	4.13	0.69
I2	0	5	Clinic 1	Mosaic/Euploid	12.4	0.21
I20	0	5	Clinic 1	Euploid	5.48	0.46
I21	0	6	Clinic 1	Euploid	14.6	0.24
I22	0	6	Clinic 1	Aneuploid	2.66	0.68
I23	0	6	Clinic 1	Euploid	15.7	0.25
I24	0	6	Clinic 1	Aneuploid	10.8	0.29
I25	3	5	Clinic 1	Euploid	33.2	0.43
I26	0	6	Clinic 1	Aneuploid	2.08	0.49
I27	0	5	Clinic 1	Euploid	11.3	0.91
I28	0	5	Clinic 1	Euploid	14.9	0.56
I29	3	5	Clinic 1	Aneuploid	10.5	0.21
I30	0	5	Clinic 1	Aneuploid	24	0.15
I31	0	5	Clinic 1	Euploid	18.1	0.26
I7	0	6	Clinic 1	Aneuploid	12.8	0.33
B10	0	6	Clinic 2	Aneuploid	24.6	0.15
B11	0	5	Clinic 2	Aneuploid	2.87	0.4
B13	0	5	Clinic 2	Aneuploid	6.6	0.29
B16	0	6	Clinic 2	Aneuploid	11.6	0.26
B2	0	5	Clinic 2	Mosaic/Euploid	2.58	0.69
B20	3	5	Clinic 2	Aneuploid	3.99	1.37
B21	0	5	Clinic 2	Euploid	7.53	0.26
B24	0	6	Clinic 2	Euploid	18.1	0.17
B25	0	5	Clinic 2	*No result*	3.17	0.72
B26	3	6	Clinic 2	Aneuploid	12.7	0.25
B27	3	6	Clinic 2	*No result*	4.34	0.45
B28	3	6	Clinic 2	Aneuploid	2.16	0.68
B29	3	6	Clinic 2	Euploid	10.3	0.22

B3	3	5	Clinic 2	Aneuploid	3.85	0.35
B30	3	5	Clinic 2	Euploid	4.29	0.3
B32	3	6	Clinic 2	Mosaic/Euploid	2.42	0.57
B33	1	6	Clinic 2	Euploid	14.5	0.3
B34	3	5	Clinic 2	Aneuploid	4.57	1.38
B35	3	5	Clinic 2	Aneuploid	36.2	0.1
B36	3	5	Clinic 2	Aneuploid	2.13	1.04
B37	3	5	Clinic 2	Mosaic/Euploid	4.93	0.78
B4	3	5	Clinic 2	Euploid	1.2	1.31
B40	3	5	Clinic 2	Aneuploid	10.4	0.67
B41	0	5	Clinic 2	Aneuploid	2.66	0.9
B42	1	6	Clinic 2	Mosaic/Euploid	5.09	0.32
B43	1	6	Clinic 2	Aneuploid	11.2	0.22
B44	1	5	Clinic 2	Aneuploid	1.81	1.63
B48	1	5	Clinic 2	Euploid	12.9	0.2
B6	1	5	Clinic 2	Aneuploid	3.04	0.52
B7	0	6	Clinic 2	*No result*	7.4	0.33
B8	0	5	Clinic 2	Euploid	8.56	0.25
S1	0	6	Clinic 3	Aneuploid	22	0.17
S2	0	6	Clinic 3	*No result*	4.71	0.36
S3	0	6	Clinic 3	Euploid	11.5	0.21
S4	4	5	Clinic 3	Euploid	2.28	0.51
S5	4	5	Clinic 3	Euploid	2.72	1.57
S6	2	6	Clinic 3	Aneuploid	11.6	0.29
S7	2	6	Clinic 3	Aneuploid	0.44	0.52
9P (46, XX control)	*N/A*	*N/A*	*N/A*	*N/A*	7.96	0.43

Table 56. Concordance between TE (PGT-A) and SCM (niPGT) karyotypes. The PG-Seq Rapid Non-Invasive PG pipeline and the PG-Find Software are not developed to identify embryo mosaicism accurately from SCM. For this reason, the mosaicism results from the PGT-A results were not considered for the concordance analysis (except for one potential case of reciprocity), neither were the apparent mosaicism results in the chromosome plots produced by the PG-Find Software. Samples that produced a 'no result' for PGT-A analysis were removed.

