Investigation of DNA repair in human oocytes and preimplantation embryos

by

Souraya Jaroudi

Thesis submitted for the degree of Doctor of Philosophy
at University College London

October 2009

Supervisors: Dr. Sioban SenGupta
Dr. Joyce Harper

UCL Centre for PGD, Institute for Women's Health,
University College London

I, Souraya Jaroudi, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

ACKNOWLEDGEMENTS

I would first like to thank my supervisor, Dr. Sioban SenGupta, for providing me with a motivating project, for her continuous support, guidance and generous giving of her time. I would like to thank my secondary supervisor, Dr. Joyce Harper, for her direction and initially welcoming me to the PGD group and Professor Joy Delhanty for her expert advice and for inspiring the researcher in me.

I would like to acknowledge the whole team at the UCL Centre for PGD, the clinical scientists at Chenies Mews, the nurse specialist in PGD, Karen Fordham, and the embryologists and clinicians at the Assisted Conception Unit. I am very grateful to Mr. Paul Serhal who generously funded the microarray project.

I would like to thank Dr. Chris Jones (*Bioinformatics and Microarray Group, UCL)* for his advice on our preliminary study design, Dr. Paul Smyth (*Trinity Centre for Health Sciences, Dublin*) for processing our microarrays and for his generous assistance with the analysis of our raw data and Catherine King (*The Wolfson Institute for Biomedical Research, UCL*) for her help with the assessment of our RNA samples.

Special recognition goes to the PCR team, foremost, Dr. Georgia Kakourou, with whom I did all the RNA work for the microarray experiments. Thank you for keeping me sane during the long and sometimes late-night hours at Chenies Mews! I am extremely grateful to Seema Dhanjal and Thalia Mamas who taught me about clinical PCR for PGD (among many other things). Many thanks go the rest of my colleagues Dr Anna Mantzouratou, Dr. Barbara Smith, Dr. Hanan Sultan, Harita Ghevaria, Leoni Xanthopoulou, Dr. Soha Tashkandi, Dr. Stavros Glentis and Dr. Wael Elmahaishi (in alphabetical order) who have made my work in the lab more enjoyable.

Finally, I would like to thank my family and friends (particularly Dr. Nicoletta Charolidi and Dr. Soha Tashkandi) for their support, patience and understanding, especially during the stressful times. I am very grateful to Dr. Mounir Bohsali for proofreading many chapters of this thesis and to my sister Nadia and cousins Ranna and Muna for helping me check and format specific sections. Special thanks go to aunty Nahida and ammo Wael for their continuous encouragement and care. My deepest gratitude goes to my mom and dad for their financial and moral support throughout the years. Thank you for believing in me - I could not have done this without you.

ABSTRACT

DNA repair genes are expressed in mammalian embryos and in human germinal vesicles, however, little is known about DNA repair in human preimplantation embryos.

This project had three aims: 1) to produce a DNA repair profile of human MII oocytes and blastocysts using expression arrays and identify repair pathways that may be active before and after embryonic genome activation; 2) to design an *in vitro* functional assay that targeted mismatch repair and which could be applied to human oocytes and embryos; 3) to investigate the effect of germline mutations in DNA repair genes on early human embryonic development by initiating a preimplantation genetic diagnosis (PGD) service for these genes.

Microarray profiling showed that all DNA repair pathways were potentially functional in human oocytes and blastocysts. Higher mRNA levels were detected for most repair genes in oocytes compared to blastocysts.

The functional assay for insertion/deletion loops (IDL) and mismatch repair was more sensitive than available methods. Repair was detected for 3, 21 and 24-nucleotide IDLs and single base mismatches. Optimisation is necessary to improve sensitivity of the assay before it can be applied to nuclear extracts from pooled oocytes or embryos.

PGD protocols were developed for two *BRCA1* mutations (c.68-69delAG and c.3339T>G) and one *MSH2* mutation (c.1277-?_1386+?del). One *BRCA1* protocol was clinically applied resulting in the birth of a healthy child. Morphological assessment of embryos from this treatment cycle showed that the presence of the germline mutation adversely affected development between days 3 and 5 post fertilisation. The other two PGD protocols were licensed for treatment by the HFEA.

The availability of donated oocytes and embryos from this PGD service will allow the investigation of repair pathways by real time PCR to examine specific genes selected from the microarray expression profiles and by the application of the functional assay.

ABBREVIATIONS

A	adenine nucleotide base
ACU	assisted conception unit
ADO	allele dropout
AE	amplification efficiency
ALB	alkaline lysis buffer
AP	apurinic/apyrimidic
APEX	AP endonuclease
ATP	adenosine 5'-triphosphate
B	biotin
BER	base excision repair
BLAST	Basic Local Alignment Search Tool
β-ME	beta-mercaptoethanol
B&W	binding and washing
bp	base pair
BSA	bovine serum albumin
C	cytosine nucleotide base
ºC	degrees Celsius
CCC	cell cycle checkpoint
CCCC	cell cycle checkpoint control
cDNA	complementary deoxyribonucleic acid
cRNA	complementary ribonucleic acid
CVS	chorionic villus sampling
dATP	deoxyadenosine triphosphate
DB	dissociation buffer
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
dH_2O	distilled water
DO	dragonfly orange
DSB	double strand break
DSBR	double strand break repair
DM1	Dystrophia Myotonica type 1
DIG	digoxigenin
DIG-UTP	digoxigenin-11-uridine-5′-triphosphate
DMPK	Dystrophia Myotonica protein kinase
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxyribonucleoside triphosphates
DOP-PCR	degenerate oligonucleotide primed - polymerase chain reaction
dsDNA	double stranded deoxyribonucleic acid
ds cDNA	double stranded complementary deoxyribonucleic acid

DTT	dithiothreitol
dTTP	deoxythymidine triphosphate
EDTA	ethylene-di-amine-tetra-acetic acid
EGA	embryonic genome activation
ESC	embryonic stem cell
ESHRE	European Society for Human Reproduction and Embryology
Exo1	exonuclease 1
FAM	6-carboxyfluorescein
FISH	fluorescent *in situ* hybridisation
FN rate	false positive rate
FP rate	false negative rate
F-PCR	fluorescent polymerase chain reaction
g	gram
G	guanine nucleotide base
GI	gastrointestinal
GV	germinal vesicle
h	hour
HCl	hydrochloric acid
hCG	human chorionic gonadotrophin
hESC	human embryonic stem cell
HEX	4,7,2',4',5',7'- hexachloro-6-carboxyfluorescein
HFEA	Human Fertilisation and Embryology Authority
HNPCC	hereditary nonpolyposis colorectal cancer
HR	homologous recombination
ICM	inner cell mass
ICSI	intracytoplasmic sperm injection
i/d	insertion/deletion
IDL	insertion/deletion loop
IVF	*in vitro* fertilisation
Kb	kilobase
KDa	kiloDalton
M	molar
MI	meiosis I
MII	meiosis II
MII oocyte	oocyte at metaphase II
mA	milliampere
MDA	multiple displacement amplification
mESC	mouse embryonic stem cell
MEF	mouse embryonic fibroblast
µg	microgram
mg	milligram
min	minute
µl	microlitre

ml	millilitre
µM	micromolar
mm	millimetre
mM	millimolar
MMR	mismatch repair
MPC	magnetic particle concentrator
mRNA	messenger ribonucleic acid
MSI	microsatellite instability
NE	nuclear extracts
NER	nucleotide excision repair
NHEJ	non-homologous end joining
ng	nanogram
nm	nanometre
nt	nucleotide
OCM	outer cell mass
OD	optical density
Oligo(s)	oligonucleotide(s)
Oligo-dT	oligodeoxythymidylic acid
PAR	poly(ADP-ribose)
PBS	phosphate buffered saline
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PGC	primordial germ cell
PGD	preimplantation genetic diagnosis
PGS	preimplantation genetic screening
PVA	polyvinyl alcohol
RNA	ribonucleic acid
RNase	ribonuclease
ROS	reactive oxygen species
ROX	6-carboxyl-X-rhodamine
RPA	replication protein A
rpm	revolutions per minute
RT-PCR	reverse-transcriptase polymerase chain reaction
s	second
SAP	shrimp alkaline phosphatase
SDS	sodium dodecyl sulphate
siRNA	small interfering ribonucleic acid
S.N.	signal to noise
SNP	single nucleotide polymorphism
SSB	single strand break
SSCP	single-strand conformation polymorphism
ssDNA	single stranded deoxyribonucleic acid
STR	short tandem repeat

T	tyrosine nucleotide base
TAMRA	N, N, N', N'-tetramehtyl-6-carboxyrhodamine
Taq	Thermus aquaticus DNA polymerase
TE	trophectoderm
TET	tetrachlorofluorescein phosphoramidite
Tris	tris(hydroxymethyl)methanamine
U/µl	units per microlitre
UCLH	University College London Hospital
UTP	uridine-5′-triphosphate
UV	ultraviolet
V	Volt
v/v	volume to volume
Vh	Volt-hour
W	Watt
w/v	weight to volume
YY	Yakima yellow
ZP	zona pellucida

Table of Contents

List of Figures

List of Tables

1 INTRODUCTION

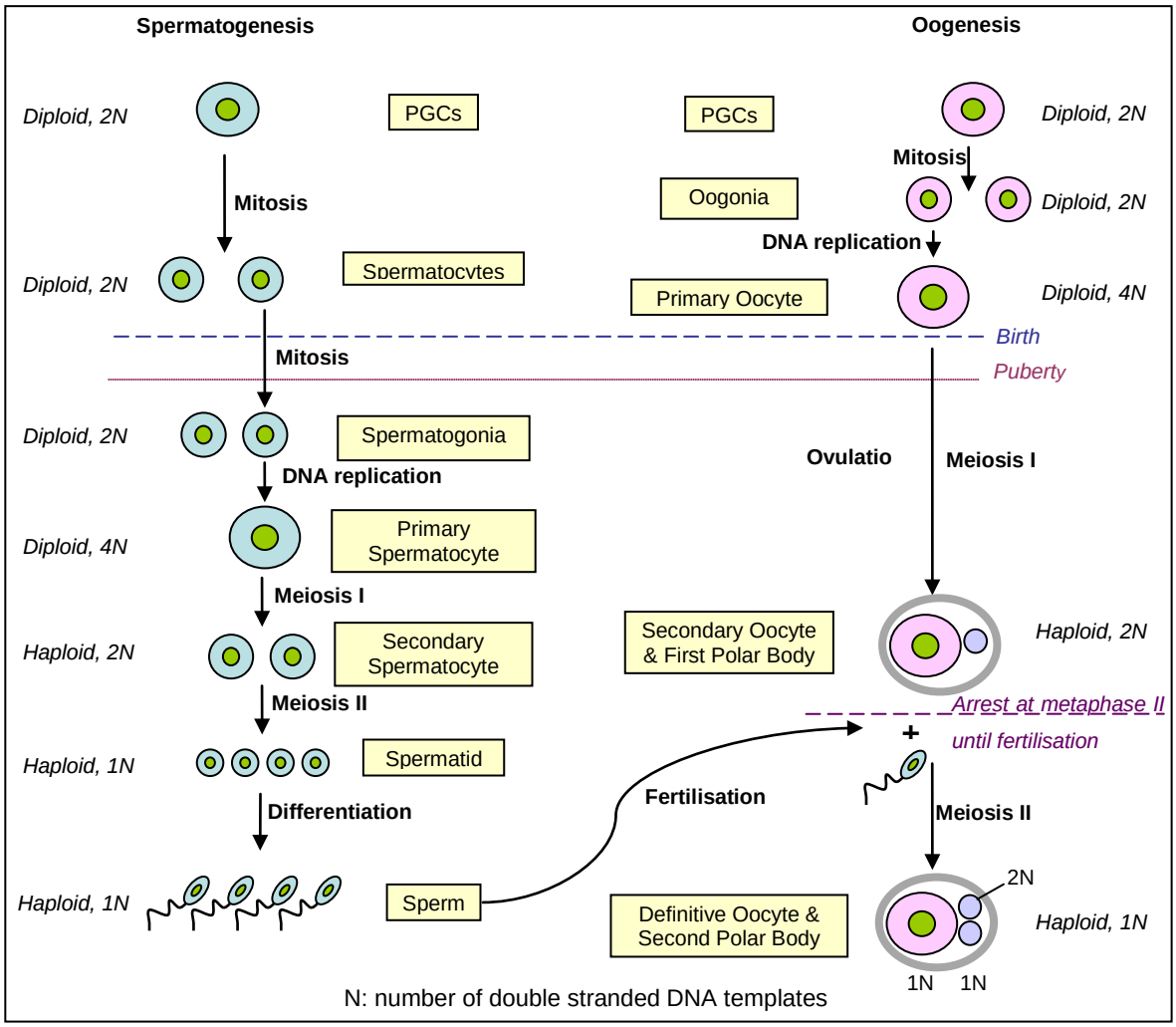
1.1 Early stages of embryonic development

1.1.1 Gametogenesis, fertilisation and the preimplantation embryo

During gametogenesis, a process involving meiosis and cell differentiation, primordial germ cells (PGCs) are converted to mature male and female gametes (spermatozoa and definitive oocytes). In males, PGCs remain dormant during embryonic development until puberty when the seminiferous tubules mature and germ cells differentiate into spermatogonia. Spermatogonia enter mitosis throughout life, but not continuously. Starting at puberty and until death, successive waves of spermatogonia undergo mitosis followed by meiosis and mature into spermatozoa. Each primary spermatocyte yields four spermatozoa. In females, however, PGCs undergo a few mitotic divisions, differentiate into oogonia and begin meiosis during fetal development. Meiotic arrest takes place during an early phase of meiosis I (prophase I) and all cells remain dormant as primary oocytes until puberty. Starting at puberty, a single primary oocyte completes meiosis I and matures into a secondary oocyte (with a first polar body) and is ovulated each month. This oocyte enters a second phase of meiotic arrest and only completes meiosis if fertilisation occurs. Each primary oocyte has the potential to yield one definitive oocyte and two polar bodies. The monthly cycles continue until the female reaches menopause. The processes of spermatogenesis and oogenesis are summarized in Figure 1.1.

Fertilisation involves the fusion of the spermatozoon cell membrane with the oocyte membrane, allowing the sperm nucleus to enter the oocyte. The oocyte then completes meiosis and is called a zygote. The chromosomes of the oocyte and sperm are enclosed in the female and male pronuclei, respectively. However, the membranes quickly disappear as the maternal and paternal chromosomes are replicated in preparation for the first cleavage of the zygote. The maternally inherited chromosomes are decondensed. The paternally inherited genome also decondenses and extensive chromatin remodelling occurs. Programming of both parental chromosome sets must then take place to create the embryonic genome and initiate embryo development (Zheng *et al.*, 2005). During the first days of development, the zygote is split into several smaller daughter cells called blastomeres as it travels down the oviduct without increasing in size. This happens through a series of mitotic divisions, which are not preceded by cell growth.

Figure 1.1: Schematic diagram summarising the major steps in mammalian gametogenesis in male (left) and female (right)



Adapted from Jaroudi and SenGupta (2007)

There are three main transitions in preimplantation development; each is distinguished by changes in gene expression patterns (Zheng *et al.*, 2005). The first transition is the maternal to embryonic transition during which embryonic transcripts replace maternal transcripts. Maternal mRNA transcripts and proteins are laid down in the primary oocyte during follicular growth, when the oocyte is in close contact with granulosa cells that supply ~85% of the occyte's metabolic requirements via gap junctions (Buccione *et al.*, 1990), but transcription stops when the germinal vesicle (GV) undergoes breakdown during meiosis (Bell *et al.*, 2008). The mature oocyte contains enough transcripts and proteins to undergo fertilisation and the first two cell divisions, i.e. until the embryo's genome is activated (Mtango *et al.*, 2008).

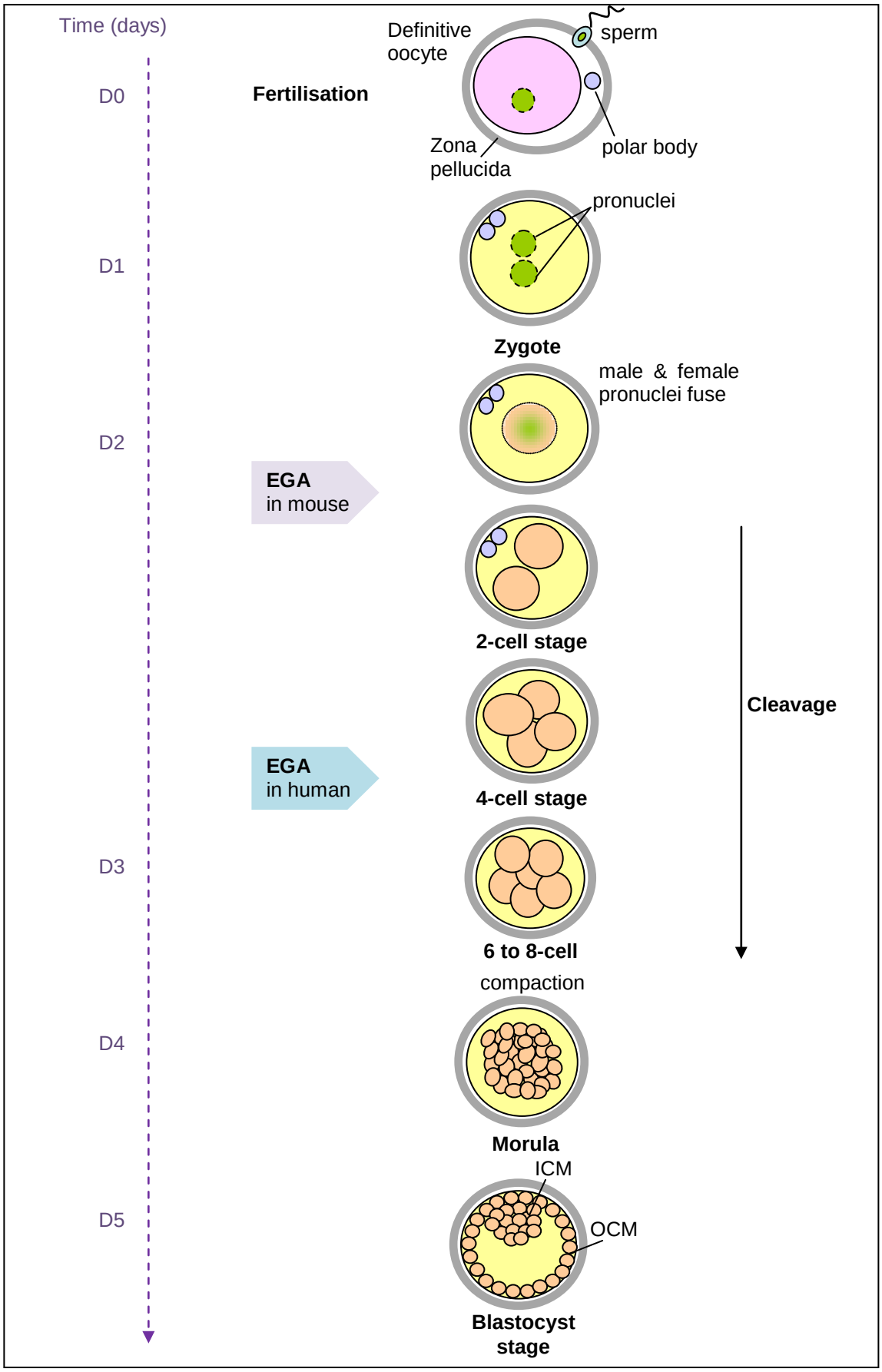
Embryonic genome activation (EGA) begins during the 1-cell stage and is evident by the 2-cell stage in mice (Schultz, 2002). In human preimplantation embryos, global EGA occurs at the 4 to 8-cell stage (Braude *et al.*, 1988; Dobson *et al.*, 2004; Telford *et al.*, 1990; Tesarik *et al.*, 1986; Tesarik *et al.*, 1987; Tesarik *et al.*, 1988); however, Ao *et al.* (1994) showed that paternal Y chromosomal genes are expressed in the zygote and at the 2-cell stage.

Gene expression profiling of mouse preimplantation embryos showed characteristic patterns of maternal RNA depletion and revealed that EGA activation happens in two phases: an initial weak transcription from the new zygotic genome followed by major EGA allowing dramatic reprogramming of expression patterns, which support development to the blastocysts stage, as reviewed by Hamatani *et al.* (2006). The factors controlling chromatin remodelling and reprogramming of the embryonic genome include changes in DNA methylation, histone acetylation, transcription, translation and microRNA regulation, which have been investigated by a large number of studies in recent years and are described in recent reviews by Bell *et al.* (2008), Duranthon *et al.* (2008), Jeanblanc *et al.* (2008), Morgan *et al.* (2005) and Reik *et al.* (2001). Epigenetic reprogramming occurs via active and passive DNA demethylation of the paternal and maternal genomes, respectively (Jeanblanc *et al.*, 2008; Morgan *et al.*, 2005). The precise mechanisms of demethylation are still not understood. Active demethylation takes place right after sperm chromatin is remodelled in the oocyte's cytoplasm, which involves the replacement of protamines by acetylated histones, and is complete prior to DNA replication (Jeanblanc *et al.*, 2008; Reik *et al.*, 2001). This is then followed by further reorganisation by histone modifications (Morgan *et al.*, 2005). The maternal genome undergoes passive demethylation that happens progressively with cell division (Jeanblanc *et al.*, 2008; Morgan *et al.*, 2005). The exact timings of demethylation and remethylation are not determined for the human embryo. In the mouse, active demethylation of the paternal genome occurs in the early cleavage stages and *de novo* methylation takes place in the ICM of the expanded blastocyst (Morgan *et al.*, 2005; Reik *et al.*, 2001).

The second transition in development is compaction at the 8-cell stage (Fleming *et al.*, 2001), which is the first morphological differentiation that occurs during preimplantation development. At approximately the 16-cell stage, the embryo is called a morula. The morula develops a fluid-filled cavity through the formation of tight junctions and gap junctions between blastomeres and is transformed into a blastocyst.