Sample ID	Clinic	PGT-A karyotype of corresponding TE biopsy	niPGT karyotype of SCM
I10	Clinic 1	47,XX,+17(mos)	46,XX
I11	Clinic 1	49,XY,-6q11.1q27(mos),+15,+17,+22	47,XY,+15
I14	Clinic 1	46,XY	46,XY
I15	Clinic 1	45,XX,-3	46,XX,+6q12q16.1
I18	Clinic 1	46,XY	46,XY
I19	Clinic 1	44,XY,-7(mos),-13(mos)	46,XX
I2	Clinic 1	45,XY,+7q21.11q36.3(mos),-16	45,XY,-16
I20	Clinic 1	45,XX,-14	46,XX
I21	Clinic 1	46,XX	46,XX
I22	Clinic 1	50,XY,+Y(mos),+4(mos),+6(mos),+12(mos),-15,+21(mos)	45,XY,-15
I23	Clinic 1	46,XX,-14q21.2q32.33	46,XX
I24	Clinic 1	46,XY	46,XY
I25	Clinic 1	47,XX,+15	46,XX,+7q11.23q36.3(4 copies)
I26	Clinic 1	46,XX	46,XX
I27	Clinic 1	46,XX,-16q12.1q24.2(mos)	46,XX,+10q11.1q26.3(4 copies)
I28	Clinic 1	46,XX	46,XX
I29	Clinic 1	46,XX	46,XX
I30	Clinic 1	47,XY,-5q13.2q35.3(mos),+22	46,XX
I31	Clinic 1	46,XX	46,XX
I7	Clinic 1	47,XX,+8	47,XX,+8
A02	Clinic 1	46,XY,-3p11.1q29(mos)	46,XY
B01	Clinic 1	45,XX,-21	46,XX
C03	Clinic 1	46, XX	46,XX
D01	Clinic 1	46,XY	46,XY
D03	Clinic 1	47, XX,+22	46,XX
E02	Clinic 1	46,XY	46,XY,-Yp11.32q12,+11p15.5p15.2

E03	Clinic 1	46, XY	46,XY
H02	Clinic 1	45,XX,-20	46,XX
B10	Clinic 2	46,XX	46,XX
B11	Clinic 2	46,XX,+9,-17	46,XX,+5q21.3q35.3
B16	Clinic 2	46,XY	46,XY
B2	Clinic 2	50,XX,+10,+11,+13,+17	46,XX
B20	Clinic 2	47,XY,+22	46,XY
B21	Clinic 2	46,XX	48,XX,-5p15.33q35.3,+8p23.3q11.1,+8q11.1q24.3+12(4 copies),+15,-16
B24	Clinic 2	43,XX,-4(mos),-21,-22(mos)	46,XX
B25	Clinic 2	48,XX,+12,+22	47,XX,+12
B26	Clinic 2	46,XX,+1,-6q14.1,+9q22.1q34.3(mos),+13q31.3q34(mos),-21	45,XX,-21
B27	Clinic 2	46,XX,-6q14.1	46,XX
B28	Clinic 2	46,XY,-6q14.1	47,XXY,+X
B29	Clinic 2	46,XX,+9(mos)	46,XX
B30	Clinic 2	46,XX	46,XX,+Xp22.32p11.3
B32	Clinic 2	46,XX	46,XX
B34	Clinic 2	40,XY,-2(mos),-4(mos),-5(mos),-7(mos),-11,-18(mos),-19(mos),+21	46,XY,-11,+21
B36	Clinic 2	47,XY,+16	47,XY,+8q23.3q24.3,+16
B37	Clinic 2	46,XY	46,XY,+3q22.1q25.32
B4	Clinic 2	47,XY,+18	46,XX
B40	Clinic 2	46,XX	46,XX
B41	Clinic 2	45,XY,-19(mos)	46,XX
B42	Clinic 2	46,XY	46,XY
B44	Clinic 2	47,XY,-16q12.1q24.2,+21	46,XX
B6	Clinic 2	46,XY,-4q(mos),-20p11.21q13.33(mos)	46,XX
B8	Clinic 2	46,XY,+19,-22	46,XX
A04	Clinic 2	46,XX	46,XX
A05	Clinic 2	45,XX,-22	46,XX
B02	Clinic 2	46,XY	46,XY,+6p25.3p23