The last transition is blastocyst formation which takes place at the 32 to 64-cell stage (Zheng *et al.*, 2005) and leads to the allocation of some cells to the inside of the developing morula. The centrally placed blastomeres become the inner cell mass (ICM), which gives rise to the embryo proper, and the ones at the periphery constitutes the trophectoderm (TE). The TE forms the extraembryonic tissue that is responsible for the formation of the blastocoele cavity and is essential for the development and differentiation of the ICM (Watson and Barcroft, 2001). Blastocyst formation is the first stage during which differentiation takes place and is characterised by differences in gene expression between the ICM and TE cells (Zheng *et al.*, 2005). In vivo, the embryo would have reached the uterus at this stage. The blastocyst then hatches from the zona pellucida and is able to interact directly with the endometrium resulting in implantation. The main stages of preimplantation development are shown in Figure 1.2.

Figure 1.2: Schematic diagram outlining the main steps of preimplantation embryonic development: MII oocyte, zygote, cleavage stage, morula and blastocyst



Adapted from Jaroudi and SenGupta (2007)

1.1.2 *In vitro* fertilised human preimplantation embryos

The preimplantation period encompasses the time of oocyte fertilisation until the embryo reaches the blastocyst stage and is ready to implant. IVF derived embryos have diverse morphological qualities and are graded according to their cell number (growth/cleavage rate), blastomere symmetry and size uniformity, and level of fragmentation (Bolton *et al.*, 1989). Early preimplantation embryos may stop developing or arrest at various stages due to biochemical and/or cytogenetic problems of maternal or paternal origins. These problems occur both *in vitro* and *in vivo*; however, the *in vitro* environment (i.e. the culture media) is an additional factor that can affect development (Devreker and Englert, 2000; Menezo, 2006). Most preimplantation embryos with abnormalities that would not allow sustained development undergo arrest at time of EGA; embryos reaching the blastocyst stage must therefore possess the appropriate biochemical/genetic tools to activate the embryonic genome (Menezo, 2004). Only about half of human embryos obtained after *in vitro* fertilisation (IVF) reach the blastocyst stage by day 6 (post fertilisation). Blastocyst formation varies between 48% and 65%, depending on the laboratory, the culture media and patient background (Thomas *et al.*, 2009). Increased cell fragmentation in human IVF embryos is associated with reduced blastocyst formation and implantation rates (Hardy *et al.*, 2003) as well as increased chromosomal abnormalities (Moayeri *et al.*, 2008; Ziebe *et al.*, 2003). Therefore, other than suboptimal culture conditions and intrinsic defects in the oocyte or embryo, chromosomal abnormalities explain the high rate of embryonic arrest.

A large proportion of human preimplantation embryos (40%-60%) at different stages of development exhibit chromosomal abnormalities (Daphnis *et al.*, 2008; Munne *et al.*, 1995; Munne *et al.*, 1998) and chromosome breaks, which can persist to the blastocyst stage (Daphnis *et al.*, 2008). A recent study by Vanneste *et al.* (2009) has identified chromosome instability during human cleavage stage embryogenesis as the leading cause of constitutional chromosomal disorders. Mosaicism, aneuploidies, uniparental disomies, frequent segmental deletions, duplications and amplifications that were reciprocal in sister blastomeres (suggesting the occurrence of breakage-fusion-bridge cycles) were all commonly observed in cleavage stage human embryos (Vanneste *et al.*, 2009).

A study by Daphnis *et al.* (2008) showed that embryos with chromosomal abnormalities detected on day 3 post fertilisation seem to have a poor developmental fate as they exhibit increased chromosomal abnormalities or arrest by day 5, however, those that appeared to be normal on day 3 were still normal on day 5. This highlights the developmental importance of day 3 (around the time of EGA) in human preimplantation embryos. Chromosomal abnormalities may result from inappropriate mRNA transcripts or protein supplies (laid down by the oocyte), but they can also lead to imbalances in mRNA transcription levels and protein synthesis which can affect subsequent preimplantation development.

The chromosomal instability observed in human cleavage stage embryos is higher than would be expected based on current IVF implantation and pregnancy rates (as these rates are greater than the rates of chromosomally normal embryos) (Vanneste *et al.*, 2009). It is possible that the chromosomally mosaic embryos that contain normal blastomeres may "self-correct" and become chromosomally normal blastocysts or foetuses (Li *et al.*, 2005; Munne *et al.*, 2005; Staessen *et al.*, 2004a). A recent study by Fragouli *et al.* (2008) showed that the aneuploidy rate in human blastocysts (38.8%) is significantly lower than in embryos at earlier stages (51%).

1.2 DNA damage detection and repair

DNA repair is responsible for protecting the genome's integrity from endogenous metabolites or exogenous agents causing DNA damage. DNA repair enzymes are present in all organisms examined to date (Brendel *et al.*, 1997). Their evolutionary conservation confirms their necessity for the normal function and reproduction of living organisms (Taylor and Lehmann, 1998). Currently, over 150 human DNA repair genes have been cloned and sequenced (Ronen and Glickman, 2001; Wood *et al.*, 2001; Wood *et al.*, 2005). These genes code for DNA repair enzymes, which are involved in the cellular response to DNA damage or are known to be mutated in human diseases associated with DNA sensitivity; not all of them have been assigned functions yet.

Cells are continuously subjected to different types of DNA damage; in fact, due to normal cellular metabolism alone, each human cell is expected to experience approximately 2×10^4 lesions per day (Ames and Shigenaga, 1992). Several DNA repair proteins act together in elaborate cellular pathways in order to detect and repair replication errors and transient or accumulated damage caused by genotoxic stress. Activation of the appropriate pathway is made possible via transient cell cycle arrest. The main DNA repair pathways active in mammalian cells are: base excision repair (BER), double strand break (DSB) repair (DSBR), mismatch repair (MMR), nucleotide excision repair (NER) and post-replication repair (PRR). Alkylation damage and UV-induced pyrimidine dimers can be repaired in a single step by a few specialised enzymes that carry out direct reversal of damage (Table 1.1).

Various DNA repair proteins target different types of lesions (based on their substrate specificities) in order to trigger the necessary repair pathway. There is some degree of redundancy between the different proteins and pathways. Therefore, the absence of a given DNA repair protein (that is crucial for one repair pathway) may be overcome by another fully functional DNA repair pathway. Table 1.1 lists the main DNA repair pathways and their commonly targeted substrates.

Table 1.1: DNA repair pathways and their main substrates

DNA repair pathway	Main damage recognition proteins	Target DNA lesions
BER	DNA glycosylases (Table A.1, Appendix)	Mismatches 8-oxoguanine 3-methyladenine
Direct reversal of damage	MGMT AGT ABH2 ABH3	O^6-methylguanine O^6-alkylguanine 1-methyladenine 3-methylcytosine
DSBR: HR NHEJ & HR	RAD52 KU70 (G22P1)-KU80 (XRCC5) PARP-1	Crosslinks, DSBs DSBs SSBs
NER	XPA-RPA & XPC-HR23B	Pyrimidine dimmers Bulky crosslink adducts
MMR	MSH2-MSH6 MSH2-MSH3	Mismatched or damaged bases, small IDLs Larger IDLs

IDL: insertion/deletion loop; DSB: double strand break; SSBs: single-strand breaks

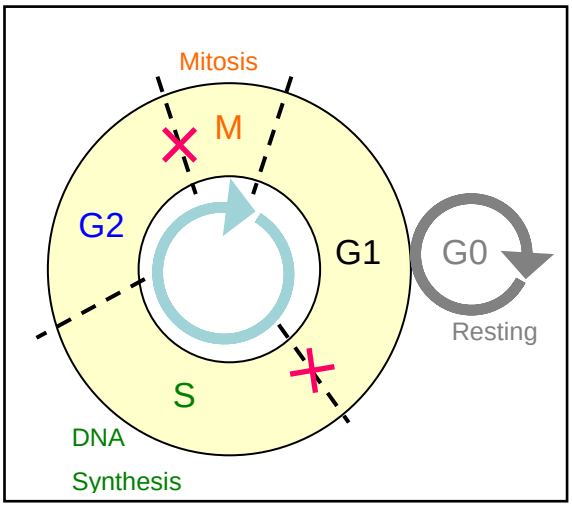
The mechanisms of mammalian DNA repair have been reviewed extensively (Christmann *et al.*, 2003; Cline and Hanawalt, 2003; Hakem, 2008; Sancar *et al.*, 2004). There are three steps common to the three excision repair pathways (BER, NER and MMR): 1) the DNA lesion is detected; 2) a DNA endonulease (BER and NER) or exonuclease (MMR) excises a fragment on the targeted strand containing the lesion; 3) a DNA polymerase re-synthesises a new copy to fill the gap (based on sequence of complementary strand) and a DNA ligase seals the remaining nick.

1.2.1 DNA damage induced cell cycle checkpoints

Coordination of the cell cycle is crucial in the response to DNA damage. Cell cycle arrest occurs at the G1/S or G2/M checkpoints and during replication (Figure 1.3), allowing DNA repair to take place before commencing DNA synthesis or committing to cell division (Lukas *et al.*, 2004) . All four major DNA repair pathways (NER, BER, MMR and DSB repair) along with genotoxic stress sensors interact with components of the cell cycle machinery. Furthermore, some DNA damage sensors trigger a cell cycle checkpoint that activates the apoptotic cell death pathway (Vinson and Hales, 2002). Although the roles of cell cycle machinery involved in the response to genotoxic stress have been studied extensively in cancer models, little is known about their regulation or activity during development (Vinson and Hales, 2002). In fact, fluorescent *in situ* hybridisation (FISH) analysis of cleavage stage human embryos suggested that cell cycle checkpoints may not function during early cleavage development (Delhanty and Handyside, 1995; Harrison *et al.*, 2000).

MMR has been suggested to be part of the G2 checkpoint following some types of DNA damage in human cancer cell lines (Carethers *et al.*, 1996; Hawn *et al.*, 1995) and single primary mouse embryonic fibroblasts (Marquez *et al.*, 2003). MMR can be involved in checkpoint functions in at least two ways: MMR proteins can interact directly with checkpoint proteins to activate cell cycle arrest (Hawn *et al.*, 1995) or MMR is required to process certain types of DNA damage into a form that activates the checkpoint response (Carethers *et al.*, 1996). However, other models suggest that neither interaction with nor activation of MMR complexes are necessary for the G2/M checkpoint to function (Aquilina *et al.*, 1999). Control of DNA repair throughout the cell cycle was recently reviewed by Branzei and Foiani (2008).

Figure 1.3: Diagram showing different phases of the cell cycle (Jaroudi and SenGupta, 2007).



The critical checkpoints are marked with a red cross: the G1/S checkpoint between the first gap phase (G1) and the DNA synthesis phase (S) and the G2/M checkpoint between the second gap phase (G2) and mitosis (M). The cell may arrest at G1 or G2 preventing its commitment to DNA synthesis or mitosis respectively.

1.2.2 DNA repair pathways

1.2.2.1 Base excision repair (BER)

The BER pathway removes damaged DNA bases arising spontaneously within the cell during inflammatory responses and exposure to exogenous agents including radiation. The primary substrates for BER are oxidised DNA bases induced by reactive oxygen species (ROS). BER operates via two pathways that involve different repair enzymes and result in the removal and replacement of one (short-patch BER) or several (long-patch BER) nucleotides (reviewed by Robertson *et al.* (2009)).

Substrate-specific DNA glycosylases (Table A.1, Appendix), like the 8-oxoguanine DNA glycosylase (OGG1) and uracil DNA glycosylase (UNG) that initiate repair of damage induced by ROS and deamination, respectively, remove the damaged bases from the DNA leaving an apurinic/apyrimidic (AP) site that is subsequently processed in steps involving AP endonuclease (APEX), DNA polymerase β and δ or ε, flap endonuclease (FEN1) stimulated by the proliferating cell nuclear antigen (PCNA) and the DNA ligases 1 or 3.

Functional BER is required for early development since knockout mice for essential BER proteins (such as DNA polymerase β, Apex, DNA ligase 1 and Fen1) are embryonic lethal. This may explain the absence of human genetic syndromes related to complete loss of BER proteins (Park and Gerson, 2005). However, biallelic-inherited mutations in the human MYH gene, which encodes for a glycosylase that recognises an adenines misincorporated opposite 7,8-dihydro-8-oxoguanineresidue, have been linked to increased susceptibility to colorectal cancer (Jones *et al.*, 2002; Lipton *et al.*, 2003; Sampson *et al.*, 2005).

1.2.2.2 Double strand break repair (DSBR)

DNA double-strand breaks are among the most dangerous inducers of genotoxic damage and cell death via apoptosis (Rich *et al.*, 2000). Failure to rejoin breaks properly can lead to loss of chromosome segments or rearrangements (Dikomey *et al.*, 1998; Pfeiffer *et al.*, 2000). Furthermore, interactions between two ends from different DSBs can generate tumorigenic chromosome translocations (Aten *et al.*, 2004). DNA DSBs occur when the sugar backbones of both strands are broken close enough to disrupt Watson-Crick pairings and release two DNA ends (Bassing and Alt, 2004). DSBR genes are ubiquitously expressed and highly conserved across eukaryotes (Bassing and Alt, 2004).

Two sub pathways are involved in the repair of DSBs: homologous recombination (HR), which is error-free and active in late S and G2-M cells where a sister chromatid would act as a template (Johnson and Jasin, 2000; Takata *et al.*, 1998), and non-homologous end-joining (NHEJ), which is error-prone and the predominant pathway in mammalian cells at G0/G1 stages of the cell cycle (Cromie *et al.*, 2001; Haber, 2000; Johnson and Jasin, 2000).

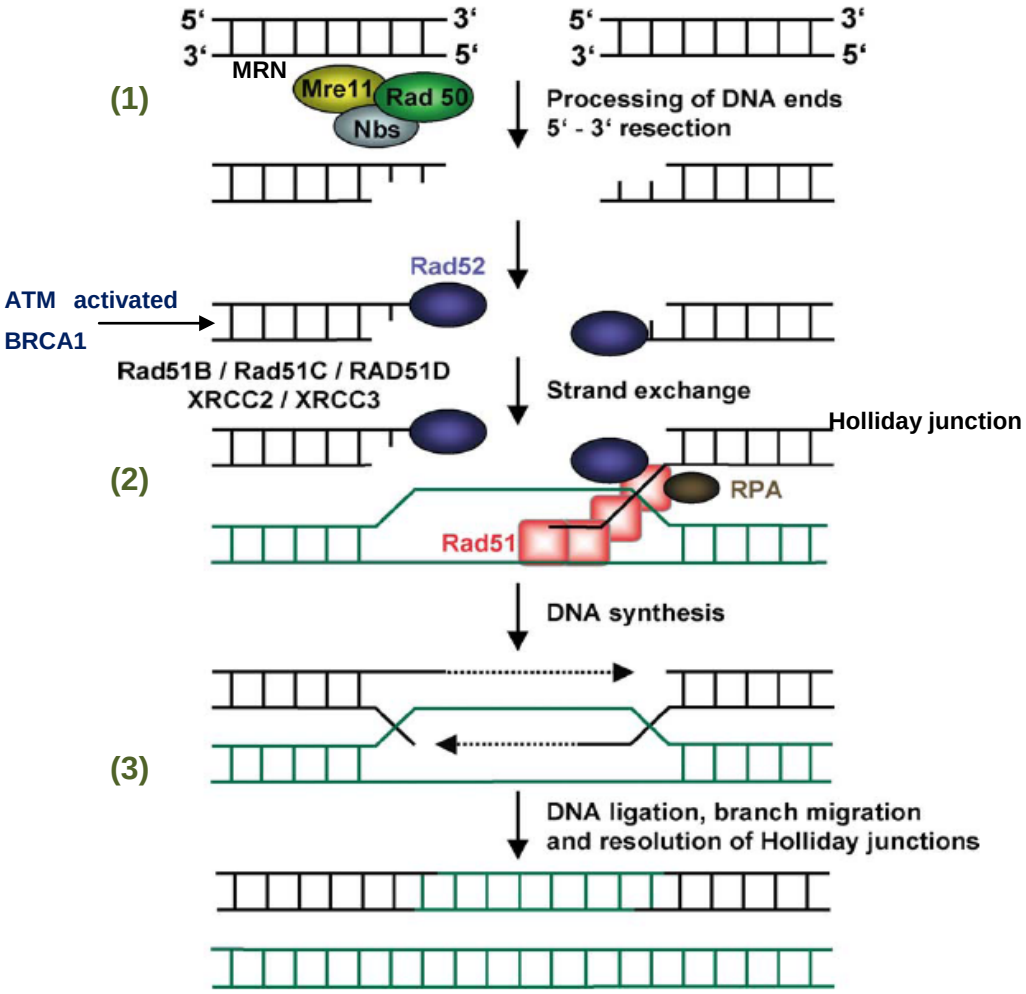
Although HR could use the homologous chromosome in G1 cells, this mechanism may be inhibited in somatic cells during the G1 phase to minimise loss of heterozygosity (Bassing and Alt, 2004). Defective HR repair leads to high incidence of leukemia, breast-ovarian cancers, Werner's and Bloom's syndromes and premature aging (Valerie and Povirk, 2003). Defective NHEJ repair is associated with severe combined immunodeficiency (SCID) (Park and Gerson, 2005).

1.2.2.2.1 Homologous Recombination (HR) Repair

During HR repair, the damaged chromosome comes in close contact with an undamaged DNA molecule which has a homologous sequence and uses it as a template for repair (Sonoda *et al.*, 2001). HR is initiated by binding of the MRE11-RAD50-NBS1 (MRN) complex to the damaged DNA fragment; this promotes a nucleolytic resection of the DSB in the 5'→3' direction and reveals single-stranded DNA ends. Subsequently, RAD51 and the single-strand binding replication protein A (RPA) bind to RAD52 and the single stranded DNA (ssDNA) fragments to form a nucleoprotein filament (Kagawa *et al.*, 2001; Park *et al.*, 1996).

BRCA1, which is activated by ATM, helps RAD51 and BRCA2 bind to the DNA overhang and attract RAD52, facilitated by BLM/WRN (Valerie and Povirk, 2003). Binding of the RAD52 protein heptameric ring complex (Singleton *et al.*, 2002; Stasiak *et al.*, 2000) to the ssDNA ends protects against exonucleolytic digestion (Christmann *et al.*, 2003). The formed nucleoprotein filament invades the duplex template DNA and forms a joint molecule stimulating the strand exchange activity of RAD51 (Baumann and West, 1999; Benson *et al.*, 1998; New *et al.*, 1998; Van Dyck *et al.*, 2001). RAD51 catalyses strand-exchange events with the complementary strand, displacing one strand as a D-loop (Baumann and West, 1997). The strands are extended by a polymerase and sealed by DNA ligase 1 (Bassing and Alt, 2004). This is followed by branch migration and resolution of the Holliday junctions by DNA cleavage and ligation which releases the repaired DNA molecule (Singleton and Jeggo, 1999). The mechanism of DSBR via HR is summarised in Figure 1.4.

Figure 1.4: Diagram outlining the main steps involved in DSBR via homologous recombination



Adapted from Toxicology, 193/1-2, Christmann et al., Mechanisms of human DNA repair, pages 3-32, Copyright (2003), with permission from Elsevier

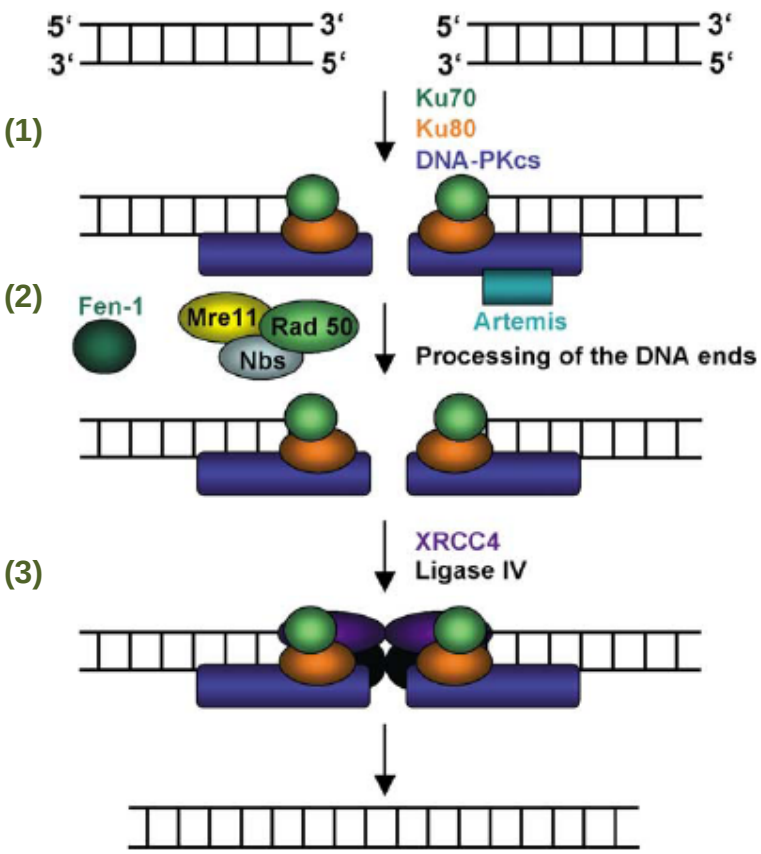
1) Resection of DSB in the 5'→3' direction revealing ssDNA ends to which RAD52 binds with RAD51 & RPA (via BRCA1 activation)
2) Formation of Holliday junctions allowing DNA resynthesis
3) Resolution of Holliday junctions and ligation of gaps in DNA

1.2.2.2.2 Non-homologous end-joining (NHEJ) repair

NHEJ repair begins with the heterodimeric Ku complex (Ku70-Ku80) (Jeggo *et al.*, 1992) binding to the damaged DNA (Reeves, Sthoeger and Lahita, 1989). This protects the DNA ends from exonuclease digestion. The Ku complex then associates with the catalytic subunit of the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) (Hartley *et al.*, 1995; Sipley *et al.*, 1995) and forms the active DNA-PK holoenzyme (Gottlieb and Jackson, 1993; Smith and Jackson, 1999). Interaction with single-stranded DNA at the site of DSBs triggers the DNA-PKcs' kinase activity (Hammarsten *et al.*, 2000; Martensson and Hammarsten, 2002). DNA-PKcs can also join together two broken ends without an active kinase domain (Bassing and Alt, 2004). DNA-PKcs targets XRCC4 which forms a stable complex with DNA ligase 4 (LIG4) (Leber *et al.*, 1998) and link together duplex DNA molecules with complementary but non-ligatable ends.

Processing of most DSBs is necessary prior to ligation by the XRCC4-LIG4 complex (Lee *et al.*, 2003). This is performed mainly by the MRN complex, which removes excess DNA at 3' flaps (Maser *et al.*, 1997; Nelms *et al.*, 1998), or flap endonuclease 1 (FEN1), which removes 5' flaps. FEN1 deficiency leads to a major reduction in the usage of the NHEJ pathway (Wu *et al.*, 1999). Another NHEJ factor involved in processing overhangs is the protein Artemis, which exhibits a single-strand-specific exonuclease activity and acts in a complex with DNA-PK (Moshous *et al.*, 2001). DNA-PKcs binds, phosphorylates and activates Artemis forming an endonuclease that is essential for processing 5' and 3' overhangs during NHEJ (Bassing and Alt, 2004; Christmann *et al.*, 2003). The mechanism of DSBR via NHEJ is summarised in Figure 1.5.

Figure 1.5: Diagram outlining the main steps involved in DSBR via non-homologous end-joining



Adapted from Toxicology, 193/1-2, Christmann et al., Mechanisms of human DNA repair, pages 3-32, Copyright (2003), with permission from Elsevier

1) Ku70–Ku80 complex recognises and binds damaged DNA then binds DNA–PKcs.
2) DNA ends are processed by MRN complex (via FEN1 and Artemis)
3) XRCC4–ligase IV (activated by DNA–PK) links broken DNA ends together.

1.2.2.3 Nucleotide excision repair (NER)

NER corrects the majority of DNA damage (photoproducts from UV radiation and other bulky lesions) (Friedberg *et al.*, 2000). NER is not essential for viability; many mouse strains with targeted NER mutations are viable and develop normally (Friedberg and Meira, 2004). However, mutations in NER are responsible for many DNA repair genetic disorders, such as Xeroderma pigmentosum (*XP* genes mutations), Cockayne syndrome (*CSA* and *CSB* mutations) and Trichothiodystrophy (*XPB* or *XPD* mutations) (Lehmann, 2003). Specific null mutations in murine NER homologues (like Xpd) can lead to early embryonic lethality or a reduced life span (Friedberg and Meira, 2000), showing that NER gene expression is important for normal embryo development as well as counteracting exposure to genotoxic stress. NER is divided into two subpathways: global genomic repair (GGR) and transcription-coupled repair (TCR), which repairs lesions in actively transcribed DNA (Leadon, 1999).

The main difference between GGR and TCR is the requirement for different factors during the initial recognition step; GGR initiation requires UV-DDB (DDB1 and DDB2) as well as XPC-HR23B and TCR is initiated by RNA polymerase II stalled at a lesion (Christmann *et al.*, 2003; Shuck *et al.*, 2008).

In NER, large multi-enzyme complexes scan the DNA molecule for a distortion in the double helix rather than for a specific base change. Once a bulky lesion is recognised, other proteins (XPA, RPA, RNA Polymerase II and XPG) bind to the site of the damaged base. This is followed by the binding of the ERCC1-XPF heterodimer, which produces the complete NER multi-protein complex. The phosphodiester backbone of the abnormal strand is then cleaved on both sides of the distortion. A portion of the strand containing the lesion, approximately 24-32 nucleotides (Barnes *et al.*, 1993), is removed from the double helix by the DNA helicase activities of XPB and XPD (Barnes *et al.*, 1993; van Brabant *et al.*, 2000), which unwind the DNA around the lesion (de Laat *et al.*, 1999) and allow dual incision at either side of lesion by XPG and XPF-ERCC1. The gap is then resynthesised with DNA polymerases δ or ε and other replication proteins, including PCNA, RPA and RFC. The repair process is completed by DNA ligase 1, which seals the nicks on the corrected strand by reforming the broken phosphodiester bonds.

1.2.2.4 Post-replication repair (PRR)

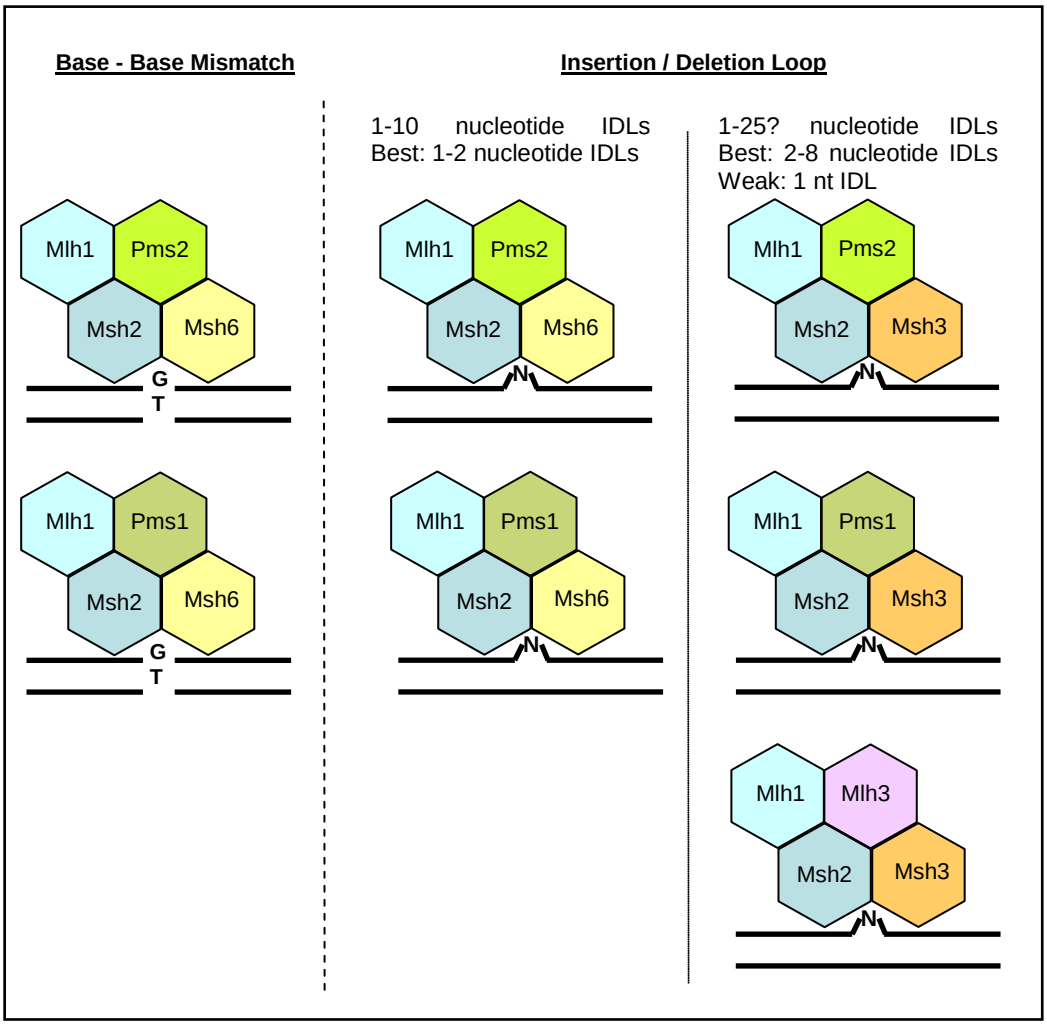
Post-replication repair, which occurs via the RAD6 pathway, does not exert any DNA repair activity and is better described as DNA damage bypass. The PRR pathway is activated when the DNA replication machinery is stalled at a lesion and allows mitotic somatic cells in S phase to proceed with DNA replication over a DNA template containing damaged bases, in an error-free (using the newly synthesized sister chromatid as a template to bypass DNA damage) or in an error-prone manner (using specialised translesion synthesis (TLS) polymerases). This mechanism is important as it avoids prolonged stalling at DNA replication forks, which could result in double strand breaks (DSBs), or premature termination of DNA replication in cells that would lead to unscheduled cell death (Ulrich, 2005).

1.2.2.5 Mismatch repair (MMR)

Mismatch repair corrects mismatched bases resulting from base deamination, oxidation or methylation or from DNA replication errors causing incorrect alignment of two individually normal strands (Christmann *et al.*, 2003; Karran and Bignami, 1999; Modrich and Lahue, 1996). The most common targets of MMR are G.T, G.G, A.C and C.C base mismatches which arise from deamination of 5-methylcytosine (Fang and Modrich, 1993). MMR increases the fidelity of DNA replication by correcting errors of the DNA polymerase that have escaped the 3'5' exonuclease proofreading activity (Stojic *et al.*, 2004). The importance of the mismatch repair pathway in maintaining genome stability is demonstrated by the association of inactive human MMR with a strong predisposition to tumour development (de Wind *et al.*, 1995; Hsieh and Yamane, 2008; Lipkin *et al.*, 2000; Lynch *et al.*, 2004; Wimmer and Etzler, 2008). Human germline mutations in MMR genes lead to Lynch syndrome (also known as hereditary nonpolyposis colorectal cancer or HNPCC). Approximately 10% of sporadic colorectal cancers are caused by MMR defects other than MMR gene mutations such as silencing of genes by promoter methylation (Kemp *et al.*, 2004). In mice that are *Mlh1*, *Mlh3*, *Pms2* or *Exo1* defective, increased tumour incidence can be accompanied by defects in meiosis which may cause sterility (Baker *et al.*, 1995; Baker *et al.*, 1996; Edelmann *et al.*, 1996; Li, 2008; Lipkin *et al.*, 2002) (see Table 1.2). Clearly MMR plays a critical role in meiosis and gamete formation in mice; however, the effects of MMR defects on meiosis in humans are still poorly characterised (Li, 2008).

In humans there are at least three MutS homologous proteins that participate in mismatch recognition and binding: MSH2, MSH3 and MSH6. MSH2 and MSH6 (also known as GT-binding protein, GTBP) (Palombo *et al.*, 1995) form the MutSα heterodimer and MSH2 and MSH3 form the MutSβ heterodimer (Palombo *et al.*, 1996). MSH2 plays a central role in MMR (Savouret *et al.*, 2003). The two complexes are ATPases that play a crucial role in the initiation of MMR and have different substrate specificities. MutSα (MSH2:MSH6) recognises base mismatches and 1-2 nucleotide insertion/deletion loops (IDLs). MutSβ (MSH2:MSH3) recognises 2-8 nucleotides IDLs most efficiently, is weakly active on single nucleotide IDLs and does not seem to be involved in base mismatches (Marti *et al.*, 2002) (Figure 1.6). The MSH2:MSH6 heterodimer is predominant in human cells (Jiricny, 1998; Stojic *et al.*, 2004).

Figure 1.6: Different DNA MMR complexes and their recognised substrates



Adapted by permission from Macmillan Publishers Ltd from Nature Genetics 24, Lipkin et al., MLH3: a DNA mismatch repair gene associated with mammalian microsatelliteinstability, pages 27-35, copyright (2000)

Unlike in BER and NER, in MMR neither nucleotide at the mismatch site is usually damaged. This makes it difficult to identify the strand that should be corrected. The MMR proteins need to distinguish the parental DNA strand from the newly synthesised 'incorrect' strand and this cannot be done at the site of the mispair (Stojic *et al.*, 2004). Although it is still unclear how MMR discriminates between the strands, it is presumed that the new strand is identified by the presence of single-strand breaks (SSB), i.e. the gaps between the Okazaki fragments on the lagging strand, or by the 3'-terminus on the leading strand. The SSB and the mismatch can be far apart. The MutSα complex triggers an ADP→ATP transition upon binding to a mismatch. This is followed by an ATP-dependent conformational change, forming a hydrolysis-independent sliding clamp capable of moving along the DNA backbone (Gradia *et al.*, 1999; Gradia *et al.*, 2000).

The hydrolysis of ATP allows the hMutSα.ATP.DNA complex to bind to a second heterodimer MutLα, composed of MLH1 and PMS2 (Alani *et al.*, 1997; Iaccarino *et al.*, 1998). MLH1 also forms a heterodimer with PMS1 or MLH3. The specific contributions of the MutL heterodimers to MMR function are still poorly understood, however, a functional redundancy between MLH3, PMS1 and PMS2 has been suspected (Lipkin *et al.*, 2000).

Another model suggests that the MutSα complex uses the energy gained by ATP hydrolysis to translocate actively along the DNA in either direction in search of a site that shows strand specificity (probably SSB). At this signalling site, the MutLα and MutSα come together (Blackwell *et al.*, 1998; Blackwell, Bjornson and Modrich, 1998). The excision of the strand containing the 'incorrect' base is carried out by exonuclease I (Genschel *et al.*, 2002) and the new synthesis to fill the formed gap is performed by DNA Polymerase δ (Longley *et al.*, 1997). The nick that is left after the polymerase has filled the gap is sealed by DNA ligase 1.

1.2.2.5.1 Microsatellite instability (MSI) and defects in MMR

Microsatellite sequences are tandem repeats of short DNA motifs ubiquitously present in the genome (Eckert and Hile, 2009). These sequences are highly unstable in both somatic and germ cell lineages and their expansions are known to be a molecular basis of genetic anticipation (Martorell *et al.*, 1997; White *et al.*, 1999). The human genome contains ~3% microsatellite sequences, which have an average density of 14bp/kbp for every chromosome (Subramanian *et al.*, 2003). The most abundant class of microsatellites in the human genome are mononucleotide repeats, mostly poly-A/T tracks, followed by dinucleotide and tetranucleotide repeats. Trinucleotide repeats, however, are less abundant (Eckert and Hile, 2009; Subramanian *et al.*, 2003). Many microsatellite loci are highly polymorphic, which explains their use as markers in genetic diagnosis and forensics. Mutation rates in non-tumourigenic human cells (estimated from direct observations of allele size changes between parents and offspring) range from 10^{-6} to 10^{-2} per locus per generation and depend primarily on repeat sequence composition, motif size and allele length; the location of the allele within the genome has a less significant effect (Ellegren, 2004).

1.2.2.5.1.1 Mechanisms of MSI

The mechanism by which somatic expansions occur is not well understood. A few hypotheses have been suggested; the main ones are listed below.

1. DNA polymerase slippage (during DNA replication or repair) can lead to additions or deletions within the microsatellite locus (Eckert and Hile, 2009; Li *et al.*, 2002).

Two different mechanisms can lead to slipped strand mispairing and microsatellite mutations: unequal crossing over between repetitive sequences on separate DNA molecules during recombination and transient DNA polymerase slippage during DNA synthesis (Subramanian *et al.*, 2003). Single strand breaks (SSBs) can increase the risk of microsatellite mutations.

2. Inappropriate DNA MMR leads to higher MSI. MMR is an important pathway that affects MSI as it processes replication errors with greater efficiency than the 3'→5' exonuclease proofreading of DNA polymerases (Vilkki *et al.*, 2001). MMR efficiency is dependent on motif size and sequence composition (Eckert and Hile, 2009).

3. Stabilisation of secondary structures in DNA by MutSα or β proteins mediation is another MMR factor that might affect MSI, however, the exact mechanism is still not well established. In a sequence and length dependent manner, common microsatellites can adopt non-B form DNA secondary structures, which can pause the replication by DNA polymerases, consequently increasing the production of errors during DNA synthesis (Eckert and Hile, 2009).

1.2.2.5.1.2 Defects in MMR

The primary function of MMR is the correction of persistent DNA replication errors in order to avoid the accumulation of deleterious mutations (Aquilina *et al.*, 1999). Defects in MMR lead to increased spontaneous mutation rates (Narayanan *et al.*, 1997). *Pms2*$^{-/-}$ mice develop spontaneous lymphomas and sarcomas, however, *Pms2*$^{+/-}$ develop normally without predisposition in tumour formation (Baker *et al.*, 1995). Studies on male transgenic mice crossed with *Pms2*$^{-/-}$ females suggested that the paternally derived *Pms2* gene provided normal levels of PMS2 protein to embryos by the 8-cell stage; however, smaller embryos derived from PMS2 deficient eggs lacked PMS2 function (Larson *et al.*, 2004).

This is in agreement with a similar study by Gurtu *et al.* (2002), which investigated the consequences of maternal MMR deficiency on genetic stability in the embryo. Mosaicism for the paternal alleles most probably resulted from PMS2 deficiency during the early cleavage divisions; additionally, the absence of MMR in one-cell embryos allowed the formation of unrepaired replication errors during the initial cell divisions in the zygote. Post-zygotic mutation in the early mouse embryo implied that PMS2 deficiency is due to lack of maternal transcripts (Gurtu *et al.*, 2002).

Human embryos that exhibit microsatellite instability, associated with spontaneous abortions, possibly harbour MMR mutations (Spandidos *et al.*, 1998). Specific mutations could be inherited from the parents and lead to microsatellite instability in normal cells as a result of MMR deficiency (Parsons *et al.*, 1995). In mice, *Msh2*, *Msh3*, *Mlh1* and *Mlh3* deficiencies have been associated with microsatellite instability (MSI) (Baker *et al.*, 1996; Lipkin *et al.*, 2000; Prolla *et al.*, 1998; Reitmair *et al.*, 1995). Absence of *Msh6* displays increased single base-pair mutation rates and lower penetrance of colon cancer susceptibility (Edelmann *et al.*, 1997). The effects of various MMR gene defects are summarised in Table 1.2.

Table 1.2: DNA mismatch repair functional deficiency associated with MSI and tumour phenotypes in the mouse and human (Li, 2008; Lipkin *et al.*, 2000; Wimmer and Etzler, 2008)

MMR gene	Phenotypes in knockout mice	Phenotypes resulting from homozygous or compound heterozygous mutations in human
MLH1	high MSI; lymphoma, GI + other tumors; infertile	high MSI; severe colon cancer susceptibility, GI + other tumors
MSH2	high MSI; lymphoma, GI + other tumors, fertile	high MSI; severe colon cancer susceptibility, lymphoma, GI + other tumors
MSH3	low MSI; tumour free or GI tumours at late age; fertile	unknown
MSH6	low MSI in dinucleotide repeats; lymphoma, GI + other tumors; fertile	low MSI, less intense colon cancer susceptibility, lymphoma, GI + other tumors
PMS1	MSI in mononucleotide repeats only; no tumours; fertile	low MSI, less intense colon cancer susceptibility
PMS2	MSI; lymphoma and sarcoma; male infertile	low MSI, less intense colon cancer susceptibility, lymphoma, GI + other tumors
MLH3	MSI, unknown tumour phenotype, Infertile	unknown
EXO1	MSI, lymphoma and sarcoma, infertile	unknown

GI: gastrointestinal

1.2.2.5.1.3 Expansion of trinucleotide repeats

Over 15 out of approximately 30 human hereditary disorders caused by DNA repeat dynamic mutations involve trinucleotide repeats (Mirkin, 2007; Pearson *et al.*, 2005). These consist mainly of neurodegenerative and neuromuscular disorders such as Myotonic dystrophy type 1 (DM1), Huntington's disease (HD), spinocerebellar ataxias and fragile X syndrome, which are caused by expansions in CNG trinucleotide repeats (Mirkin, 2006; Savouret *et al.*, 2004).

DM1, a neuromuscular condition involving myotonia and progressive muscle wasting, results from the expansion of an unstable CTG repeat in the 3' untranslated region of the dystrophia myotonica-protein kinase (*DMPK*) gene at 19q13.3 (Brook *et al.*, 1992). Somatic CTG repeat instability is widely observed in tissues of DM1 patients, with a strong bias towards expansions. This instability appears in early embryogenesis, increases after 16 weeks of gestation and continues into adulthood (Martorell *et al.*, 1997; Savouret *et al.*, 2003). Intergenerational contractions and expansions of a large CTG repeat were observed both in the germline and after fertilisation in mice (Savouret *et al.*, 2003). These dynamic mutations result from germinal mosaicism in the transmitting parent and an MSH2-dependent instability that takes place in the zygote just after fertilisation (Savouret *et al.*, 2004).

Lack of MSH2 shifts the instability towards contractions rather than expansions but a single *Msh2* allele is sufficient to drive instability towards expansions (Savouret *et al.*, 2003). Further studies showed that both MSH2 and MSH3 are necessary for the somatic expansions of CTG repeats in transgenic mouse models carrying expanded repeats (Manley *et al.*, 1999; Savouret *et al.*, 2003; van den Broek *et al.*, 2002). However, MSH6 activity suppresses the accumulation of repeat length variation, possibly because MSH6 competes more effectively than MSH3 over MSH2 and therefore limits the amount of MutSβ complex (van den Broek *et al.*, 2002). This suggests a key role for MutS homologues in triplet repeat expansions in mammalian cells (Gomes-Pereira *et al.*, 2004).

Therefore to sum up, triplet repeat expansions appear in cells prior to meiosis and the mechanism generating expansions is meiosis independent, occurs throughout life and involves MMR proteins. The differences in the dynamics and timing of trinucleotide repeat expansion could be due to the different genomic settings surrounding the repeats in these models (Libby *et al.*, 2003).

1.3 DNA repair in mammalian preimplantation development

The importance of DNA repair gene products during development is highlighted by the phenotypes observed for animals lacking key repair genes (see mouse database by Friedberg and Meira (2006)) and the human genetic disorders associated with DNA damage response defects (reviewed in Hales (2005)). Many animals lacking DNA repair enzymes are not viable, with preimplantation death being a common result; however, much of the embryonic loss takes place around the time of implantation, showing the necessity for DNA repair ability when embryonic cells proliferate rapidly and differentiate (Vinson and Hales, 2002).