B04	Clinic 2	46,XX	46,XX
B05	Clinic 2	44,XX,-10,-20	44,XX,-10,-20
B06	Clinic 2	46,XY	42,XY,-10,-12,-14,-17,-18,+20
C02	Clinic 2	46,XY	46,XX
C04	Clinic 2	46,XX	41,XX,-2(both copies),-7,-11,-15
C06	Clinic 2	46,XX	46,XX
D02	Clinic 2	46,XY,-1q(mos)	46,XY,-8q23.3q24.3
D04	Clinic 2	46,XY	46,XX
D06	Clinic 2	46,XY	46,XY
E04	Clinic 2	46,XY,+11p15.5p13(mos)	45,XY,-18
F04	Clinic 2	46,XX,-Xq26.2q28(mos)	46,XX
F06	Clinic 2	46,XX	46,XX
G04	Clinic 2	45,XX,-4p16.3p15.33,-12	46,XX,+4p16.3p15.32
G05	Clinic 2	46,XX	46,XX
G06	Clinic 2	46,XY	46,XX
H03	Clinic 2	46,XY,-5q13.3q23.1(mos),-5q23.1q35.3	46,XY,+5q14.3q23.1,+12q23.3q24.23
H04	Clinic 2	*XX,+15,-Xq(mos)	46,XX
H06	Clinic 2	45,XX,-22	46,XX
S1	Clinic 3	47,XY,+22	46,XY
S3	Clinic 3	46,XY	46,XY
S4	Clinic 3	46,XY	46,XX
S6	Clinic 3	45,XY,-21	47,XY,-1p32.2p13.3,+6p25.1p22.3,+7p22.3,+7p21.3q36.3,+9,+11q13.5q25,+15,-20p13.p11.22
S7	Clinic 3	47,XX,+16	47,XX,+6
9P (46,XX control)	N/A	N/A	46,XX
D05 (47,XX,+18 control)	N/A	N/A	47,XX,+18
F03 (46,XY,-10p14p12.1 control)	N/A	N/A	46,XY,-10p14p12.1

Appendix D

Table 57. List of 46 human DMRs included in the panel for the study methylation perturbations in imprinted loci. Coordinates are from the GRCh37/hg19 genome assembly. This list was originally compiled by Ph.D. Dhruti Babariya.

DMR name	Extent of DMR (genomi coordinates)		Size (bp)	Number of CpGs	Methylation origin	Type of DMR
	Start	Finish				
DIRAS3	Chr1:68,515,433	Chr1:68,517,545	2,112	88	Maternal	Oocyte gDMR
DIRAS3 Ex2	Chr1:68,512,505	Chr1:68,513,486	981	39	Maternal	Oocyte gDMR
ERLIN2	Chr8:37,604,992	Chr8:37,606,088	1,096	37	Maternal	Oocyte gDMR
FAM50B	Chr6:3,849,082	Chr6:3,850,359	1,277	90	Maternal	Oocyte gDMR
FANCC	Chr9:98,075,400	Chr8:98,075,744	344	26	Maternal	Oocyte gDMR
GNAS A/B	Chr20:57,463,265	Chr20:57,465,201	1,936	198	Maternal	Secondary DMR
GNAS-AS1	Chr20:57,425,649	Chr20:57,428,033	2,384	128	Maternal	Oocyte gDMR
GNAS-NESP	Chr20:57,414,039	Chr20:57,418,612	4,573	257	Paternal	Secondary DMR
GNAS-XL	Chr20:57,428,905	Chr20:57,431,463	2,558	200	Maternal	Oocyte gDMR
GPR1-AS	Chr2:207,066,967	Chr2:207,069,445	2,478	86	Maternal	Oocyte gDMR
GRB10	Chr7:50,848,726	Chr7:50,851,312	2,586	171	Maternal	Oocyte gDMR
H19	Chr11:2,018,812	Chr11:2,024,740	5,928	250	Paternal	Sperm gDMR
HTR5A	Chr7:154,862,719	Chr7:154,863,382	663	55	Maternal	Oocyte gDMR
IG-DMR	Chr14:101,275,427	Chr14:101,278,058	2,631	64	Paternal	Sperm gDMR
IGF1R	Chr15:99,408,496	Chr15:99,409,650	1,154	55	Maternal	Oocyte gDMR
IGF2 3'UTR	Chr11:2,153,991	Chr11:2,155,112	1,121	63	Paternal	Secondary DMR
IGF2 DMR0	Chr11:2,168,333	Chr11:2,169,768	1,435	33	Paternal	Sperm gDMR
IGF2R	Chr6:160,426,558	Chr6:160,427,561	1,003	74	Maternal	Oocyte gDMR
INPP5F	Chr10:121,578,046	Chr10:121,578,727	681	52	Maternal	Oocyte gDMR
KCNQ1OT1	Chr11:2,719,948	Chr11:2,722,259	2,311	192	Maternal	Oocyte gDMR
L3MBTL1	Chr20:42,142,365	Chr20:42,144,040	1,675	84	Maternal	Oocyte gDMR