The integrity of the genome is at greater risk during early embryonic development and the efficiency of DNA repair at those early stages is of great significance. Robust DNA damage response mechanisms are particularly important in undifferentiated embryonic cells to prevent any unnecessary mutations and avoid critical changes that can affect many different cell types in the organism and may be passed on to progeny. Indeed, enhanced repair of many types of DNA lesions resulting in distinct differences in mutation types and lower mutation frequencies have been reported in human and murine embryonic stem cells compared to somatic cells (Maynard *et al.*, 2008; Tichy and Stambrook, 2008).

Many recent studies have investigated DNA repair in mammalian preimplantation embryos; the main findings are described in a review by Jaroudi and SenGupta (2007). The next section summarises the key points and Table A.2 (Appendix) provides an updated list of all DNA repair genes found to be expressed in mammalian oocytes and preimplantation embryos.

1.3.1 Current understanding of DNA repair in oocytes and the preimplantation embryo

The integrity of the genome in germline cells ensures faithful transmission of genetic information from one generation to the next. However, some mutations are tolerated by gametogenesis; otherwise evolution would not be possible. DNA repair pathways ensure that the two opposite requirements, DNA sequence stability versus instability, are achieved and specific functions of those pathways may operate during gametogenesis (Baarends *et al.*, 2001).

Damage to the maternally inherited genome may arise due to the long delay between entry into meiosis and fertilisation of the oocytes (Drost and Lee, 1995). A large number of genes encoding proteins involved in different DNA repair pathways show enhanced or specialised expression during gametogenesis (Baarends *et al.*, 2001). DNA damage response and repair genes are overrepresented in the oocyte, compared to the preimplantation mouse embryo (1-cell to blastocyst stages). This may reflect the oocyte's response to pressures to ensure genomic integrity, especially after a long arrest in the first meiotic prophase (Zheng *et al.*, 2005). Using microarray analysis, Menezo *et al.* showed that human oocytes at the GV stage express DNA repair genes at high levels allowing low tolerance for DNA decays.

DNA repair genes have been shown to be expressed in the early stages of mammalian development. However, DNA repair in the newly fertilised preimplantation embryo is believed to rely entirely on the maternal mRNAs and proteins deposited and stored in the oocyte before ovulation (Vinson and Hales, 2002). DNA repair is probably essential during the formation of the embryonic genome. If the oocyte is not adequately equipped or if the embryonic gene expression does not start at the correct time, DNA damage will persist and the embryo's only recourse will be the activation of cell death machinery (Jurisicova *et al.*, 1998). However, during the first few cell cycles, cell death would diminish the survival chances of the embryo. Indeed, it seems that mammalian embryos may be inhibited from undergoing apoptosis before the late cleavage or blastocyst stages (Brison, 2000; Hardy, 1999).

The expression of DNA repair genes can determine whether or not the embryo survives following genotoxic stress during development. DNA repair genes of different repair pathways may be required by the embryo in different ways and at different stages of development (Jaroudi and SenGupta, 2007). mRNA expression data suggest that although many of the genes required for DNA repair are expressed, the embryo's ability to repair DNA may be highly limited. For instance, little or absence of *NBS1* and *OGG1* mRNA templates could limit BER in the rhesus monkey preimplantation embryo (Zheng *et al.*, 2005). The chromosomal abnormalities observed in cleavage stage human embryos may be due to deficiencies in recombination repair (Jackson, 2002; Pfeiffer *et al.*, 2000). However, as development continues after EGA, the embryo acquires greater ability to respond to DNA damage by regulating and activating DNA damage control genes (Zheng *et al.*, 2005).

The capacity of the mammalian embryo to respond to and repair damaged DNA and its selective sensitivity to specific lesions is still not well understood. Many gaps exist in our current knowledge concerning the precise roles and expression timings of several DNA repair genes in the early stages of embryonic development. The stage-specific variations in DNA repair gene expression transcripts and proteins point out the complexity of the regulation of these pathways during development. Many of the proteins involved in DNA damage control exist in large multimeric complexes that include proteins binding to damaged DNA, proteins participating in normal DNA replication, DNA repair proteins, proteins participating in meiotic recombination and proteins interacting with the cell cycle control machinery. The exact nature of interactions among the various proteins remains to be clarified. Additionally, there is some degree of redundancy in some of the DNA damage response pathways.

In vitro manipulations of human oocytes/embryos during *in vitro* fertilisation (IVF) using intracytoplasmic sperm injection (ICSI) are expected to increase the risk of DNA damage (by eliminating the selection against poor sperm DNA integrity) making the DNA repair ability of fertilised oocytes more crucial especially before embryonic genome activation (EGA) (Menezo, 2006). Whether or not embryo culture disconcerts the expression of genes involved in DNA damage control needs to be confirmed, as it raises questions about the utility and potential adverse effects of prolonged *in vitro* culture of human embryos (Zheng *et al.*, 2005).

It is important to understand precisely how and when preimplantation embryos in culture may experience changes in the ability to regulate DNA repair; as this will be extremely helpful for developing improved methods in embryo research, applied reproductive biology and assisted reproduction in clinical medicine.

1.4 Methods of investigating DNA repair in human oocytes and preimplantation embryos

1.4.1 Gene expression analysis

Gene expression analysis during embryonic development has been extensively studied in several species. However, little is known about the gene expression profiles in early human embryos, especially concerning DNA repair genes. The main limiting factor is the scarcity of material due to ethical issues concerning the use of human embryos for research. The previous section described many results obtained from the use of this technique for the investigation of DNA repair genes in mammalian oocytes and preimplantation embryos.

Most of the studies that analysed gene expression in early mammalian embryos employed the following procedure. First, embryos and/or oocytes were collected and tubed. The RNA was extracted from the samples, often pooled oocytes or embryos, however, individual oocytes, embryos or blastomeres can and have been used as well. The RNA was converted to cDNA by reverse transcription. The final step involved real-time quantitative PCR followed by automated quantification analysis (Mamo *et al.*, 2006b; Mamo *et al.*, 2007; May *et al.*, 2009; Wells *et al.*, 2005a; Wells *et al.*, 2005b; Zeng *et al.*, 2004; Zheng *et al.*, 2005).

The main limitation of real-time PCR for mRNA quantitation is that only a few genes can be investigated for a single sample (up to 12 genes from individual oocytes, embryos or blastomere) as at least three replicates need to be used per sample (Mamo *et al.*, 2006a; Mamo *et al.*, 2007). In order to overcome this problem, multiplex real-time PCR can be used (May *et al.*, 2009), which may allow simultaneous analysis of a slightly larger number of genes.

Quantitative analysis requires the selection of suitable housekeeping genes to be used as a reference for normalisation. This is not always straightforward, especially when analysing oocytes and embryos at different stages of development. Mamo *et al.* evaluated twelve commonly used housekeeping genes in mouse oocytes and embryos cultured *in vivo* and *in vitro* and recommend the use of the geometric average of the three most stable genes (Ppia, H2afz and Hprtl) for normalisation in preimplantation-stage expression studies (Mamo *et al.*, 2007).

Some studies employed microarray analysis to obtain mRNA transcript profiles. This requires pooling of oocytes or embryos prior to RNA extraction. The number of pooled samples in published studies ranged from as few as 1-5 oocytes or embryos (Wells and Patrizio, 2008; Zheng *et al.*, 2004) to as many as 400-500 mouse oocytes or embryos (Hamatani *et al.*, 2006; Mamo *et al.*, 2006b). This method also requires an amplification step to generate sufficient target amounts of RNA (in micrograms) for hybridisation onto arrays starting with picograms of extracted total RNA.

There are two main amplification techniques for expression studies: *in vitro* transcription (IVT) using RNA polymerases to amplify RNA (linear amplification) and polymerase chain reaction (PCR) using Taq polymerases to amplify cDNA (exponential amplification). Both seem to generate distortions but can show true differential expression between embryos at different stages of development (Degrelle, 2008). As Taq polymerases have higher error rates than RNA polymerases, IVT amplification may be favoured over PCR (Degrelle, 2008).

Amplification bias, which distorts the true proportional differences of candidate genes analyzed, can affect the validity of the experiment. Several recent studies have investigated RNA amplification efficiency for gene expression analysis using cDNA microarrays. Duftner *et al.* (2008) found minor amplification bias when using a T7 polymerase-based technique to amplify RNA. Some sequence-specific properties were found to significantly affect RNA amplification, these include: GC content, folding energy, hairpin length and number and lengths of poly(A) or poly(T) stretches. Patel *et al.* (2005) demonstrated that the use of amplified RNA using a T7 polymerase-based technique in microarray experiments retains the fidelity of detection of differential gene expression that is relatively comparable to experiments carried out on unamplified RNA.

Rudnicki *et al.* (2004) showed that performing two rounds of IVT amplification using T7 oligo primers resulted in amplified RNA with a high degree of linearity and reproducibility. Comparison of amplified versus unamplified mRNA showed a correlation of 0.868 and intra-amplification consistencies showed correlations of 0.968, 0.907 and 0.912 for the first round, the second round and two successive rounds of amplification (Rudnicki *et al.*, 2004). Another study showed loss of 30% of differentially expressed genes when two rounds of amplification are performed using T7 oligo primers in the first round and random hexamer primers in the second round (Boelens *et al.*, 2007). Therefore, although there are some discrepancies in the reported findings, it seems that the main limitation of microarray analysis following RNA amplification is the decreased capacity of detection of differentially expressed genes.

Pooling samples together overcomes sample-to-sample variation providing satisfactory statistical power and can improve the efficiency and cost-effectiveness of microarray experiments (Peng *et al.*, 2003) as it increases the number of samples investigated without increasing the number of arrays. Having biological and technical replicates, with a minimum requirement of three (Lee *et al.*, 2000), is also necessary to be able to validate the results and carry out statistical tests.

The main advantage of microarray analysis is that it provides information on the expression levels of potentially every gene in the genome from a single experiment. However, it is still not sensitive enough to allow quantitative analysis that can be obtained using real-time PCR (Wong and Medrano, 2005) from oocytes, preimplantation embryos or single blastomeres. For these reasons, the best and generally adopted strategy for gene expression analysis is the use of microarrays to generate a global expression profile from the samples of interest and subsequently, based on the obtained results, investigate a number of target genes in greater detail using real-time PCR. As real-time PCR does not necessitate an RNA amplification step, any amplification bias that may have affected microarray results should be picked up by real-time PCR. This formed the basis of the first aim of this project: ‘DNA repair gene expression profiling in human oocytes and blastocysts’ using microarrays.

1.4.2 Proteomic and metabolomic analysis

Expression of DNA repair mRNA transcripts alone does not necessarily indicate translation of all templates (if any) into potentially functional proteins. Several factors, including post-transcriptional processing by microRNAs and targeted mRNA or protein degradation, can result in discrepancies between gene expression levels and the protein content in the cell.

Direct detection of proteins involved in DNA repair can be conducted on single oocytes, embryos or blastomeres using immunofluorescence analysis, which allows visualisation and localisation of the targeted protein within the cell under a fluorescent microscope. Several studies have used this technique for the investigation of specific DNA repair proteins in mammalian oocytes, embryos or single mouse embryonic fibroblasts (Adiga *et al.*, 2007; Barton *et al.*, 2007; Fernandez-Gonzalez *et al.*, 2008; Roig *et al.*, 2004; Wirthner *et al.*, 2008). The main advantage of this technique is that it allows the localisation of proteins within the cell, detecting protein-protein or protein-DNA interactions; however, its main limitation is that it allows the analysis of a limited number of proteins in a given sample.

A larger number of proteins can be analysed from oocyte or embryo lysates; however, despite the recent advancements in proteomic technologies, the available platforms are still not sensitive enough to fully investigate the limited amounts of protein templates available in individual mammalian preimplantation embryos. Two-dimensional (2-D) gel electrophoresis accompanied by computerised analysis was initially used to generate protein databases for preimplantation mouse embryos (using lysates from ~30 pooled embryos) at different stages of development (Latham *et al.*, 1992; Shi *et al.*, 1994). Western blotting can be used to detect the expression of known proteins or specific post-translational modifications in preimplantation embryos; this technique also requires pooling of large numbers of embryos in order to reach the necessary protein amounts (Cho *et al.*, 2003; Wang *et al.*, 2005).

Embryo lysates (from single or pooled oocytes and preimplantation embryos) can be used to analyse a large number of proteins at once with higher sensitivity using mass spectrometry (MS). MS using surface-enhanced lazer desorption/ionisation coupled to time of flight (SELDI-TOF MS) allows the analysis of small samples in the picomole to femtomle range (Katz-Jaffe *et al.*, 2009).

SELDI-TOF MS was used to generate protein profiles of individual human embryos at the early and expanded blastocyst stages (Katz-Jaffe *et al.*, 2005). The main disadvantage of MS is the need to carry out additional experiments like Western blotting and bioinformatics techniques for protein identification.

Currently, different methods for proteomic and metabolomic analysis of the preimplantation embryo secretome, consisting of proteins and metabolites secreted by the embryo into the culture media, are under investigation as tools for non-invasive assessment of the embryo's viability. Protein profiling of the human embryonic secretome revealed distinct signatures for the different stages of development, when SELDI-TOF MS analysis was used (Katz-Jaffe, Schoolcraft and Gardner, 2006), and identified proteins that may be associated with embryo development and implantation potential, when protein microarray analysis targeting 120 proteins was conducted (Dominguez *et al.*, 2008). Protein microarray analysis of the embryo secretome can complement gene expression microarray results, allowing more accurate biological interpretations. Recent reviews on proteomic analysis and profiling of metabolites in the embryonic secretome were presented by Brison *et al.* (2007) and Katz-Jaffe *et al.* (2009).

Another non-invasive method for embryo selection is measuring amino acid turnover in the culture media using reverse-phase high performance liquid chromatography, reviewed by Sturmey *et al.* (2008). This method seems to be simpler and more cost-effective than metabolic profiling, which looks at all metabolites, making it better suited for embryo selection in clinical IVF (Sturmey *et al.*, 2008). A recent study by Sturmey *et al.* (2008) found a correlation between the amount of DNA damage and the metabolic activity, measured as amino acid turnover, in human, bovine and porcine preimplantation embryos. However, extent of DNA damage did not correlate with embryo grade. Their finding was consistent with the "quiet embryo hypothesis", which states that embryos with less active metabolisms are more viable (Leese, 2002). Baumann *et al.* (2007) speculated that this is because the extent of DNA damage and the RNA and protein content of the immature oocyte can determine the embryo's potential to reach the blastocysts stage.

Therefore, an embryo with low metabolic activity is either exposed to fewer insults or is better equipped to efficiently deal with those insults; however, an embryo with higher levels of damage to the genome, proteome or transcriptome exerts higher metabolic activity as it is less efficient at correcting damage and maintaining development (Baumann *et al.*, 2007). Amino acid turnover can act as a non-invasive marker of DNA damage in the human blastocyst (Sturmey *et al.*, 2009).

Detection of a large number of DNA repair proteins in a single or a few human oocytes or preimplantation embryos is still difficult; however, with the advancement of protein microarrays it may be possible to obtain a DNA repair protein profile from human oocytes or preimplantation embryos in the near future. Detecting the presence of a particular DNA repair protein does not guarantee functional activity of that protein as the activation of the repair pathway may be dependent on other proteins involved in DNA repair and cell cycle control. The completion of repair can involve over 30 different proteins. Measuring DNA damage can be indicative of exogenous stress rather than just DNA repair ability. For all the above reasons, it seems logical that the best way to assess the embryo's DNA repair ability is by using functional assays.

1.4.3 Functional analysis

The functions of many DNA repair genes were deduced from phenotypes of animals with germline mutations in a particular repair gene. A database of mouse mutant strains for genes affecting the cellular response to DNA damage generated by targeted gene replacement or transgenic technologies was initiated in 1997. The latest update of the database is Version 7 (Friedberg and Meira, 2006). Oocytes in mouse *Spo11*$^{-/-}$ mutants that fail to form DSBs for the initiation of recombination and *Msh5*$^{-/-}$ and *Dmc1*$^{-/-}$ mutants that fail to process DSBs showed that responses to recombination errors in mammalian oocytes can be DNA damage dependent or independent (Di Giacomo *et al.*, 2005). More recently, mice harbouring deletions in MMR genes (Mlh1, Mlh3, Exo1 and ATPase deficient variant of Mlh1) were used to show that MMR function is necessary for normal formation and stabilisation of crossovers in mammalian oocytes (Kan *et al.*, 2008). *Parp1* and *Parg* deficient mice, which exhibit morphological and functional sperm abnormalities, showed the importance of the poly(ADP-ribose) (PAR) metabolism for normal sperm chromatin quality (Meyer-Ficca *et al.*, 2009).

Using cell sorting to separate cells at specific times after irradiation, Wu *et al.* (2008) measured DSBR with pulse-field gel electrophoresis in G1 and G2 phase Chinese hamster cells. While wild type and mutant cells for the HR repair genes *Xrcc2* and *Xrcc3* had similar repair efficiencies in G1 and G2, mutants of the DNA-PKcs, *Ku80* and *Xrcc4* showed greater repair in G2 than in G1, which indicated enhanced function of the backup NHEJ pathway involving DNA ligase III and PARP1 (Wu *et al.*, 2008). These are a few examples among numerous studies reported in the literature, many of which investigated DNA repair in cell lines and non-mammalian mutants (e.g. Drosophila melanogaster).

Small interfering RNA (siRNA) can be used to knock down the expression of specific DNA repair genes. For example, stable siRNA mediated silencing of Smug1 in mouse embryo fibroblasts (MEFs) resulted in increased mutation rates (An *et al.*, 2005). The use of specific siRNA in mouse embryonic stem cells (mESCs) showed that APE1 is necessary for normal embryonic hematopoiesis due to the critical redox function but not the repair endonuclease function of APE1 (Zou *et al.*, 2007) and PRMT1 is necessary for the recruitment of the HR RAD51 recombinase to DNA damage foci (Yu *et al.*, 2009). Combined with expression analysis and immunocytology, RNA interference mediated depletion of BRCA1 in young mouse oocytes was carried out to confirm the role of BRCA1 in chromosome segregation (Pan *et al.*, 2008).

Immunofluorescence (which can be combined with time lapse imagery) is used to detect specific DNA repair proteins in individual cells and oocytes. Derijck *et al.* (2008) demonstrated that HR repair is predominantly active in the male pronuclei during the first S phase of mouse zygotes by tracking γH2AX foci, which are indicative of DSBs, and RAD51 protein at different stages of the cell cycle. Another group used immunofluorescent staining to investigate DSBR by HR in relation to chromosome dynamics in human foetal oocytes by detecting γH2AX, RPA and MLH1 foci, which are markers for the occurrence of DSBs, progress and completion of HR repair, respectively (Roig *et al.*, 2004).

A commonly used system for DNA repair assessment employs plasmid/bacteriophage circular DNA substrates (which may be radiolabelled) that are transfected into cells for *in vivo* repair or exposed to nuclear or whole cell extracts for *in vitro* repair.

In order to investigate the role of PARP-1 and p53 in DSBR, Susse *et al.* (2004) transfected DNA plasmids into primate cells and assessed repair by immunoblot analysis and flow cytometric quantification of recombination frequencies. Another study used an *in vivo* DSBR assay to investigate the role of RAD54 in mESCs (Dronkert *et al.*, 2000).

Post exposure to ~50-1500μg of protein extracts, substrates (carrying the *Escherichia coli lacZ* gene) can be electroporated into MMR deficient *E. coli* in order to assess repair. The transformed bacteria are plated onto minimal agar supplemented with X-gal and repair efficiencies are assessed by counting blue/white plaques. This technique was mostly used for BER (Waters and Akman, 2001; Zhang and Dianov, 2005) and MMR (Fang and Modrich, 1993; Matheson and Hall, 2003).

Alternatively, repair is assessed with immunoblot analysis and gel shift assays. This approach has been used to investigate the role of Ku80 in NHEJ (Feldmann *et al.*, 2000), OGG1 incision activity (Riis *et al.*, 2002) and interaction of APE1 and other DNA repair proteins with specific DNA substrates that match BER intermediates (Dyrkheeva *et al.*, 2008) or with restriction endonuclease digestion followed by electrophoresis for mismatch and insertion/deletion loop repair (Kadyrov *et al.*, 2009; Littman *et al.*, 1999; McCulloch *et al.*, 2003a; Wang and Hays, 2002a). Many studies combine different techniques especially when investigating a pathway that is not fully understood. A recent study by Wu *et al.* (2008b) investigated DSBR in G1 and G2 MEF mutants using both *in vivo* and *in vitro* plasmid end-joining assays and showed a new cell cycle regulation of the backup NHEJ pathway in G2. Clearly, the *in vivo* repair system cannot be used for DNA repair assessment in human oocyte or preimplantation embryos. Repair carried out *in vitro* using protein extracts may be used, however, if the assays had suitable sensitivities.

Several studies used functional assays with cell-free extracts from Xenopus and Drosophila eggs or embryos to assess NER (Oda *et al.*, 1996), MMR (Petranovic *et al.*, 2000; Varlet *et al.*, 1996; Varlet *et al.*, 1990) and NHEJ (Labhart, 1999); however, there are no studies showing assessment of DNA repair in cell-free extracts of mammalian oocytes or preimplantation embryos. A recent study measured OGG1 activity in protein extracts from human embryonic stem cells (hESCs) (Maynard *et al.*, 2008).

Mismatch repair is probably the most investigated pathway because it is important in rapidly dividing cells and its substrates can be easily synthesised. Many studies have targeted the functional assessment of mismatch repair in human cell/nuclear extracts to investigate the repair abilities of various cancer cell lines that are deficient in different MMR proteins (Bennett *et al.*, 1997; Matheson and Hall, 2003; Thomas *et al.*, 1996).

Currently, there are no available DNA repair assays suitable for direct application on cell-free extracts from human oocytes or preimplantation embryos. Assessing the capacity of the human preimplantation embryo to detect and repair different types of DNA damage would be extremely useful. The second aim of this project was thus to develop an *in vitro* functional assay for mismatch repair and insertion/deletion loop repair that could be applied to nuclear extracts from human oocytes and embryos.