MAGEL2	Chr15:23,892,425	Chr15:23,894,029	1,604	51	Maternal	Secondary DMR
MCTS2P	Chr20:30,134,663	Chr20:30,135,933	1,270	47	Maternal	Oocyte gDMR
MEG3	Chr14:101,290,524	Chr14:101,293,978	3,454	188	Paternal	Secondary DMR
MEG8	Chr14:101,370,741	Chr14:101,371,419	678	43	Maternal	Secondary DMR
MEST	Chr7:130,130,122	Chr7:130,134,388	4,266	226	Maternal	Oocyte gDMR
MKRN3	Chr15:23,807,086	Chr15:23,812,495	5,409	109	Maternal	Oocyte gDMR/secondary DMR
NAP1L5	Chr4:89,618,184	Chr4:89,619,237	1,053	57	Maternal	Oocyte gDMR
NDN	Chr15:23,931,451	Chr15:23,932,759	1,308	108	Maternal	Secondary DMR
NHP2L1	Chr22:42,077,774	Chr22:42,078,873	1,099	63	Maternal	Oocyte gDMR
NNAT	Chr20:36,148,604	Chr20:36,150,528	1,924	135	Maternal	Oocyte gDMR
PEG10	Chr7:94,285,537	Chr7:94,287,960	2,423	119	Maternal	Oocyte gDMR
PEG13	Chr8:141,108,147	Chr8:141,111,081	2,934	193	Maternal	Oocyte gDMR
PEG3	Chr19:57,348,493	Chr19:57,353,271	4,778	221	Maternal	Oocyte gDMR
PLAGL1	Chr6:144,328,078	Chr6:144,329,888	1,810	143	Maternal	Oocyte gDMR
PPIEL	Chr1:40,024,626	Chr1:40,025,540	914	39	Maternal	Oocyte gDMR
RB1	Chr13:48,892,341	Chr13:48,895,763	3,422	195	Maternal	Oocyte gDMR
SNRPN variant 4	Chr15:25,068,564	Chr15:25,069,481	917	19	Maternal	Secondary DMR
SNRPN/SNURF	Chr15:25,200,004	Chr15:25,201,976	1,972	113	Maternal	Oocyte gDMR
VTRNA2-1	Chr5:135,414,802	Chr5:135,416,645	1,843	76	Maternal	Oocyte gDMR
WDR27	Chr6:170,054,504	Chr6:170,055,618	1,114	58	Maternal	Oocyte gDMR
WRB	Chr21:40,757,510	Chr21:40,758,276	766	43	Maternal	Oocyte gDMR
ZDBF2	Chr2:207,114,583	Chr2:207,136,544	21,961	439	Paternal	Sperm gDMR/Secondary DMR
ZNF331 variant 7	Chr19:54,057,086	Chr19:54,058,425	1,339	102	Maternal	Oocyte gDMR
ZNF331 variants 4-6	Chr19:54,040,510	Chr19:54,042,212	1,702	125	Maternal	Oocyte gDMR
ZNF597	Chr16:3,481,801	Chr16:3,482,388	587	29	Maternal	Oocyte gDMR

ZNF597/NAA60	Chr16:3,492,828	Chr16:3,494,463	1,635	76	Paternal	Secondary DMR

Table 58. Additional SNPs identified during the CpGs methylation analysis of DMRs. Coordinates are from the GRCh37/hg19 genome assembly.

Gene/DMR	SNP (ID, coordinates and population frequency)
PEG3	rs3795005 (chr19:57,349,494) (T: 87.93%; G: 12.07%)
NDN	rs3743340 (chr15:23,932,467) (G:50.92%;A:49.07%)
NAP1L5	rs1734800419 (chr4:89,618,918) (G:100%;A:0%)
ZDBF2	rs1362467 (chr2:207,116,449) (A:44.54%;T:0%)
WRB	rs2244352 (chr21:40,757,973) (T:32.49%;G:67.51%)
MKRN3	rs957259435 (chr15:23,810,517) (G:100%;A:0%)
MAGEL2	rs116252228 (chr15:23,893,130) (G:96.94%;A:3.06%)

Table 59. Methylation analysis of 46 DMRs results using 100 ng of commercial gDNA and an equimolar primer pool. '-' denotes not enough coverage for analysis.

DMR	Before deduplication						After deduplication					
	X15X20 cycles			X20X20 cycles			X15X20 cycles			X20X20 cycles		
	Average No. of reads	Average CpGs meth.	SNP Allele freq.	Average No. of reads	Average CpGs meth.	SNP Allele freq.	Average No. of reads	Average CpGs meth.	SNP Allele freq.	Average No. of reads	Average CpGs meth.	SNP Allele freq.
DIRAS3	6,160	56%	*N/A*	3,350	62%	*N/A*	8	73%	*N/A*	6	81%	*N/A*
DIRAS3 Ex2	203	71%	*N/A*	130	67%	*N/A*	3	91%	*N/A*	4	91%	*N/A*
ERLIN2	19	46%	*N/A*	6	59%	*N/A*	0	-	*N/A*	0	-	*N/A*
FAM50B	12,837	45%	*N/A*	15,332	49%	*N/A*	10	28%	*N/A*	18	16%	*N/A*
GNAS A/B	5,191	37%	*N/A*	6,970	37%	*N/A*	5	72%	*N/A*	7	32%	*N/A*
GNAS-AS1	2,001	44%	*N/A*	1,188	50%	*N/A*	7	51%	*N/A*	6	66%	*N/A*
GNAS-NESP	60,106	51%	*N/A*	106,291	53%	*N/A*	16	60%	*N/A*	20	60%	*N/A*
GNAS-XL	44,696	50%	*N/A*	72,277	53%	*N/A*	11	39%	*N/A*	17	46%	*N/A*
GPR1-AS	-	-	*N/A*	3	85%	*N/A*	0	-	*N/A*	2	90%	*N/A*
GRB10	151	79%	*N/A*	51	77%	*N/A*	2	83%	*N/A*	0		*N/A*
H19	9,503	58%	37%	10,245	58%	32%	8	45%	18%	13	50%	19%
HTR5A	17,327	59%	*N/A*	22,728	58%	*N/A*	12	56%	*N/A*	15	63%	*N/A*
IG-DMR	17,390	54%	79%	20,830	53%	80%	8	52%	60%	14	61%	69%
IGF1R	852	37%	*N/A*	399	41%	*N/A*	4	75%	*N/A*	3	33%	*N/A*
IGF2 3'UTR	35,318	64%	*N/A*	56,272	65%	*N/A*	17	64%	*N/A*	21	71%	*N/A*
IGF2 DMR0	6,701	45%	*N/A*	12,797	47%	*N/A*	7	35%	*N/A*	10	43%	*N/A*
IGF2R	-	-	*N/A*	-	-	*N/A*	0	-	*N/A*	0	-	*N/A*
INPP5F	20,453	68%	*N/A*	18,599	69%	*N/A*	11	60%	*N/A*	15	43%	*N/A*
KCNQ1OT1	12,381	44%	*N/A*	15,022	45%	*N/A*	10	28%	*N/A*	10	7%	*N/A*
L3MBTL1	10,577	43%	*N/A*	15,400	47%	*N/A*	8	26%	*N/A*	11	59%	*N/A*
MAGEL2	296	56%	92%	159	60%	92%	5	65%	60%	2	64%	-
MCTS2P	2,442	39%	*N/A*	1,925	45%	*N/A*	6	48%	*N/A*	4	44%	*N/A*
MEG3	1,634	46%	*N/A*	1,641	49%	*N/A*	6	32%	*N/A*	7	30%	*N/A*
MEG8	29,820	40%	*N/A*	101,343	41%	*N/A*	15	48%	*N/A*	27	40%	*N/A*
MEST	47,779	50%	*N/A*	74,423	53%	*N/A*	17	43%	*N/A*	19	41%	*N/A*

MKRN3	994	89%	72%	641	86%	71%	6	83%	-	5	93%	-
NAP1L5	987	52%	64%	1,694	57%	60%	5	60%	80%	6	54%	100%
NDN	10,471	31%	42%	18,027	32%	41%	12	23%	42%	24	29%	31%
NHP2L1	118	33%	*N/A*	91	39%	*N/A*	3	67%	*N/A*	2	8%	*N/A*
NNAT	1,792	67%	*N/A*	1,153	70%	*N/A*	9	67%	*N/A*	9	55%	*N/A*
PEG10	826	37%	*N/A*	660	34%	*N/A*	7	2%	*N/A*	7	31%	*N/A*
PEG3	64,739	44%	26%	161,092	42%	26%	19	37%	10%	19	49%	37%
PEG3_UMI	20,284	44%	*N/A*	23,883	43%	*N/A*	14	27%	*N/A*	17	28%	*N/A*
PLAGL1	190	18%	*N/A*	252	17%	*N/A*	4	2%	*N/A*	5	6%	*N/A*
PPIEL	1,377	48%	*N/A*	786	50%	*N/A*	7	48%	*N/A*	5	40%	*N/A*
RB1	22	82%	*N/A*	19	72%	*N/A*	0		*N/A*	2	50%	*N/A*
SNRPN variant 4	15	59%	*N/A*	4	56%	*N/A*	2	50%	*N/A*	0	-	*N/A*
SNRPN/SNURF	23,788	34%	19%	39,525	39%	21%	13	42%	20%	22	23%	4%
VTRNA2-1	1,650	31%	*N/A*	1,744	34%	*N/A*	7	52%	*N/A*	5	41%	*N/A*
WDR27	10,033	58%	*N/A*	11,995	64%	*N/A*	12	73%	*N/A*	13	75%	*N/A*
WRB	3,648	57%	65%	2,004	58%	67%	4	50%	50%	9	67%	56%
ZDBF2	21,894	58%	39%	28,411	58%	39%	12	61%	86%	12	37%	29%
ZNF331 variant 7	19,133	54%	*N/A*	21,529	58%	*N/A*	12	76%	*N/A*	15	66%	*N/A*
ZNF331 variants 4-6	1,029	44%	*N/A*	1,196	43%	*N/A*	5	72%	*N/A*	4	67%	*N/A*
ZNF597	18,984	49%	*N/A*	20,898	51%	*N/A*	18	35%	*N/A*	24	27%	*N/A*
ZNF597/NAA60	348	70%	*N/A*	149	70%	*N/A*	3	98%	*N/A*	3	87%	*N/A*