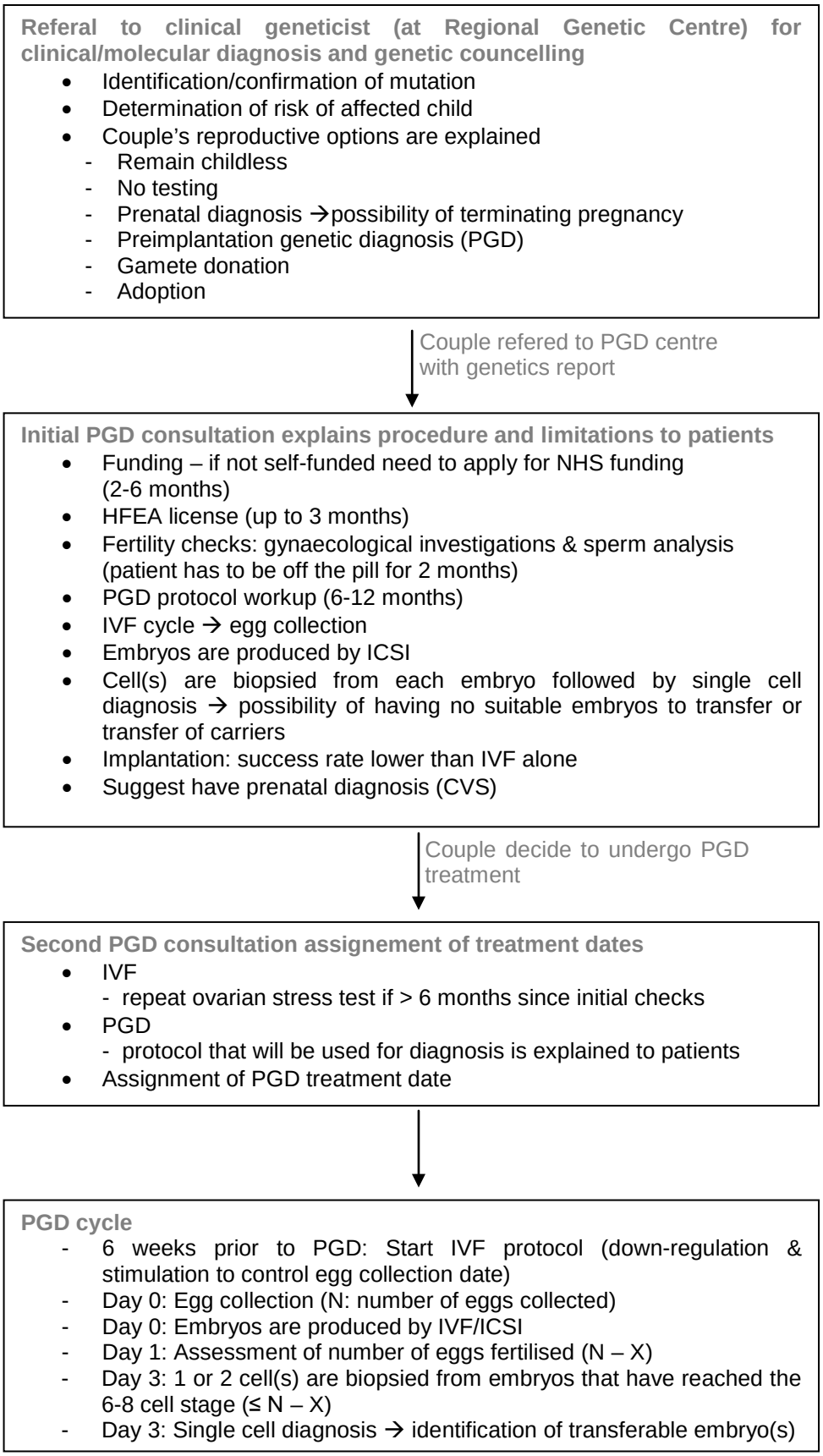
1.5 Preimplantation genetic diagnosis (PGD) for cancer predispositions caused by DNA repair deficiencies

1.5.1 Preimplantation genetic diagnosis (PGD)

Preimplantation genetic diagnosis (PGD) was first developed in the UK around 20 years ago (Delhanty and Harper, 2000; Handyside *et al.*, 1990). It is an established method for the genetic analysis of embryos prior to implantation and pregnancy. Individual blastomeres are genetically tested to select the unaffected embryos for transfer into the uterus. PGD can be considered an alternative to prenatal genetic diagnosis for couples at high risk of transmitting a genetic disorder to their offspring. Its main advantage is that it eliminates the possibility of having to choose whether or not to terminate an affected pregnancy.

Patients requesting PGD undergo *in vitro* fertilisation (IVF) treatment so that multiple embryos can be generated, giving an increased probability that a disease free embryo will be identified (Wells and Delhanty, 2001). Intracytoplasmic sperm injection (ICSI) is used to reduce the risk of paternal contamination from sperm when blastomeres are biopsied. The main steps involved in patient councelling and treatment are shown in Figure 1.7. PGD centres in the UK are controlled by the Human Fertilisation and Embryology Authority (HFEA) (Delhanty and Harper, 2000).

Figure 1.7: Flowchart outlining the main steps preceding patients' PGD treatment cycle



Genetic analysis is performed using fluorescent in situ hybridisation (FISH) for the detection of chromosomal abnormalities and sex selection (for X-linked disorders with no specific molecular diagnosis) or using polymerase chain reaction (PCR) for the detection of mutations responsible for monogenic disorders. FISH is also used for preimplantation genetic screening (PGS). The indications for PGS are advanced maternal age, repeated IVF failure and recurrent miscarriage when both parents have normal karyotypes. PGS is also offered to couples undergoing ICSI because the male partner has a severe oligospermia and/or a meiotic anomaly (Egozcue *et al.*, 2000; Wilton, 2002). At present, there is no evidence that shows that PGS improves pregnancy rates (Anderson and Pickering, 2008; Harper *et al.*, 2008). To date, nine randomised control studies have been conducted on 'good prognosis' patients (Jansen *et al.*, 2008; Mersereau *et al.*, 2008; Meyer *et al.*, 2009; Staessen *et al.*, 2008) and 'poor prognosis' patients (maternal age≥38 years) (Debrock *et al.*, 2009; Hardarson *et al.*, 2008; Mastenbroek *et al.*, 2007; Staessen *et al.*, 2004; Stevens *et al.*, 2004); all have shown that PGS does not improve live birth rates compared to control groups.

Currently, most PGD centres employ cleavage stage biopsy on day 3 post fertilisation (at the 6-8 cell stage), where 1 or 2 cells are extracted from each embryo (Goossens *et al.*, 2009). Blastocysts stage biopsy on day 5 post fertilisation is a possible alternative that overcomes the problem of small cell number (Basille *et al.*, 2009), however, it allows less time for the diagnosis (as the embryo needs to be transferred by day 6) or would require embryo freezing prior to transfer into the uterus in a future cycle. Alternatively, biopsy of both the first and second polar bodies can be performed when the woman is the carrier of a mutation or a chromosomal aberration (Munne *et al.*, 2000; Strom *et al.*, 2000; Verlinsky *et al.*, 1998). This is necessary in countries where cleavage stage biopsy is not allowed (e.g. Germany). The main advantage of blastomere analysis is that it allows testing of both parental genomes. However, the main disadvantage is mosaicism, where different blastomeres have a different chromosomal complement resulting from post fertilisation mitotic errors. Blastocyst stage biopsy overcomes this problem as the embryo is more stable at that stage. Additionally, since the cells are taken from the trophectoderm (TE), the inner cell mass is left intact avoiding any harmful stress to the embryo (Basille *et al.*, 2009).

One of the problems of PGD is the possibility of misdiagnosis as the genetic analysis from a single cell is technically challenging (Delhanty and Harper, 2000). The causes of misdiagnosis include failure of the genetic test (e.g. due to unsuitable probes), chromosomal mosaicism, parental contamination, confusion of embryo/cell number and transfer of the incorrect embryo (Wilton *et al.*, 2009). As the embryos subjected to PGD must be generated by IVF, certain difficulties are encountered and the success of PGD will always be dependent on the success of IVF/ICSI. The latest European Society of Human Reproduction and Embryology (ESHRE) PGD consortium data show that clinical pregnancy rate is 17-26% per oocyte retrieval and 29-33% per embryo transfer (22-23% implantation rate) and the delivery rate is 15-23% per oocyte retrieval and 26-29% per embryo transfer (Goossens *et al.*, 2009).

1.5.2 PGD for monogenic disorders

The most common method of genetic analysis for monogenic disorders is PCR. Mutation detection can be direct (when the amplicon sequence includes the mutation site) or indirect via linkage analysis, which involves the amplification of informative highly polymorphic microsatellite markers or single nucleotide polymorphisms (SNPs) that are very close/linked to or within the gene of interest. The simultaneous amplification of different loci using multiplex PCR has become the gold standard in PGD for monogenic disorders (Spits and Sermon, 2009). The minute amounts of starting genomic DNA, obtained from a single blastomere, make the problems that can be encountered in conventional PCR additionally challenging in PGD. PCR protocols used in PGD cases have to be optimised with great care to minimise allele dropout (ADO), preferential amplification and allow the detection of contamination.

ADO is the random failure of amplification of one allele in a heterozygous sample due to the primers annealing to one allele with lower efficiency. ADO can lead to a misdiagnosis as a heterozygous embryo can be diagnosed as homozygous. While optimisation of PCR conditions can minimise its occurrence, ADO is a random phenomenon that occurs even in optimally standardised PCR protocols because of degradation of the template DNA at the target site. For this reason, PCR-based genotyping protocols should include internal monitoring in the form of informative linked markers.

Ideally, at least two linked markers (located as close as possible to the gene of interest or mutation site) should be used with a marker on either side of the mutation/gene to detect recombination (Wilton *et al.*, 2009). A distance of 1Mb limits the chances of recombination to 1% (Collins, 2009). Initial analysis of parental DNAs determines the alleles to be expected in the embryos as well as which marker alleles segregate together with the mutation (the phase alleles).

Another potential source of misdiagnosis, which can also be detected when using a suitable multiplex PCR protocol, is contamination. Contamination from cumulus cells or sperm can be avoided by stripping the oocytes of cumulus cells and using ICSI for fertilisation (Thornhill *et al.*, 2005). Contamination at the PCR stage can involve the operator's DNA or amplicon carry over from previous PCR experiments. In order to minimise the risk of contamination, laboratories performing PGD take special measures such as isolating the working areas where embryo biopsy and tubing of blastomeres take place, where the PCR is setup and where post PCR analysis is performed, working in laminar flow hoods and areas that are cleaned and UV treated, aliquoting PCR reagents and using dedicated laboratory equipment for single cell PCR.

1.5.2.1 Methods of molecular analysis

Fluorescent PCR (F-PCR) uses primers that are labelled at the 5'-end with a fluorochrome in order to produce labelled PCR products that can be analysed on an automated genetic analyzer, which allows very accurate sizing of DNA fragments and simultaneous detection of several amplicons. Using this method, the alleles are determined based on fragment size. Fragment size analysis is used for the detection of small insertions or deletions or linkage analysis using STR markers. Indirect mutation detection via linkage analysis may involve the use of SNP sites instead of microsatellite markers. Mini-sequencing uses primers that anneal one base before the target site and are elongated by only one dideoxy nucleotide (which is labelled with one of four fluorochromes depending on the base). Automated sequencing analysis allows the determination of the base at the SNP site of interest. Single-strand conformation polymorphism (SSCP) is a simple and rapid mutation detection method that is capable of detecting single base substitutions, small insertions/deletions and rearrangements in DNA fragments ranging in size from 100 to 500 base pairs (bp) (Orita *et al.*, 1989). Single stranded DNA adopts specific conformations which are stabilised by intra-strand interactions and are uniquely dependent on sequence composition (Orita *et al.*, 1989).

1.5.3 PGD for *MSH2* and *BRCA1*

PGD for *MSH2* or *BRCA1* was not available in the UK when this project was started. In fact, PGD for cancer predispositions with a late age of onset and incomplete penetrance (like breast and ovarian cancer) was not allowed in the UK until recently. The HFEA carried out a public consultation between December 2005 and March 2006 and decided it was appropriate for PGD to be available for hereditary breast and ovarian cancer (Menon *et al.*, 2007). No reports on PGD for *BRCA1* had been published prior to 2007 (Spits *et al.*, 2007).

The basis of the third aim of this project was to investigate embryos donated for research following PGD cycles for germline mutations in DNA repair genes, particularly *MSH2* and *BRCA1*.

1.5.3.1 MSH2 mutations and Lynch syndrome

The *MSH2* gene (OMIM *609309) maps to human chromosome 2p22-21 and consists of 16 exons spanning 80.10 kilobases (Kb) of DNA (Fishel *et al.*, 1993; Leach *et al.*, 1993). MSH2 plays an important role in DNA MMR and stability of microsatellite repeats. Heterozygous mutations in *MSH2* can lead to Lynch syndrome (LS) (OMIM #120435), also known as hereditary nonpolyposis colorectal cancer (HNPCC) (Liu *et al.*, 1996). This autosomal dominant heritable cancer syndrome is characterised by early age (~45 years) at onset of colorectal and other gastrointestinal cancers as well as endometrial and other cancers.

Approximately 500 different MMR mutations lead to LS; 40% involve *MSH2*, 50% involve *MLH1* and 10% involve other MMR genes (*MSH6* and *PMS2*) (Papp *et al.*, 2007). The lifetime risk of developing cancer is 80% for individuals carrying a MMR mutation (Peltomaki and Vasen, 1997; Watson and Lynch, 1993). Carriers of different mutations in various MMR genes may have different cancer risks (Peltomaki *et al.*, 2001). For example, carriers of an intron 5 splice site mutation in *MSH2* have an 89% risk of developing cancer by the age of 60, whereas carriers of the exon 8 and exon 4-16 deletions have an 81% and 85% risk, respectively (Stuckless *et al.*, 2007). Many of the *MSH2* mutations causing LS involve large deletions and genomic rearrangements (Charbonnier *et al.*, 2000; Charbonnier *et al.*, 2002; Papp *et al.*, 2007; Wang *et al.*, 2003; Wijnen *et al.*, 1998).

1.5.3.2 BRCA1 mutations and breast cancer

BRCA1 (OMIM *113705) is a tumour-suppressor gene that maps to human chromosome 17q21 and confers a high risk of breast and ovarian cancer (OMIM #114480) (Miki *et al.*, 1994). The gene consists of 22 exons spanning ~110 kilobases (Kb) of DNA (Miki *et al.*, 1994) and encodes a 1863 amino acid nuclear protein (220KDa) (Koonin *et al.*, 1996). BRCA1 interacts with many proteins involved in transcriptional regulation, cell cycle checkpoint control, chromatin remodelling and DNA repair (Coupier *et al.*, 2004; Venkitaraman, 2002).

Up to 10% of breast cancers are hereditary and most are caused by mutations in *BRCA1* or *BRCA2* (Liebens *et al.*, 2007). Hereditary breast or ovarian cancers are characterised by an early age at onset (<45 years). More than 660 mutations have been identified in *BRCA1*, many involving small deletions (Dumitrescu and Cotarla, 2005) and more than 75% of *BRCA1* mutations result in truncated protein (Hogervorst *et al.*, 1995). Some *BRCA1* founder mutations have high prevalence rates in specific ethnic groups and populations (e.g. Ashkenazi Jews and Netherlands).

Women carrying *BRCA1* mutations have an estimated 60-85% lifetime risk of developing breast cancer and 26-54% lifetime risk of developing ovarian cancer (Antoniou *et al.*, 2003; Brose *et al.*, 2002; King *et al.*, 2003; Liebens *et al.*, 2007). *BRCA1* deficient or heterozygous mutant cells show increased radiosensitivity (Buchholz *et al.*, 2002; Foray *et al.*, 1999; Rothfuss *et al.*, 2000) and impaired DSBR via HR (Cousineau and Belmaaza, 2007) or NHEJ (Baldeyron *et al.*, 2002; Coupier *et al.*, 2004).

1.6 Aims

1) DNA repair gene expression profiling in human oocytes and blastocysts

The first aim of this project was to obtain a global expression profile of DNA repair genes in human *in vitro* derived preimplantation embryos before and after embryonic genome activation (EGA) using microarrays. Prior to EGA, the embryo relies on maternal transcripts. Mature (MII) oocytes were selected to represent the mRNA pool of pre EGA embryos. The blastocyst stage, which is a key developmental stage, was used for the analysis of post EGA expression.

A previous expression study showed high expression levels of most DNA repair genes in human GV oocytes (Menezo *et al.*, 2007). However, there is no evidence indicating whether the blastocyst has an intrinsic supply of DNA repair mRNA under routine IVF conditions.

Hypothesis 1: The human oocyte expresses most DNA repair genes to support the early preimplantation embryo and limit DNA damage.
Hypothesis 2: Due to the high rate of replication and the onset of differentiation in the blastocyst, the expression profile of DNA repair genes may be different to oocytes.

In order to identify the DNA repair pathways that are active pre and post EGA, mRNA from triplicate sets of pooled human *in vitro* matured oocytes (MII) and blastocysts was analysed using the Human Genome Survey Microarrays V2.0 (Applied Biosystems^{TM}).

2) Development of a functional assay for MMR and IDL repair in human nuclear extracts

The gene expression profile alone cannot determine activity or functionality of any given DNA repair pathway. The second aim of this project was thus to develop an *in vitro* DNA repair functional assay that was more sensitive than currently available assays and could be applied to cell-free extracts from a limited number of human oocytes or embryos. Mismatch repair was the targeted pathway because the rates of cell proliferation and DNA replication are high in the early embryo.

Mismatched heteroduplex DNA constructs containing single base mismatches or insertion deletion loops (IDLs) were created from PCR products or synthetic

oligonucleotides to act as substrates for repair. The constructs were exposed to human nuclear extracts and repair was assessed by detection of newly formed homoduplexes. The flexible strategy developed for the formation of heteroduplexes could be used to generate constructs of different sizes with small or large IDLs and various single base mismatches by targeting different STR or SNP loci.

The DM1 $(CTG)_n$ sequence was targeted in the creation of heteroduplexes with IDLs as this triplet repeat is known to show anticipation. DM1 represents the largest group of PGD cases carried out at the UCL Centre for PGD; thus, the application of the designed assay on protein extracts from DM1 affected embryos may help elucidate the mechanism of expansion of this CTG repeat during preimplantation development. An SNP located in the *MSH2* gene (rs1981929) was used to form heteroduplex constructs with an A.C or G.T mismatch.

3) Development of PGD protocols for *BRCA1* and *MSH2* mutations

The third aim of this project was to investigate the effect of germline mutations in DNA repair genes on early embryonic development. In order to achieve this aim, clinical PGD for *MSH2* and *BRCA1* was initiated and spare embryos donated for research were collected.

Comparison of the expression profiles of embryos with mutations in DNA repair and cell cycle checkpoint genes with the profiles obtained from routine IVF embryos may help identify factors that are important for normal development and implantation. These embryos could also be investigated using functional assays, specifically assessing mismatch repair in embryos with *MSH2* mutations.

2 MATERIALS AND METHODS

The methods have been divided into five themes: sample collection and processing, preparation of nucleic acids, amplification methods, assessment of amplified products, and processing and analysis of amplified products. The methodology required to achieve each of the three aims of the project involved a variety of techniques spanning several of the themes. The workflow for each of the aims is presented in Table 2.1.

Table 2.1: Outline of methodology used to achieve each of the three aims of the project

Method theme	**Aim**		
	Gene expression profiling of DNA repair genes using microarrays	Development of a functional assay for MMR and IDL repair	Development of PGD protocols for *MSH2* and *BRCA1*
Sample collection & processing	IVF protocol, collection of immature oocytes & grading of embryos (2.2.1)		IVF protocol collection of immature oocytes & grading of embryos (2.2.1)
	Tubing of oocytes & blastocysts (2.2.2)	Collection & processing of blood samples (2.2.3)	Collection & processing of blood samples (2.2.3)
			Separation of lymphocytes (2.2.3.1)
			Single cell isolation & tubing of lymphocytes (2.2.3.2)
			Embryo biopsy & blastomere tubing (2.2.4)
Preparation of nucleic acids	RNA extraction (2.3.1)	DNA extraction from whole blood (2.3.2)	DNA extraction from whole blood (2.3.2)
		Measurement of DNA concentration (2.3.3)	Measurement of DNA concentration (2.3.3)
Amplification methods	RNA amplification & labeling (2.4.1)		
		DNA amplification using PCR (2.4.2.1 & 2.4.2.2)	DNA amplification using PCR (2.4.2.3)
Assessment of amplified products	Agarose gel electrophoresis (2.5.1)	Agarose gel electrophoresis (2.5.1)	Agarose gel electrophoresis (2.5.1)
	Assessment of RNA (2.5.2)		
Processing & analysis of amplified products	Microarray analysis (2.6.1)		
		Formation of heteroduplex DNA constructs (2.6.2)	
		Exposure of heteroduplex constructs to nuclear extracts (2.6.3)	
		ABI Prism genetic analysis (2.6.4/2.6.4.1)	ABI Prism genetic analysis (2.6.4/2.6.4.2)
		SSCP/heteroduplex analysis (2.6.5)	SSCP/heteroduplex analysis (2.6.5)
		Sequencing (2.6.6)	Sequencing (2.6.6)
			Mini-sequencing (2.6.7)

Numbers in brackets represent the section number where the methodology is described in this chapter.

2.1 General workflow and laboratory practice

In order to minimise DNA contamination, a separate laboratory was used for single cell isolations, single cell PCR preparations and all oocyte/embryo manipulations, including RNA amplifications. This 'single cell' room was kept under positive pressure to reduce the entry of contaminants. The circulating air in the room was passed through a filter and was completely changed at a rate of 20 times per hour. Primer dilutions, buffer preparations and other PCR solutions were prepared in a laminar flow cabinet within the 'single cell' room. In order to avoid repeated freezing and thawing of reagents, aliquots were prepared whenever possible.