Table 60. Methylation analysis of 46 DMRs results using clumps of cells and an equimolar primer pool. '-' denotes not enough coverage for analysis.

DMR	Before deduplication						After deduplication					
	X15X20 cycles			X20X20 cycles			X15X20 cycles			X20X20 cycles		
	Average No. of reads	Average CpGs meth.	SNP Allele freq.	Average No. of reads	Average CpGs meth.	SNP Allele freq.	Average No. of reads	Average CpGs meth.	SNP Allele freq.	Average No. of reads	Average CpGs meth.	SNP Allele freq.
DIRAS3	5,267	15%	*N/A*	5,092	63%	*N/A*	4	8%	*N/A*	7	36%	*N/A*
DIRAS3 Ex2	2,425	74%	*N/A*	357	92%	*N/A*	2	73%	*N/A*	3	88%	*N/A*
ERLIN2	124	100%	*N/A*	16	81%	*N/A*	2	100%	*N/A*	-	-	*N/A*
FAM50B	9	96%	*N/A*	4,524	57%	*N/A*	2	98%	*N/A*	10	51%	*N/A*
GNAS A/B	3,452	0%	*N/A*	1,289	58%	*N/A*	3	0%	*N/A*	4	44%	*N/A*
GNAS-AS1	1,065	86%	*N/A*	3,475	70%	*N/A*	2	89%	*N/A*	3	86%	*N/A*
GNAS-NESP	86,326	7%	*N/A*	33,721	38%	*N/A*	11	5%	*N/A*	12	32%	*N/A*
GNAS-XL	155	16%	*N/A*	25,227	93%	*N/A*	-	-	*N/A*	9	86%	*N/A*
GPR1-AS	-	-	*N/A*	-		*N/A*	-	-	*N/A*	-		*N/A*
GRB10	41	83%	*N/A*	80	83%	*N/A*	-	-	*N/A*	-		*N/A*
H19	1,335	19%	*N/A*	918	52%	*N/A*	3	19%	*N/A*	4	38%	*N/A*
HTR5A	18,192	56%	*N/A*	4,954	84%	*N/A*	6	61%	*N/A*	5	81%	*N/A*
IG-DMR	78,905	58%	27%	18,626	69%	56%	9	50%	22%	7	65%	38%
IGF1R	1,102	42%	*N/A*	250	90%	*N/A*	2	96%	*N/A*	4	100%	*N/A*
IGF2 3'UTR	35,714	72%	*N/A*	19,270	58%	*N/A*	9	73%	*N/A*	13	59%	*N/A*
IGF2 DMR0	3,074	26%	*N/A*	7,599	45%	*N/A*	4	20%	*N/A*	8	4%	*N/A*
IGF2R	-	-	*N/A*	-	-	*N/A*	-		*N/A*	-	-	*N/A*
INPP5F	23,445	4%	*N/A*	27,783	50%	*N/A*	5	3%	*N/A*	10	49%	*N/A*
KCNQ1OT1	14,941	5%	*N/A*	4,014	33%	*N/A*	6	4%	*N/A*	4	0%	*N/A*
L3MBTL1	4	50%	*N/A*	7,131	9%	*N/A*	-		*N/A*	8	5%	*N/A*
MAGEL2	3,339	76%	99%	213	8%	17%	4	78%	100%	3	4%	0%
MCTS2P	1,261	29%	*N/A*	446	40%	*N/A*	2	0%	*N/A*	4	25%	*N/A*
MEG3	7	34%	*N/A*	613	21%	*N/A*	-		*N/A*	3	5%	*N/A*
MEG8	130,352	23%	*N/A*	26,483	18%	*N/A*	17	25%	*N/A*	14	15%	*N/A*
MEST	8,707	0%	*N/A*	42,576	48%	*N/A*	7	0%	*N/A*	15	42%	*N/A*