DNase and RNase free microcentrifuge tubes (Eppendorf, UK) were used for all RNA, embryo and single cell work. All RNA work, including RNA extraction, amplification and handling, was conducted in an RNase-free area using nuclease-free tips, water and reagents. RNase Zap (Ambion, UK) solution or wipes were used to clean pipettes, beakers and surfaces of all working areas before beginning any RNA work.

Dedicated equipment (such as thermal cyclers and cold racks) and reagents for single cell PCR, which have not been in contact with previously amplified DNA, were used to avoid any PCR template carry-over. All racks and consumables were exposed to UV irradiation at 254 nm in a Template Tamer to degrade any contaminant DNA.

Second-round PCRs for PGD cases were set up in a laminar flow cabinet in a second laboratory. Post-amplification, samples were handled in a special area of a third laboratory on a different floor to the single cell room.

Human oocytes, embryos and blastomeres were collected from the Assisted Conception Unit (ACU) at University College London Hospital (UCLH) and transported in a closed thermally insulated case (IsoTherm-System®, Eppendorf, UK) back to the UCL Centre for PGD for processing and analysis. Oocytes and embryos were kept in their culture dishes and the tubed blastomeres for PGD were placed on a cold rack during transport.

2.2 Sample collection and processing

2.2.1 IVF protocol, collection of immature oocytes and grading of embryos

Following assessment of the patient's ovarian reserve, the gonadotrophin stimulation was initiated and dose adjustments were made based on the patient's response as described by Muttukrishna *et al.* (2005). Ultrasound guided vaginal collection of immatute (MI) oocytes was conducted at 37 hours post human chorionic gonadotrophin (hCG) administration at the ACU at UCLH. IVF/ICSI was performed at ~ 40 hours post hCG and fertilisation was evaluated at 16–20 hours post insemination. Observation of two pronuclei and two polar bodies was indicative of a normally fertilized oocyte (Sahu *et al.*, 2008). Embryos were cultured in G-1/G-2 PLUS media (Vitrolife, UK).

Preimplantation embryos were graded according to (Bolton *et al.*, 1989) as follows:

Grade 1 Embryo at the correct stage of *in vitro* development with perfect symmetrical and even-sized blastomeres with no fragmentation

Grade 1^{-} Embryo at the correct stage of *in vitro* development with perfect symmetrical and even-sized blastomeres with less than 10% fragmentation

Grade 2^{+} Development with unequally sized blastomeres with less than 20% fragmentation

Grade 2 Retarded development with unequally sized blastomeres with 25%-50% fragmentation

Grade 3 Retarded development with unequally sized blastomeres with more than 50% fragmentation

2.2.2 Tubing of oocytes and blastocysts for gene expression analysis

Donated immature oocytes and surplus cryopreserved blastocysts from routine IVF cases were collected from the ACU (UCLH) for mRNA expression analysis. Informed written consent was obtained from all patients whose oocytes and embryos were used for research purposes.

Immature oocytes at ~ 40 hours post hCG injection were kept in culture in G-IVF PLUS medium (Vitrolife, UK) for a maximum of four hours before rescoring. Fifteen oocytes that had matured in this time were selected for mRNA analysis. Twelve slow frozen blastocysts were thawed and allowed to recover for four hours in G-2 PLUS culture medium (Vitrolife, UK) prior to collection for mRNA analysis.

Using a mouth pipette with a 0.34mm diameter polycarbonate capillary (Biohit, UK), the embryo or oocyte was placed in a drop (10µl) of acidified Tyrode's solution (Medi-Cult, UK) to remove the zona pellucida (ZP) in order to make sure the sample was completely denuded of cumulus cells. The embryo/oocyte was continuously observed under an inverted microscope. Once the ZP had lysed, the embryo/oocyte was immediately moved into a new drop of phosphate buffered saline (PBS, Mg^{2+} and Ca^{2+} free) (Invitrogen/Gibco, UK) containing 0.1% polyvinyl alcohol (PVA) (Sigma, UK) and 0.3 U/µl RNasin® Plus ribonuclease inhibitor (Promega, UK). The RNase inhibitor was added to prevent RNA degradation as the samples were collected for future gene expression analysis. Each sample was washed in fresh PBS/PVA/RNasin® drops at least three times to minimise contamination and to get rid of any remaining cumulus cells. The sample was then transferred in minimum volume (less than 2µl) of PBS/PVA/RNasin® buffer into an empty 0.2ml MicroAmp reaction tube. The presence of a polar body was confirmed in all oocytes during manipulation, indicating that all immature oocytes had matured into MII oocytes during incubation in the standard culture media, prior to tubing. All samples collected were clearly labelled, spun-down in a benchtop microcentrifuge (MSE Microcentaur, Sanyo, UK) and immediately frozen at -80ºC until RNA extraction. All manipulations were conducted rapidly to minimize changes to gene expression, while samples were outside of incubators.

2.2.3 Collection and processing of blood samples

Blood samples were obtained from couples undergoing PGD and from laboratory members who donated blood for the production of anonymous control DNA samples. Blood that was used for the isolation and tubing of single lymphocytes for the PGD workups (described in sections 2.2.3.1 and 2.2.3.2) was collected in lithium heparin tubes. Blood that was used for DNA extractions (described in section 2.3.2) was collected in sodium EDTA tubes. All samples were processed on the day they were received or collected.

2.2.3.1 Separation of lymphocytes from whole blood for PGD workups

Single lymphocytes were isolated from bloods of couples undergoing PGD as part of the protocol workup. The lymphocyte separation from whole blood was carried out in a laminar flow cabinet and all steps were performed at room temperature.

Six millilitres of blood were mixed with 6ml of 0.9% NaCl solution in a 14ml centrifuge tube (Falcon), using a plastic Pasteur pipette. In a separate 14ml centrifuge tube, 6ml of Ficoll-Paque PLUS (GE Healthcare Life Sciences, UK) were added. The tube was tilted and 8ml of the diluted blood was trickled down the side of the tube using a Pasteur pipette so that the blood formed a layer on top of Ficoll-Paque without mixing with it. The tube was centrifuged at 1300rpm for 30min at room temperature in a benchtop centrifuge (Labofuge 400, Heraeus Instruments, UK); the centrifuge was stopped slowly without the use of a brake. Due to differential migration, four layers were formed in the tube consisting of: plasma, buffy coat (lymphocytes and platelets), mixture of Ficoll-Paque PLUS and plasma, and erythrocytes and mononucleocytes (from top to bottom).

The lymphocytes/buffy coat layer was collected using a Pasteur pipette in a fresh 14ml centrifuge tube, which was then filled with 0.9% NaCl up to 12ml. The solutions were mixed and the tube was centrifuged at 1300 rpm for 15minutes. The supernatant was discarded and 2ml 0.9% NaCl was used to resuspend the pellet. The tube was filled with 0.9% NaCl up to 12ml and the centrifugation was repeated. The supernatant was discarded and the lymphocyte pellet was resuspended in 2ml of 0.9% NaCl. Lymphocytes were stored at 4°C and used for single cell isolations within two days from separation.

2.2.3.2 Single cell isolation and tubing of lymphocytes for PGD workups

All manipulations were carried out under an inverted microscope using a mouth pipette with a 0.19mm diameter capillary (Biohit, UK) in the 'single cell' room. One to 5µl of lymphocytes suspended in 0.9% NaCl were sequentially diluted in two to three drops of dissociation buffer (Appendix, section B.1) in order to pick up a single cell. Once a single cell was isolated, it was washed in at least three new drops (10µl) of dissociation buffer (DB). The cell was transferred in minimum volume (~ 1µl) of DB into a 0.2ml MicroAmp reaction tube containing 2.5µl of alkaline lysis buffer (ALB) (Appendix, section B.1), which had been freshly prepared and kept on ice. A small volume (~ 1µl) was taken from the last drop used to wash the cell as the negative control for that cell. The tubes were stored at -80^{o}C for at least 30 minutes (and a maximum period of two weeks) prior to cell lysis and DNA amplification by PCR.

2.2.4 Tubing of blastomeres following embryo biopsy for PGD

Three days post egg collection, embryos that had reached at least the 4-cell stage underwent laser-assisted biopsy in a Mg^{2+} and Ca^{2+} free medium (G-PGD, Vitrolife, UK). Two blastomeres were biopsied from embryos that had 6 to 8 cells. Only one cell was taken from embryos with fewer than 6 cells. The presence of a nucleus was noted at time of biopsy and tubing. If no nucleus was observed, an extra cell was biopsied from the same embryo when possible. Biopsied blastomeres were tubed at the ACU (as described in section 2.2.3.2) using a hand pipette with a 0.19mm diameter capillary and transported to the UCL Centre for PGD laboratory for PCR analysis.

2.3 Preparation of nucleic acids

2.3.1 RNA extraction

Total RNA was extracted from 15 MII oocytes and 12 blastocysts, which were pooled into sets of three oocytes (MIIa-MIIe) or three blastocysts (Ba-Bd). Selection of samples that were pooled was based on the maternal age so that each set contained oocytes from a younger and older patient and the average ages of the pooled sets closely matched (see section 3.1.1 for results).

The RNA extraction was performed using the AllPrep DNA/RNA Micro kit (Quiagen, UK) following the manufacturer's instructions. All steps were carried out at room temperature and centrifugations were performed at 22ºC in a benchtop microcentrifuge (MIKRO 200R, Hettich zentrifugen, Germany). Three or four sample sets were processed together at one time.

Tubed single oocytes or blastocysts (in RNasin solution) were removed from the -80ºC freezer and 75µl of Buffer RLT Plus was added. The solution was pipetted up and down in order to lyse the cells. Pooling of oocytes/blastocysts took place at this stage. The lysate was then transferred to an AllPrep DNA spin column placed in a 2ml collection tube (supplied in kit) and centrifuged at >8000 x g (>10,000 rpm) for 30s. The AllPrep DNA spin column was discarded. One volume (usually 350µl) of 70% ethanol was added to the flow-through and mixed thoroughly by pipetting. The sample, including any precipitate that may have formed, was transferred to an RNeasy MinElute spin column placed in a new collection tube (supplied in kit) and centrifuged for 15s at >8000 x g (>10,000 rpm). 700µl Buffer RW1 were added to the RNeasy MinElute spin column and the tube was centrifuged for 15s at >8000xg (>10,000 rpm) to wash the spin column membrane. The flow-through was discarded carefully and 500µl of buffer RPE were added to the RNeasy spin column. The tube was centrifuged for 15s at >8000 x g (>10,000 rpm) to wash the spin column membrane. The flow-through was discarded, 500µl of 80% ethanol were added to the RNeasy MinElute spin column and the tube was centrifuged for 2min at >8000 x g (>10,000rpm) to wash the spin column membrane. The collection tube with the flow-through was discarded and the RNeasy MinElute spin column was placed in a new collection tube (supplied in kit) and centrifuged at full speed for 5min with the lid open. The collection tube with the flow-through was again discarded.

The RNeasy MinElute spin column was finally placed in a new 1.5ml collection tube (supplied in kit); 14µl DNase and RNase free water (Promega, UK) were added directly to the centre of the spin column membrane and the tube was centrifuged for one minute at full speed to elute the RNA. The RNA samples were stored at -80ºC until time of amplification.

2.3.2 DNA extraction from whole blood

Genomic DNA was extracted from blood samples using the protocol described by Lahiri and Nurnberger (1991). All steps were carried out at room temperature.

Five millilitres of blood was placed in a 14ml centrifuge tube and 5ml of low salt buffer TKM1 (Appendix, section B.2) was added along with 125µl of Igepal (Sigma Chemical Company, UK) to lyse the red blood cells. The contents were mixed well using a Pasteur pipette and centrifuged at 2,200 rpm (1,000g) for 10 minutes in a benchtop centrifuge (Labofuge 400, Heraeus, UK). The supernatant was gently poured off and the pellet was washed repeatedly as before (with 5ml TKM1 and 125µl of Igepal) until the redness of the pellet was reduced. The pellet was resuspended in 100µl of TKM1 and transferred into a 2ml microfuge tube (Eppendorf, UK). 0.8ml of TKM2 (Appendix, section B.2) was added together with 50µl of 10% (w/v) sodium dodecyl sulphate (SDS) and mixed thoroughly (by pipetting up and down) to lyse the white blood cells. The tube was firmly sealed with NescoFilm™ and incubated for at least 30 minutes at 55º C until the pellet was dissolved. 300µl of 6M NaCl were then added and the tube was mixed well. The cells were spun at 12,000 rpm (10,000g) for 5 minutes in a benchtop microcentrifuge (MSE Microcentaur, Sanyo, UK), the supernatant was transferred into a new tube and the precipitated protein pellet was discarded. Two volumes of 100% ice-cold ethanol were added to the supernatant and the tube was slowly inverted until the DNA was precipitated. The precipitated DNA was removed with a 200µl pipette tip and washed in 1ml of ice cold 70% ethanol. The tube was centrifuged at 12,000 rpm (10,000g) for 5 minutes, the supernatant was discarded and the pellet was left to dry in a class two biological safety cabinet. The DNA was resuspended in 50µl of TE buffer (pH 8.0) (Appendix, section B.2). Samples were stored at 4ºC.

2.3.3 Spectrophotometric measurements of DNA concentrations prior to amplification

Genomic DNA samples were diluted in deionised distilled water (in two consecutive steps: 1 in 10 and then 1 in 100 to reach a final dilution of 1 in 1000) prior to the measurement of their concentrations. One millilitre of each diluted sample was needed. The concentrations were calculated from the absorbance at 260nm measured using a spectrophotometer (GeneQuant pro**,** Pharmacia Boiotech, UK). A reading was also taken at 280nm allowing the GeneQuant pro DNA/RNA calculator to produce a ratio between the readings (OD_{260}/OD_{280}). The OD_{260}/OD_{280} ratio was used to estimate the purity of the nucleic acid. Pure DNA and RNA samples have OD_{260}/OD_{280} values of 1.8 and 2.0, respectively. Lower values showed protein contamination.

Quantitation of the double stranded DNA (dsDNA) relied on the following formula:

Amount of ds DNA = Optical Density (OD) at 260nm × 50μg/ml × dilution factor

DNA concentrations needed to be around 100ng/μl for optimal PCR amplification. Samples with very high concentrations were diluted before their use in PCR.

2.4 Amplification methods

2.4.1 RNA amplification and labeling

The purified RNA from oocytes and blastocysts for microarray analysis was amplified in two rounds and digoxigenin (DIG) labelled using the NanoAmp™ RT-IVT Labeling kit (Applied Biosystems, UK) following the manufacturer's instructions (see section 3.1.1 for results). Figure 2.1 summarises the main steps involved. This was performed just after completion of initial assessment of the extracted RNA on the NanoDrop® (section 2.5.2). Unless otherwise specified, all reagents were kept on ice throughout the protocol and all centrifugations were carried out at room temperature.

The first-round of amplification started with a first-srand synthesis step using reverse transcription. The following components were added to a 0.2ml MicroAmp reaction tube and mixed by pipetting up and down: 1µl T7-Oligo (dT) primer, 2µl control RNA (provided in kit) diluted to 1:50,000 and 9µl of the RNA sample. The mixture was heated in a thermal cycler (GeneAmp 9700, Applied Biosystems, UK) at 70ºC for 5 minutes (melting step) then cooled to 4ºC (primer annealing step). The rest of the components were added to the reaction tube in the following order and mixed by pipetting: 2µl 10x1st strand buffer, 4µl dNTP mix, 1µl RT enzyme and 1µl RNase inhibitor. The tube was placed in the thermal cycler with the following program settings: 25ºC for 10minutes (initiation), 42ºC for 2hours (extension), 70ºC for 5 minutes (enzyme inactivation) and hold at 4ºC.

For the synthesis of the second-strand, the following components were added to the cDNA mixture (20µl) on ice and mixed gently: 63µl nuclease-free water, 10µl 10x2nd strand buffer, 4µl dNTP mix, 2µl DNA polymerase and 1µl RNase H. The mixture was placed on the thermal cycler set to: 16ºC for 2hours (second-strand synthesis), 70ºC for 5 minutes (enzyme inactivation), hold at 4ºC. The first-round cDNA was then purified by combining the DNA binding buffer (250µl) and the entire second-strand synthesis reaction (100µl) in a new 1.5ml nuclease-free microcentrifuge tube and mixing thoroughly by pipetting up and down. The DNA Binding buffer/reaction mixture (350µl) was loaded into the DNA purification column placed in its wash tube. The tube was centrifuged at 10,000xg for one minute (MSE Microcentaur, Sanyo, UK) and the liquid flow-through was discarded. The cDNA was washed by adding 500µl of wash buffer to the column and centrifuging the tube at 10,000xg for one minute. The liquid

flow-through was discarded and the tube was centrifuged again at 10,000xg for one minute.

In order to elute the cDNA, the column was transferred to a new DNA elution tube and 10µl of nuclease-free water was pipetted onto the centre of the fibre matrix at the bottom of the column. The tube was kept at room temperature for 2 minutes then centrifuged at 10,000xg for one minute to elute approximately 9µl double-stranded (ds) cDNA. The same elution steps were repeated for a total elution volume of approximately 18µl ds cDNA.

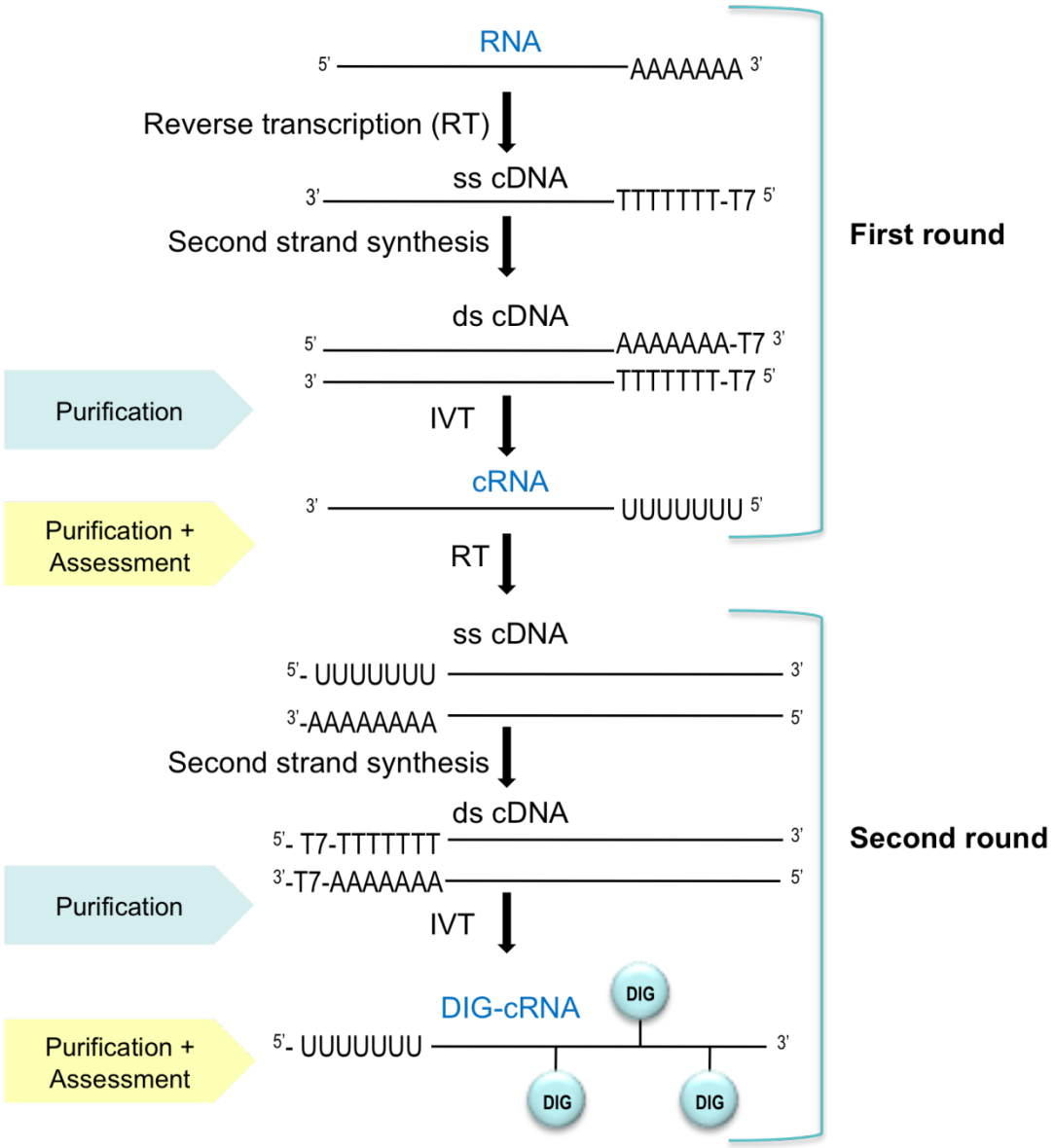
Performing the first-round *in vitro* transcription (IVT) involved adding 4µl 10xIVT buffer, 4µl NTP mix and 4µl IVT enzyme mix to 28µl ds cDNA (volume brought to 28µl with nuclease-free water) in a new MicroAmp reaction tube at room temperature and mixing gently. The tube was incubated at 37ºC for 9 hours (IVT) then cooled to 4ºC (GeneAmp 9700 thermal cycler, Applied Biosystems, UK) and the cRNA was immediately purified. In order to purify the first-round cRNA, the entire IVT reaction (40µl) was combined with 60µl nuclease-free water in a new 1.5ml nuclease-free microcentrifuge tube which was vortexed briefly to mix. The RNA binding buffer (350µl) and 100% ethanol (250µl) were then added and mixed by pipetting. The IVT reaction/RNA binding buffer/ethanol mixture (700µl) was loaded onto an RNA purification column placed in an RNA collection tube. The tube was centrifuged at 10,000xg for one minute and the liquid flow-through was discarded. 650µl of wash buffer were added to the column to wash the cRNA and the tube was centrifuged at 10,000xg for one minute. The liquid flow-through was discarded and the tube was centrifuged again at 10,000xg for one minute. The column was transferred to a new RNA collection tube and 100µl of nuclease-free water was applied onto the fiber matrix at the bottom of the column. The tube was incubated at room temperature for 2 minutes then centrifuged at 10,000xg for one minute. The eluted cRNA was stored on ice while it was transported for concentration then stored at -20ºC. All cRNA samples were concentrated from 100µl to 10µl by vacuum centrifugation (Concentrator Plus, Eppendorf, UK) upon completion of the first-round of amplification and labeling steps. Low-temperature settings were used during the drying process.

The second-round first-strand synthesis step started by thawing the cRNA sample on ice and adding 2µl second-round primers to the 10µl RNA sample in new 0.2ml MicroAmp reaction tube (mixing by pipetting up and down). The mixture was heated to 70ºC for 5minutes (melting step) then cooled to 4ºC (primer annealing) (GeneAmp 9700 thermal cycler, Applied Biosystems, UK). The rest of the reverse transcription reaction was carried out exactly as described above (for the first-round amplification).

The second-strand synthesis step started with the addition of 1µl RNase H to the cDNA mixture (20µl) on ice and mixing gently. The mixture was heated to 37ºC for 30minutes (RNA degradation) then cooled down to 4ºC in the thermal cycler. 5µl T7-Oligo (dT) Primer was added and the mixture was heated to 70ºC for 5 minutes (melting) then cooled to 4ºC. The rest of the reagents were added and mixed gently: 58µl nuclease-free water, 10µl 10x2nd strand buffer, 4µl dNTP mix and 2µl DNA polymerase. The mixture was placed on the thermal cycler set to: 16ºC for 2hours (second-strand synthesis), 70ºC for 5 minutes (enzyme inactivation) and a hold at 4ºC. The second-round cDNA was then purified as described above (for the first-round), resulting in approximately 18µl eluted cDNA.

The second-round IVT labeling involved adding in order 4µl 10× IVT buffer, 8µl DIG-UTP (Roche, UK), 4µl NTP mix, 2µl IVT Control DNA and 4µl IVT enzyme mix to 18µl ds cDNA in a new MicroAmp reaction tube at room temperature and mixing gently. The tube was incubated at 37ºC for 9 hours (IVT) then cooled to 4ºC and the cRNA was immediately purified as described above (for the first-round cRNA purification). The final yield was approximately 100µl of DIG labelled cRNA. A 6µl aliquot was taken from each sample for quality and quantity assessments (section 2.5.2) and the rest was stored at -20ºC until used for microarray analysis (section 2.6.1).

Figure 2.1: Summary of steps involved in RNA amplification (in two rounds) and digoxigenin (DIG) labeling using the NanoAmp™ RT-IVT Labeling kit (Applied Biosystems, UK)



2.4.2 DNA amplification using PCR

2.4.2.1 Design of DNA templates for the formation of heteroduplex DNA constructs using PCR

Heteroduplex DNA constructs that could be easily generated and modified were required for the assessment of MMR functional efficiency in nuclear extracts from human cells. The constructs needed to include insertion/deletion loops (IDLs) of variable sizes or different single base mismatches, which are detectable by MMR proteins. A nick 5' to the mismatch needed to be incorporated in the DNA construct to direct the repair to a specific strand.

2.4.2.1.1 Selection of the DNA sequence

Specific DNA sequences from the entire human genome were selected for amplification by PCR for the production of heteroduplex constructs. The CTG repeat, which is in the 3' untranslated region of the dystrophia myotonica-protein kinase (*DMPK*) gene on chromosome 19, was used in order to allow for size variations that are easily detectable. This particular triplet repeat region was deliberately included since a highly reliable PCR protocol for the amplication of this region had been previously established (Piyamongkol *et al.*, 2001). The rs1981929 SNP site, located upstream of exon 8 in the *MSH2* gene, was used for the formation of DNA constructs with a single base-base mismatch; however, other SNP sites in the human genome could have been used.

The *Entrez Nucleotide* (http://www.ncbi.nlm.nih.gov/sites/entrez?db=nuccore&itool) and the *Ensembl* genome (http://www.ensembl.org/index.html) databases were used to obtain the complete updated genomic DNA sequences of the selected regions (including locations of all known SNP sites).

The *NEBcutter* program (http://tools.neb.com/NEBcutter2/index.php) (Vincze *et al.*, 2003) was then used to select a nicking endonuclease with a single recognition site in the selected sequences, which needed to be upstream of the CTG repeat or SNP site.

2.4.2.1.2 Primer design

The primers were selected using the online *Primer3* software (http://frodo.wi.mit.edu/) so that the amplified product would be in the desired size range (~600bp and ~300bp) and included the desired sequence (Rozen and Skaletsky, 2000). The sequences of each primer set were confirmed to have a single common matching region in the human genome using the online Basic Local Alignment Search Tool, *BLAST*, (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

The primers were tagged so that one strand of each PCR product was biotinylated and the other was fluorescently labelled (for automated genetic analysis on the ABI Prism™ 310) or non-labelled (for SSCP/Heteroduplex analysis). The biotin molecule was incorporated in the products at the 5' ends of a single strand in order to pull out that strand using Dynabead® magnetic particles (Invitrogen, UK).

Oligonucleotides used as primers for PCR and DNA substrates for MMR assessment (Table 2.2 and 2.3) were obtained from Eurogentec Ltd. (UK) in dried form, made up to 50µM working concentrations and prepared into 20µl aliquots before storage at -20 ºC (stock at -80ºC).

Table 2.2: Sequence, chromosome location and modifications of all primers used for the DNA insertion/deletion loop (IDL) repair assay

Two main primer sets were used (DM and MMR) to identify homozygous DNA samples and form DNA constructs containing IDLs. Both Primer sets physically mapped to the long arm of chromosome 19, flanking the CTG triplet repeat in the 3' untranslated region of the *DMPK* gene. The DM primers were previously published (Brook *et al.*, 1992). MMR1 was used with MMR2 or MMR2S to produce PCR products that were ~600bp or ~300bp long, respectively.

Primer	Primer sequence (5'→ 3')	Locus	5' modification	Product size
DM-F	CTTCCCAGGCCTGCATTTGCCCATC	19q13.2-q13.3	FAM (Blue)*	122-200 bp
DM-R	GAACGGGGCTCGAAGGGTCCTTGTAGC	19q13.2-q13.3	-	
MMR1/MMR1B/MMR1FAM (F)	CAGCTCCAGTCCTGTGATCC	19q13.2-q13.3	-/Biotin/FAM (Blue)*	558-639 bp
MMR2/MMR2B/MMR2HEX/MMR2FAM (R)	GTCCTAGGTGGGGACAGACA	19q13.2-q13.3	-/Biotin/HEX (Black)*/FAM (Blue)*	
MMR1 (F)	CAGCTCCAGTCCTGTGATCC	19q13.2-q13.3	-	277-358 bp
MMR2S (R)	TGCACAAGAAAGCTTTGCAC	19q13.2-q13.3	-	

* Colour of fluorescence as visualised on the ABI Prism^{TM} 310 genetic analyzer; F: Forward primer; R: Reverse primer; B: Biotin; bp: base pairs; - : no modification

Table 2.3: Sequence, chromosome location and modifications of primers used for the DNA base-base mismatch repair assay

These primers flank the rs1981929 SNP site located upstream of exon 8 in the *MSH2* gene.

Primer	Primer sequence (5'→ 3')	Locus	5' modification	Product size
rs1981929-B-F/rs1981929-F	TGAGTGCTACATCATCTCCCTTT	2p22-p21	Biotin/-	328 bp
rs1981929-B-R/rs1981929-R	TTGCATACCTGATCCATATCTAAA	2p22-p21	-/Biotin	

F: Forward primer; R: Reverse primer; B: Biotin; bp: base pairs; - : no modification

2.4.2.2 PCR and fluorescent PCR (F-PCR) for formation of heteroduplex DNA constructs

PCR and F-PCR were used to amplify control genomic DNA samples. The PCR was performed in 25µl and 50µl reaction volumes. The general master mix was assembled on ice and consisted of 0.2µM (or up to 0.5µM) of each primer, 0.2 mM dNTPs (dATP, dTTP, dCTP and dGTP) (Promega, UK), 0.05 U/µl of DNA polymerases (AmpliTaq®, AmpliTaq Gold® (Applied Biosystems, UK) or HiFi (Expand High Fidelity PCR System, Roche, UK)), 1× Buffer without $MgCl_2$ and 1.5mM to 2.5mM $MgCl_2$. 5% (v/v) DMSO (Sigma Chemical Company, UK) was added for the MMR primers. Nuclease free water (Promega, UK) was used to adjust the volume to 24µl or 49µl per tube. Finally, 1µl of genomic DNA was added to each reaction tube. One extra tube was prepared with PCR mix only for use as a negative control for each reaction mixture. The tubes were spun down (MSE Microcentaur, Sanyo, UK) and immediately transferred to the thermal cycler (Mastercycler Gradient®, Eppendorf, UK or GeneAmp 9700 thermal cycler, Applied Biosystems, UK). The thermal cycling parameters were determined according to the primers and enzyme used (Table 2.4 lists the general cycling conditions for the different enzymes used).

Table 2.4: Conditions for the thermal cycler using different DNA polymerases

	AmpliTaq®	AmpliTaq Gold®	HiFi	Number of cycles
Initial Denaturation	94ºC for 4min 30s	96ºC for 12min	95ºC for 2min	1
Denaturation*	96ºC for 45s	96ºC for 45s	96ºC for 15s	
Annealing	*X*ºC for 45s	*X*ºC for 45s	*X*ºC for 30s	10
Elongation	72ºC for 1min	72ºC for 1min	72ºC for 1min	
Denaturation	94ºC for 45s	94ºC for 45s	94ºC for 15s	
Annealing	*X*ºC for 45s	*X*ºC for 45s	*X*ºC for 30s	25
Elongation	72ºC for 1min	72ºC for 1min	72ºC for 1min	
Final Extension	72ºC for 10min	72ºC for 10min	72ºC for 7min	1
Cooling	Hold at 4ºC	Hold at 4ºC	Hold at 4ºC	1

HiFi = Expand High Fidelity PCR System
X denotes the annealing temperature, which depended on the primers used and ranged between 50ºC and 64ºC.
* The denaturation temperature for the first 10 cycles was 96ºC to reduce allele dropout.
When AmpliTaq Gold® polymerase was used, the primary denaturation step was extended to 12 minutes to activate the enzyme.

2.4.2.2.1 Optimisation

The melting temperature (Tm) for the oligonucleotide primers was estimated by the following formula:

$$Tm = 2(A+T) + 4(G+C),$$

where A, T, C and G are the nucleotides adenine, thymine, cytosine and guanine respectively. The annealing temperature (Ta) was estimated from the melting temperature using: Ta = Tm – (5 to 10ºC). The working optimal annealing temperatures were determined empirically using the Mastercycler Gradient® thermal cycler (Eppendorf, UK). A temperature gradient PCR covering ±5ºC from the calculated Tm in 2ºC intervals was carried out. The temperature giving the most intense amplified products was considered as the optimal annealing temperature.

Other parameters were changed for the optimisation of PCR protocols: the $MgCl_2$ concentration (1.5, 2.0 and 2.5 mM), the primer concentration (mainly 0.2-0.5μM), the DNA polymerase used and the presence or absence of DMSO.

2.4.2.3 PCR and F-PCR for PGD workups

Each PGD workup involved three main steps:

1) Confirmation of the mutation in the affected partner and his/her affected relatives using primers flanking the mutation site

2) Selection of polymorphic markers that are informative for the couple

Markers are generally considered informative if both partners were heterozygous for the locus and did not share any alleles (i.e. the parental origin of any allele can be identified for the embryo). If one allele is shared, the marker was considered semi-informative. In this project, if the unaffected partner were homozygous for an allele at a linked STR marker that was not shared by the affected partner who was heterozygous at the locus, the marker was still considered informative. Similarly, any marker that was unlinked to the mutation and used as a contamination marker was considered informative as long as the female partner was heterozygous and did not share an allele with her partner.

3) Optimisation of a multiplex PCR protocol

2.4.2.3.1 Selection of markers

The *Ensembl* genome database was used to obtain the complete updated genomic DNA sequences (including locations of all known SNP sites) and to identify the nearest markers to the gene/mutation of interest.

2.4.2.3.2 Primer design

As described in section 2.4.2.1.1, the primers were designed using the online *Primer3* software and checked for matching regions in the human genome using *BLAST*. Primers (Eurogentec Ltd., UK) were made up to 50µM working concentrations and prepared into 20µl aliquots before storage at -20 ºC (stock at -80ºC). Tables 2.5-2.7 list all the primers used for the PGD workups for *MSH2* and *BRCA1*. All of the primers were initially used at 60ºC annealing temperature in singleplex reactions.

Table 2.5: Primers targeting polymorphic markers linked to the *MSH2* gene on chromosome 2p21

Primer	Primer sequence (5'→ 3')	Locus type	5' Modification	Tm (ºC)
D2S119-F	CTTGGGGAACAGAGGTCATT	STR marker	YY	60
D2S119-R	GAGAATCCCTCAATTTCTTTGGA	STR marker	-	64
D2S391-F	GTAATGGAGCCAGTAGGTTACA	STR marker	FAM	64
D2S391-R	AGAGGGTATGATGGAAAAGC	STR marker	-	58
D2S1736-F	GGAAGGCAGGCAAGCATAG	STR marker	FAM	60
D2S1736-R	GGGGAAATGTCCTGACTTGA	STR marker	-	60
D2S1715E-F	GGACACTGCAGACTTGTCCA	STR marker	YY	62
D2S1715E-R	AGACCCTCTTGGCAGCAATA	STR marker	-	60
D2S2227-F	GGGTGGCATATTCTGGTCTC	STR marker	YY	62
D2S2227-R	CGCTGTCCTTCTCTGAATGTC	STR marker	-	64
D2S2563-F	AACCCTTTTTCCATAGTGTTAACTG	STR marker	FAM	68
D2S2563-R	TTTGAAGTCACACTGCGAAGA	STR marker	-	60
D2S2495-F	CCCTGCAGTTAGCAGGATAA	STR marker	DO	60
D2S2495-R	CCGTGGTACTTTGTTATGGC	STR marker	-	60
D2S2767-F	GCCCCTGTCTGAAAATATCTCCC	STR marker	FAM	70
D2S2767-R	ATTCCATTTCCTTTTCCTTGTAGACC	STR marker	-	72
AFM196xf6-F	CATCCCCACCTATGCG	STR marker	DO	52
AFM196xf6-R	ATGAGCCACTGCTCCC	STR marker	-	52
D2S2086-F	TGCTTCGTATCTTTTGGCCT	STR marker	FAM	58
D2S2086-R	GAACACCCTGCCCATACTTG	STR marker	-	62
D2S2548-F	TAATCAACTTGTCAGGGTGTGTG	STR marker	DO	66
D2S2548-R	AAAATTTCATGTGAAGGGGTCA	STR marker	-	60
rs1981929-F	TGAGTGCTACATCATCTCCCTTT	SNP	-	66
rs1981929-R	TTGCATACCTGATCCATATCTAAA	SNP	-	64
MSH2_4SNP-F	TTTCTGGAGAAGTTTGGGAAC	4 SNPs	-	60
MSH2_4SNP-R	AGCAATCCCAGCTCTGCTAC	4 SNPs	-	62

F: Forward primer; R: Reverse primer; Tm: melting temperature; - : no modification

Table 2.6: Primers targeting *BRCA1* mutations and linked STR markers on chromosome 17q21

Primer	Primer sequence (5'→ 3')	Locus type	5' Modification	Tm (ºC)
BRCA1 c.3339T>G-F	CAGTGAGCACAATTAGCCGTA	mutation	-	62
BRCA1 c.3339T>G-R	CTCAGGTTGCAAAACCCCTA	mutation	-	60
BRCA1 c.68_69delAG-F	GCTCTTCGCGTTGAAGAAGT	mutation	FAM	60
BRCA1 c.68_69delAG -R	GGTCAATTCTGTTCATTTGCAT	mutation	-	60
D17S1338-F	TCACCTGAGATTGGGAGACC	STR marker	YY	62
D17S1338-R	GGGCAGGAATGGGTTTTAG	STR marker	-	58
D17S1185-F	GGTGACAGAACAAGACTCCATC	STR marker	FAM	66
D17S1185-R	GGGCACTGCTATGGTTTAGA	STR marker	-	60
D17S855-F	GTGCAAGACTGCGTCTCAAA	STR marker	DO	60
D17S855-R	CCTTGACAGACAGACGGACA	STR marker	-	62
D17S1343-F	AAGGGCAGTGTGACCAAAAG	STR marker	YY	60
D17S1343-R	AGCCTGGGTAACAAGAGCAA	STR marker	-	60

F: Forward primer; R: Reverse primer; Tm: melting temperature; - : no modification

Table 2.7: Primers targeting STR markers that are unlinked to the MSH2 or *BRCA1* genes

Primer	Primer sequence (5'→ 3')	5' Modification	Tm (ºC)
D17S1294-F	TGGCATGCAATTGTAGTCTC	FAM	58
D17S1294-R	TTCTTTCCTTACTAAGTTGAGAACG		68
D17S1800-F	CTAAACTAGGTTGGGTTGAAATCTC	TET	70
D17S1800-R	TCTGGCACAAAGACCTGAG		58
D17S841-F	TGGACTTTCTTACATGGCAG	HEX	58
D17S841-R	AGGTTAGTAGTCTATGTCACAGCG		70
APOC2-F	GGCTACATAGCGAGACTCCATCTCC	FAM	61
APOC2-R	GGGAGAGGGCAAAGATCGATAAAGC	-	59
DM-F	CTTCCCAGGCCTGCATTTGCCCATC	FAM	63
DM-R	GAACGGGGCTCGAAGGGTCCTTGTAGC	-	66
D19S112	GCCAGCCATTCAGTCATTTGAAG	HEX	55
D19S112	CTGAAAGACACGTCACACTGGT	-	55
D13S168-F	CTCAGGAAGAAGGAGGCTCA	FAM	62
D13S168-R	CTGCTTGCTTGTGCCTATGT		60
D13S262-F	CTAAAACAAACAAATACAACTCC	FAM	60
D13S262-R	TTGCTCAAGATGAAGTCATG		56
RBivs20-F	CTTCACCTTCTCTCCTCCCTAC	FAM	68
RBivs20-R	GGGTAACAGAGTGAGACTCTATC	-	68
RB1.20-F	GGCATTTGGACCAAGTAAGAA	DO	60
RB1.20-R	GTTGCAGTGAGCCGAGATTG	-	62
IVS26-2.3-F	AGGCCAGGAGTTCAAGACCA	FAM	62
IVS26-2.3-R	ATGAGCCACTGTGCCCAATC		62

F: Forward primer; R: Reverse primer; Tm: melting temperature; - : no modification

Genomic DNA from the couple undergoing PGD and control samples were used for direct mutation detection and to assess the informativity of the polymorphic markers. Following the identification of suitable informative markers (ideally, two linked markers with one on either side of the gene/mutation), the PCR multiplex protocol was optimised with genomic DNA and then on single lymphocytes. The optimised PCR was finally tested on 50 single lymphocytes in order to calculate amplification efficiency (AE) and ADO frequency at each locus, where AE = 'number of cells with amplified product/total number of cells' and ADO rate = 'number of cells with ADO/total number of cells with amplified product'. AE >90% and ADO <10% were considered suitable for clinical application.

The false negative and false positive rates were also calculated for the final PGD protocol using cells of known genotypes:

- the false negative rate (FN rate) = number of cells diagnosed as normal/total number of cells tested from the affected partner
- the false positive rate (FP rate) = number of cells diagnosed as affected/total number of cells tested from the unaffected partner

When the diagnosis is based on two cells, the false negative and flase positive rates are equal to $(\text{FN rate per cell})^2$ and $(\text{FP rate per cell})^2$, respectively.

PCR and F-PCR were used to amplify genomic DNA as well as DNA released from lysed single cells for PGD workups and during a PGD case. The PCRs were performed in 25µl reaction volumes and HiFi (Roche) polymerase was always used. The master mixes were set up as described in section 2.4.2.2 and the cycling conditions are listed in Table 2.8. The initial standard conditions for all PCR master mixes were: 0.2µM of each primer, 0.2mM dNTPs, 0.05 U/µl of HiFi (Expand High Fidelity PCR System) enzyme and 1× HiFi Buffer 2 (containing 1.5mM $MgCl_2$). The initial $MgCl_2$ concentration tested was 1.5mM; when higher concentrations (2.0 or 2.5mM $MgCl_2$) were necessary HiFi Buffer 3 was used. Glycerol (Sigma Chemical Company, UK) at 10% v/v was added to some multiplex PCRs. In order to amplify all targeted loci efficiently, the multiplex PCRs were sometimes split into two rounds of amplification. The first round PCR included the primers for all targeted loci (i.e. all STR markers or mutation + STR markers) and involved a fewer number of cycles (Table 2.8). The second round PCR was carried out in two amplification reactions each including one or more of the primer sets.

2.4.2.3.3 Single cell lysis and PCR

For the amplification of DNA from single cells, 200mM Tricine was added to the PCR master mix to neutralise the ALB. 21.5µl of the reaction mix were added to the tubes containing 3.5µl ALB with DNA from a lysed cell. The tubes containing the single cells were taken out of -80^{o}C and incubated at 65^{o}C for 10 minutes then cooled to 4ºC (GeneAmp 9700, Applied Biosystems, UK) to release the DNA just before adding the PCR reaction mix. In addition to the negative control for the master mix, positive controls (consisting of genomic DNAs extracted from the same individual from whom the single cells were isolated) and blanks from single cell washes were prepared for these PCRs.