MKRN3	3,032	94%	73%	1,722	51%	76%	4	100%	100%	4	13%	100%
NAP1L5	584	99%	57%	26	88%	64%	4	97%	-	-	-	-
NDN	30,506	3%	*N/A*	7,327	33%	*N/A*	13	3%	*N/A*	11	26%	*N/A*
NHP2L1	45	89%	*N/A*	126	15%	*N/A*	3	100%	*N/A*	3	25%	*N/A*
NNAT	-		*N/A*	1,302	16%	*N/A*	-		*N/A*	6	1%	*N/A*
PEG10	605	92%	*N/A*	888	18%	*N/A*	3	90%	*N/A*	9	33%	*N/A*
PEG3	56,474	24%	*N/A*	32,894	25%	*N/A*	8	20%	*N/A*	13	35%	*N/A*
PEG3_UMI	39,371	58%	*N/A*	12,293	48%	*N/A*	8	72%	*N/A*	11	38%	*N/A*
PLAGL1	7	14%	*N/A*	-		*N/A*	3	0%	*N/A*	-		*N/A*
PPIEL	3,053	14%	*N/A*	982	54%	*N/A*	4	0%	*N/A*	6	73%	*N/A*
RB1	-		*N/A*	4	94%	*N/A*	-		*N/A*	-		*N/A*
SNRPN variant 4	-		*N/A*	16	53%	*N/A*	-		*N/A*	2	50%	*N/A*
SNRPN/SNURF	2,981	0%	0%	12,523	28%	28%	4	0%	0%	10	38%	45%
VTRNA2-1	4,566	59%	*N/A*	2,198	69%	*N/A*	4	60%	*N/A*	7	71%	*N/A*
WDR27	1,801	1%	*N/A*	3,526	0%	*N/A*	4	0%	*N/A*	6	0%	*N/A*
WRB	7,767	2%	2%	3,681	36%	44%	6	12%	14%	6	50%	50%
ZDBF2	74,365	88%	100%	23,327	59%	49%	12	87%	100%	12	66%	63%
ZNF331 variant 7	199	99%	*N/A*	11,100	54%	*N/A*	-		*N/A*	9	60%	*N/A*
ZNF331 variants 4-6	2,302	93%	*N/A*	1,092	59%	*N/A*	3	90%	*N/A*	5	81%	*N/A*
ZNF597	26,824	91%	*N/A*	15,077	68%	*N/A*	7	87%	*N/A*	14	61%	*N/A*
ZNF597/NAA60	25	32%	*N/A*	296	26%	*N/A*	-	-	*N/A*	3	0%	*N/A*

Table 61. Methylation analysis of 46 DMRs results using clumps of cells and a non-equimolar primer pool. '-' denotes not enough coverage for analysis.

DMR	Before deduplication						After deduplication					
	X15X20 cycles			X20X20 cycles			X15X20 cycles			X20X20 cycles		
	Average No. of reads	Average CpGs meth.	SNP Allele freq.	Average No. of reads	Average CpGs meth.	SNP Allele freq.	Average No. of reads	Average CpGs meth.	SNP Allele freq.	Average No. of reads	Average CpGs meth.	SNP Allele freq.
DIRAS3	4,185	33%	*N/A*	2,175	31%	*N/A*	3	33%	*N/A*	0	-	*N/A*
DIRAS3 Ex2	650	92%	*N/A*	303	93%	*N/A*	2	97%	*N/A*	2	97%	*N/A*
ERLIN2	211	14%	*N/A*	198	24%	*N/A*	0	-	*N/A*	0	-	*N/A*
FAM50B	29,477	92%	*N/A*	3,277	15%	*N/A*	10	95%	*N/A*	5	60%	*N/A*
GNAS A/B	5,413	75%	*N/A*	357	73%	*N/A*	4	81%	*N/A*	2	50%	*N/A*
GNAS-AS1	2,784	0%	*N/A*	18,745	40%	*N/A*	0		*N/A*	6	45%	*N/A*
GNAS-NESP	9,540	94%	*N/A*	13,433	33%	*N/A*	6	96%	*N/A*	8	47%	*N/A*
GNAS-XL	16,168	88%	*N/A*	9,932	4%	*N/A*	4	94%	*N/A*	4	3%	*N/A*
GPR1-AS	-	-	*N/A*	-	-	*N/A*	0	-	*N/A*	0	--	*N/A*
GRB10	1,548	80%	*N/A*	2,312	78%	*N/A*	2	67%	*N/A*	3	67%	*N/A*
H19	47,403	70%	*N/A*	12,269	85%	*N/A*	10	61%	*N/A*	5	83%	*N/A*
HTR5A	254	94%	*N/A*	6,031	34%	*N/A*	0		*N/A*	5	52%	*N/A*
IG-DMR	28,299	78%	63%	65,565	57%	50%	6	100%	100%	7	57%	63%
IGF1R	5,963	53%	*N/A*	2,763	10%	*N/A*	3	89%	*N/A*	3	4%	*N/A*
IGF2 3'UTR	26,953	70%	*N/A*	70,250	73%	*N/A*	8	47%	*N/A*	14	79%	*N/A*
IGF2 DMR0	37,822	32%	*N/A*	38,211	39%	*N/A*	5	52%	*N/A*	9	25%	*N/A*
IGF2R	-	-	*N/A*	-	-	*N/A*	0	-	*N/A*	0		*N/A*
INPP5F	10,778	61%	*N/A*	23,204	63%	*N/A*	5	45%	*N/A*	4	31%	*N/A*
KCNQ1OT1	11,745	42%	*N/A*	65,153	13%	*N/A*	5	42%	*N/A*	9	5%	*N/A*
L3MBTL1	4,898	80%	*N/A*	3,850	16%	*N/A*	3	100%	*N/A*	5	23%	*N/A*
MAGEL2	2,528	10%	12%	7,160	60%	71%	4	0%	0%	5	62%	50%
MCTS2P	3,516	29%	*N/A*	2,637	32%	*N/A*	5	20%	*N/A*	4	49%	*N/A*

MEG3	19,204	54%	*N/A*	6,011	97%	*N/A*	5	19%	*N/A*	6	97%	*N/A*
MEG8	101,556	27%	*N/A*	90,615	18%	*N/A*	16	27%	*N/A*	20	17%	*N/A*
MEST	33,163	68%	*N/A*	31,988	18%	*N/A*	8	63%	*N/A*	11	26%	*N/A*
MKRN3	4,191	90%	71%	3,051	73%	71%	5	95%	-	4	64%	-
NAP1L5	1,626	0%	13%	248	2%	21%	4	1%	0%	2	7%	0%
NDN	18,590	27%	*N/A*	15,268	36%	*N/A*	8	19%	*N/A*	8	22%	*N/A*
NHP2L1	866	19%	*N/A*	2,907	99%	*N/A*	3	30%	*N/A*	6	98%	*N/A*
NNAT	1,739	100%	*N/A*	1,366	18%	*N/A*	6	100%	*N/A*	3	0%	*N/A*
PEG10	10,795	11%	*N/A*	6,370	19%	*N/A*	9	13%	*N/A*	10	11%	*N/A*
PEG3	770	62%	*N/A*	2,891	33%	*N/A*	4	67%	*N/A*	3	0%	*N/A*
PEG3_UMI	6,590	56%	*N/A*	1,408	47%	*N/A*	3	85%	*N/A*	3	32%	*N/A*
PLAGL1	8	38%	*N/A*	24	71%	*N/A*	0		*N/A*	2	0%	*N/A*
PPIEL	4,787	46%	*N/A*	2,913	62%	*N/A*	5	48%	*N/A*	4	81%	*N/A*
RB1	639	99%	*N/A*	94	95%	*N/A*	3	100%	*N/A*	0	-	*N/A*
SNRPN variant 4	269	54%	*N/A*	90	49%	*N/A*	3	78%	*N/A*	0	-	*N/A*
SNRPN/SNURF	12,231	22%	21%	16,773	56%	56%	6	5%	0%	7	59%	63%
VTRNA2-1	3,968	18%	*N/A*	6,743	23%	*N/A*	3	30%	*N/A*	5	25%	*N/A*
WDR27	45,200	78%	*N/A*	51,935	4%	*N/A*	8	67%	*N/A*	12	1%	*N/A*
WRB	18,229	59%	64%	4,680	44%	50%	7	57%	57%	5	22%	20%
ZDBF2	4,008	13%	9%	4,181	61%	46%	4	50%	100%	4	92%	100%
ZNF331 variant 7	8,122	22%	*N/A*	5,273	76%	*N/A*	6	29%	*N/A*	5	100%	*N/A*
ZNF331 variants 4-6	3,007	38%	*N/A*	2,067	61%	*N/A*	4	22%	*N/A*	5	56%	*N/A*
ZNF597	14,373	84%	*N/A*	2,713	88%	*N/A*	6	77%	*N/A*	6	82%	*N/A*

ZNF597/NAA60	2,874	74%	*N/A*	1,444	14%	*N/A*	3	78%	*N/A*	5	26%	*N/A*