Table 2.8: Conditions for the thermal cycler using the Expand High Fidelity PCR System (HiFi) for multiplex PCRs in single or split amplification reactions

PCR step	Cycling conditions	Number of cycles for single reaction	Number of cycles for multiplex PCR in split reactions	
			First round	Second round
Initial Denaturation	95ºC for 2min	1	1	1
Denaturation*	96ºC for 15s			
Annealing	*X*ºC for 30s	10	10	10
Elongation	72ºC for 1min			
Denaturation	94ºC for 15s			
Annealing	*X*ºC for 30s	30	5	30
Elongation	72ºC for 1min			
Final Extension	72ºC for 7min	1	1	1
Cooling	Hold at 4ºC	1	1	1

X denotes the annealing temperature, which depended on the primers used and ranged between 50ºC and 64ºC.
* The denaturation temperature for the first 10 cycles was 96^{o}C to reduce allele dropout.
The same cycling conditions were used for genomic DNA or single cells during PGD workups

2.5 Assessment of amplified products

2.5.1 Agarose gel electrophoresis

In order to assess RNA and DNA amplification efficiency, samples were visualised by electrophoresis on 1, 1.5 and 2% (w/v) agarose gels with 0.04µg/ml ethidium bromide. 0.5-1g agarose (Sigma Chemical Company, UK) was heated in 50ml of 1xTBE buffer (Appendix, section B.3) to which was added 1µl of ethidium bromide solution (10mg/ml). The gel was poured into a mould and allowed to set at room temperature.
RNA and DNA samples were prepared in 3µl and 5µl volumes, respectively, and mixed with 1µl loading buffer (Appendix, section B.3) before loading onto the gel. 0.8µg (8µl of 1:10 dilution, denatured at 70ºC for 10minutes) of the 0.5-10 Kilobase (Kb) RNA ladder (Invitrogen, UK) was used per well for the assessment of RNA samples and 1.5µl of the 1 Kilobase (Kb) ladder (Hyperladder IV, Bioline) was used as a reference for all PCR products. Electrophoresis was carried out at 50-75V for 20-40 minutes. The gels were visualised under ultraviolet (UV) light in the MultiImage Light Cabinet (AlphaInnotech Corporation, UK) and photographed using the AlphaImager 1220 software (version 5.04).

2.5.2 Assessment of RNA quality

The quality of amplified RNA was assessed before proceeding with the microarray analysis. Assessment was carried out on 1.5% agarose gels as well as using the NanoDrop® ND-1000 spectrophotometer (Isogen Life Science, Netherlands) and the Agilent 2100 Bioanalyzer (Agilent, CA, USA). A minimum of 3µl RNA was necessary for analysis on agarose gels, while 1.2µl of each sample was sufficient for the NanoDrop® and Agilent 2100 Biolanalyzer.

Using the Nanodrop, 1.2µl of sample was directly applied onto the sensor to obtain a reading of the concentration as well as two ratios: OD_{260}/OD_{280} and OD_{260}/OD_{230}. The OD_{260}/OD_{280} ratio was used to estimate purity; pure RNA has an OD_{260}/OD_{280} value of 2.0. The OD_{260}/OD_{230} ratio can detect the level of salt carryover and is preferred to be greater than 1.5 as lower ratios indicate greater amounts of salt in the sample.

The Agilent RNA 6000 Kit, which is specifically designed for total RNA analysis on the Agilent 2100 Bioanalyzer, was used with the Nano Chip (suitable for analysis of 5-500ng/µl total RNA) for the amplified RNA samples and the Pico Chip (suitable for analysis of 0.2-5ng/µl total RNA) for the original RNA samples (i.e. non-amplified extracted RNA).

2.6 Processing and analysis of amplified products

2.6.1 Microarray analysis

2.6.1.1 Hybridisation, microarray processing, scanning and result analysis

Triplicate sets of three pooled human MII oocytes or three pooled human blastocysts were analysed for mRNA expression using microarrays. For each sample, 10µg amplified and digoxigenin-labeled RNA was applied onto a Human Genome Survey Microarray V2.0 (Applied Biosystems, UK), which included 32,878 oligo probes interrogating 29,098 genes. Hybridisation, chemiluminescence detection (using Applied Biosystems' Chemiluminescence detection kit), image acquisition on the Applied Biosystems 1700 Chemiluminescent Microarray Analyzer and initial analysis of the raw data using the Spotfire© Integromics™ AB1700 Application Package were carried out at the Molecular Histopathology laboratory at the Trinity Centre for Health Sciences (St James's Hospital, Dublin, Ireland).

2.6.1.2 Functional annotations and detection of genes involved in DNA damage response pathways

The Panther online database (http://www.pantherdb.org/) was used for functional annotations (Mi *et al.,* 2005) and a comprehensive list of 154 DNA repair genes was obtained from a supplement table to a review by Wood *et al.* (2005) updated in February 2008. All genes listed were known to be involved in DNA repair pathways. The online table provides the OMIM links for each gene. Experimental evidence showing functional DNA repair roles of these enzymes was checked from the 'database of mouse strains carrying targeted mutations in genes affecting biological responses to DNA damage' (Friedberg and Meira, 2006).

Twelve commonly used housekeeping genes (Mamo *et al.*, 2007) were investigated as controls. Additionally, genes known to be over-expressed in cumulus cells in the human (*LHCGR*, *BMPR2*, *TNFSF11/RANKL*, *SEMA3A*, *C7*, *CD200)* (Assou *et al.*, 2006) and in the mouse (*AREG*, *EREG* and *BTC)* (Hernandez-Gonzalez *et al.*, 2006) were examined to confirm that the oocytes were completely denuded. Three genes associated with pluripotency were investigated: *NANOG*, *POU5F1* (*OCT3/4*) and *SOX2*.

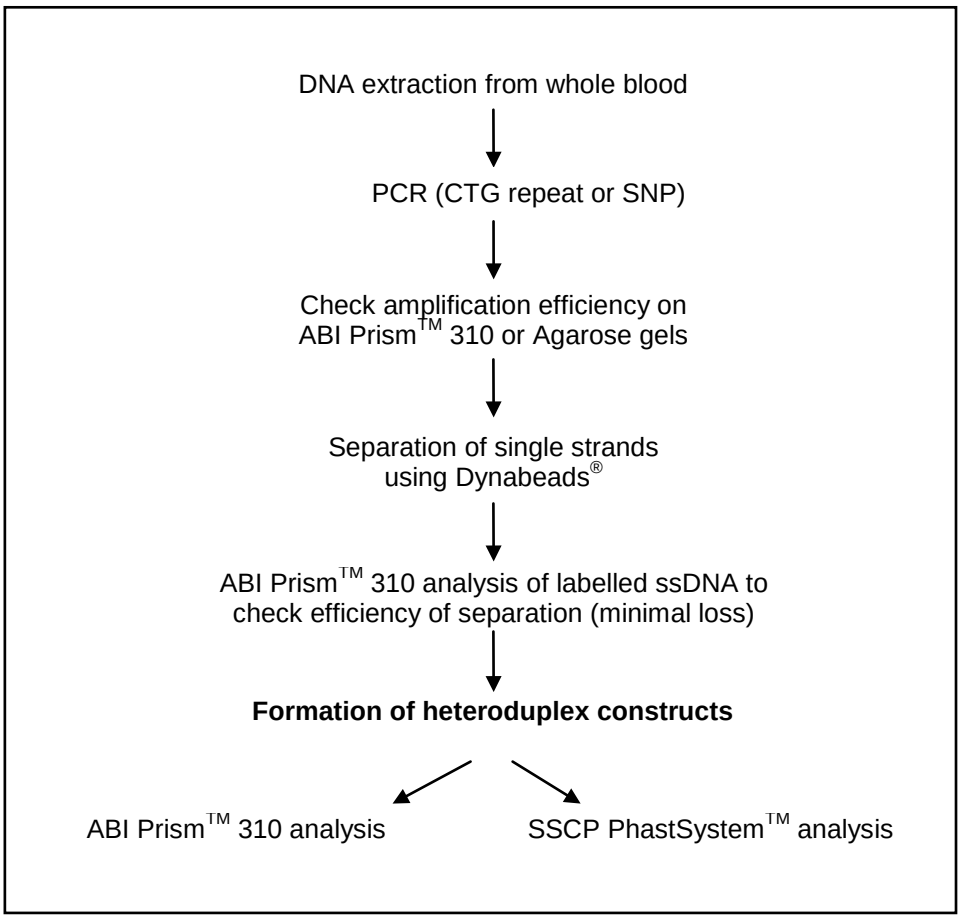
Signals from each microarray were filtered prior to analysis to eliminate all probes with a signal to noise (S.N.) ratio less than 3 and flags over 5000 from the data sets. Expression levels of all genes were grouped into three categories (high, medium and low) based on the signal values after removal of the outliers from the data (top and bottom 5% signals detected). Signals were then split into three equal groups; the medium expression level ranged from 7,000 to 29,500. A t-test analysis was performed and genes were characterized as differentially expressed for p-values <0.05.

In order to allow analysis of the relative differences in gene expression within matured oocytes and blastocysts, the structure specific flap endonuclease-1 (*FEN1*) gene was used as the reference gene. FEN1 removes 5' single-stranded flap structures, which arise during DNA replication, recombination and strand displacement DNA synthesis in BER (Liu et al., 2004); its activity is crucial for maintaining the stability of repeat sequences in the genome.

2.6.2 Formation of heteroduplex DNA constructs for functional assessment of DNA repair

Heteroduplex DNA constructs were prepared by hybridising single strands of DNA that were complementary except at the site of mismatch, which was either the rs1981929 SNP or the *DMPK* (CTG)n repeat. The single strands were prepared from homozygous PCR products. The sequence of steps involved in the formation of heteroduplex DNA molecules is described in Figure 2.2.

Figure 2.2: Summary of experimental work for the formation of heteroduplex constructs



2.6.2.1 Selection of DNA samples for use in the preparation of heteroduplex constructs

Genomic DNAs from homozygous individuals, with different alleles at the *DMPK* gene CTG repeat locus or the rs1981929 SNP site, were used to generate PCR products for the formation of heteroduplex constructs with IDLs and single base mismatches, respectively. Twelve control genomic DNA samples were amplified by fluorescent PCR (F-PCR) using the DM primers and were analysed on the ABI Prism™ 310/GeneScan™ analysis (method described in section 2.6.4 below). Homozygous samples were selected for heteroduplex formation based on allele sizes. By changing the samples used, i.e. changing the allele sizes, it was possible to generate various IDL heterologies. Nine control genomic DNA samples were amplified using the rs1981929 primer set (Table 2.3). The products were assessed using single-strand conformation polymorphism and sequencing analysis (methods described in sections 2.6.5 and 2.6.6). Two homozygous samples were selected based on the genotype (AA or GG for the SNP) for the formation G.T and A.C heteroduplex DNA constructs.

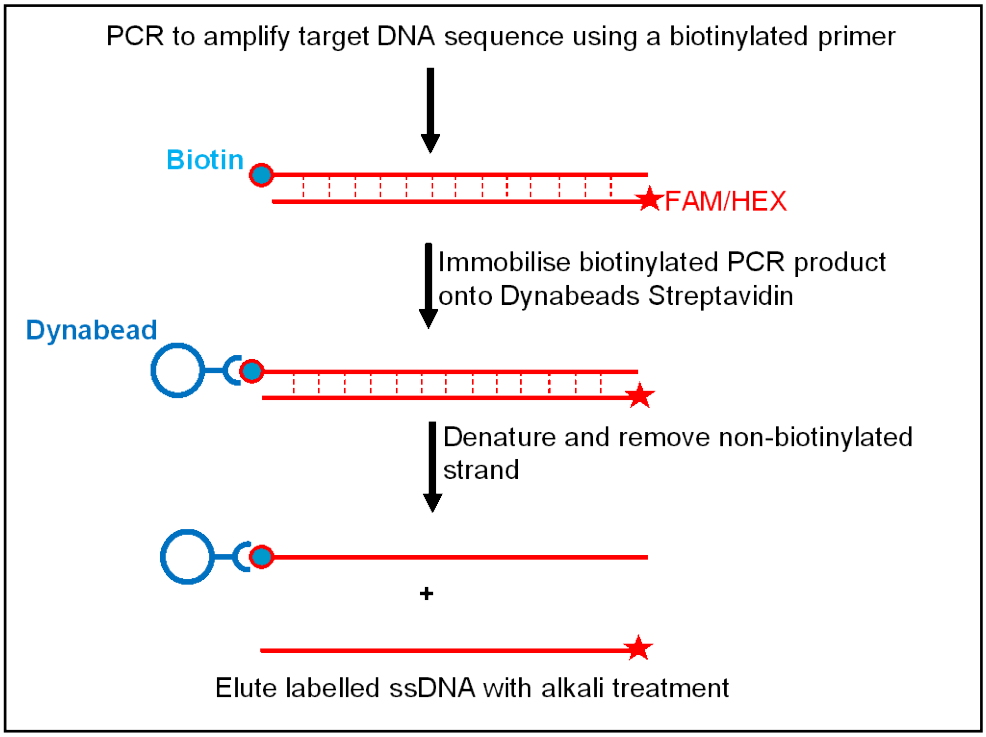
2.6.2.2 Separation of DNA strands

The separation of the strands was achieved using Dyanbeads® Streptavidin M-270 (Invitrogen, UK) in the presence of a magnetic particle concentrator (MPC) rack, following the manufacturer's instructions (see Appendix, section B.4 for solution compositions).

The Dyanbeads® were washed to remove the sodium azide (NaN_3) before adding the DNA. 20µl of beads were added to a 0.5ml tube. The tube was then placed on the MPC rack for 1 to 2 minutes. The supernatant was removed and discarded by aspiration with a pipette without touching the sides of the tube. The tube was then removed from the MPC rack and 20µl of 2× B&W (Binding and Washing) buffer were added and mixed by pipetting up and down. The washes were repeated and after the supernatant was removed, the Dyanbeads® were re-suspended in 40µl of 2× B&W buffer to a final concentration of 5µg/µl. 40µl of amplified biotinylated DNA (the PCR product) were added and the tube was incubated for 15 minutes at room temperature using gentle rotation (R100 Rotatest Shaker, Luckham Ltd., UK) to keep the beads suspended. The beads, coated with biotinylated DNA fragments, were separated using the MPC rack. This was followed by two washes with 1× B&W buffer using the MPC rack. The Dyanbeads® with immobilised PCR products were re-suspended in 40µl 1× B&W buffer. The duplex DNA was subsequently melted by exposure to 0.1M NaOH.

First, the Dyanbeads®/DNA complexes were washed twice with 40µl 2× B&W buffer. 10µl of freshly prepared 0.1M NaOH were added and the tube was incubated at room temperature for 5 minutes using gentle rotation. The tube was then placed on the MPC rack and the NaOH supernatant was transferred to a new tube. The beads with the immobilised biotinylated DNA strand were washed once with 50µl of 0.1M NaOH, once with 50µl of 1× B&W buffer and once with 50µl of TE buffer. The eluted stand was neutralised using the MinElute PCR Purification kit (QIAGEN, UK) following the manufacturer's instructions (see Appendix, section B.5). The single stranded DNA (ssDNA) was re-suspended in 10µl of deionised distilled water and stored at 4ºC until used. Figure 2.3 illustrates the main steps involved in the separation of single strands.

Figure 2.3: Summary of experimental steps involved in the separation of single strands



2.6.2.3 Formation of heteroduplex constructs

Heteroduplex molecules containing an insertion/deletion loop (IDL) or a single base mismatch were created by hybridisation of purified complementary single strands of DNA with different CTG repeat sizes or mismatched bases at the SNP site. The single strands were either purchased or isolated from two PCR products.

Synthesised oligonucleotides (180 bases in length) were obtained from Eurogentec Ltd. (UK). The ssDNA fragments were named ‘strand A’, ‘strand G’ or ‘strand T’ according to the base at the SNP site. The reverse nicked strand was also ordered in two fragments ‘strand Ta’ and ‘strand Tb’ (for full sequences see Table 2.9). All ssDNA fragments were prepared at a 0.5µM concentration before mixing the complementary strands and incubating them at 37ºC overnight to form heteroduplex molecules.

Table 2.9: Complete sequence and fragment length of synthetic oligonucleotides used for the formation homoduplex and heteroduplex DNA molecules with a G.T mismatch

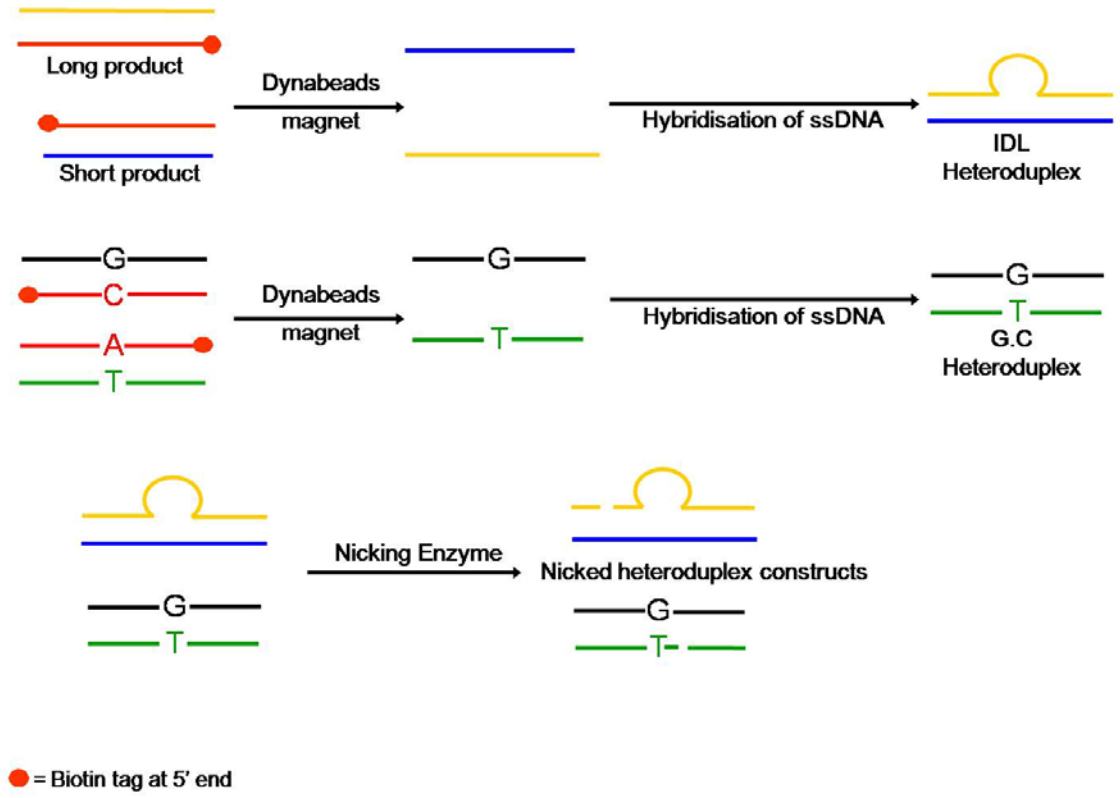
Strand	Oligonucleotide sequence	Fragment size
Strand A	5'-TAAAATAAATTGAGTACGAAACAATTTGAATTAAAACACCTGA GTAAATAGTAACTTTGGAGACCTACTGTACTATTTGTACCTTTTG GATCAAATGATGCTTGTTTATCTCAGTCAAAATTTTATGATTTGTA TTCTGTAAAATGAGATCTTTTTATTTGTTTGTTTTACTACTTTCTT-3'	180 bases
Strand G	5'-TAAAATAAATTGAGTACGAAACAATTTGAATTAAAACACCTGA GTAAATAGTAACTTTGGAGACCTGCTGTACTATTTGTACCTTTTG GATCAAATGATGCTTGTTTATCTCAGTCAAAATTTTATGATTTGTA TTCTGTAAAATGAGATCTTTTTATTTGTTTGTTTTACTACTTTCTT-3'	180 bases
Strand T	5'- AAGAAAGTAGTAAAACAAACAAATAAAAAGATCTCATTTTACAG AATACAAATCATAAAATTTTGACTGAGATAAACAAGCATCATTTGA TCCAAAAGGTACAAATAGTACAGTAGGTCTCCAAAGTTACTATTTA CTCAGGTGTTTTAATTCAAATTGTTTCGTACTCAATTTATTTTA-3'	180 bases
Strand Ta	5'-AAGAAAGTAGTAAAACAAACAAATAAAAAGATCTCATTTTACAG AATACAAATCA-3'	55 bases
Strand Tb	5'-TAAAATTTTGACTGAGATAAACAAGCATCATTTGATCCAAAAGG TACAAATAGTACAGTAGGTCTCCAAAGTTACTATTTACTCAGGTGT TTTAATTCAAATTGTTTCGTACTCAATTTATTTTA-3'	125 bases

The sequences were selected from human chromosome 2 (2p22-p21) around the same SNP site (rs1981929 in the *MSH2* gene) targeted with the primers described in Table 2.3. The A/G SNP site is highlighted in green (on the forward strand) and yellow (on the reverse strand).

Equal volumes of each single strand, isolated from PCR products of similar concentrations, were mixed in a 0.5ml tube. The concentrations were estimated based on the band intensities observed on gels after electrophoresis or from the peak heights obtained on the ABI Prism™ 310 (sections 2.6.4.1 and 2.6.4.2). Equal volumes of strands A/G, T and water or strands A/G, Ta and Tb were mixed to form non-nicked and nicked heterdoduplex molecules at 0.167µM. The mix was denatured at 95ºC for 5 minutes and incubated at 37ºC overnight to hybridise. Figure 2.4 summarises the steps involved in the formation of heteroduplex DNA constructs with an IDL or single base mismatch.

A nick was introduced upstream of the loop in order to direct the repair to a specific strand. This was accomplished by exposing the heteroduplex DNA molecules to the Nt. *Bbv*CI or Nt. *Bst*NBI restriction endonuclease (New England BioLabs, UK). In a 50µl reaction, 20µl heteroduplex sample was mixed with 24µl nuclease free water (Promega, UK), 5µl 10× NE buffer 2 (Appendix, section B.6) and 1µl Nt. BbvCI/Nt. BstNBI. The mix was incubated for one hour at 37ºC. The product was cleaned-up using the MinElute PCR Purification kit (QIAGEN, UK) prior to analysis on the ABI Prism™ and SSCP gels.

Figure 2.4: Formation of heteroduplex constructs with IDLs or single base mismatches

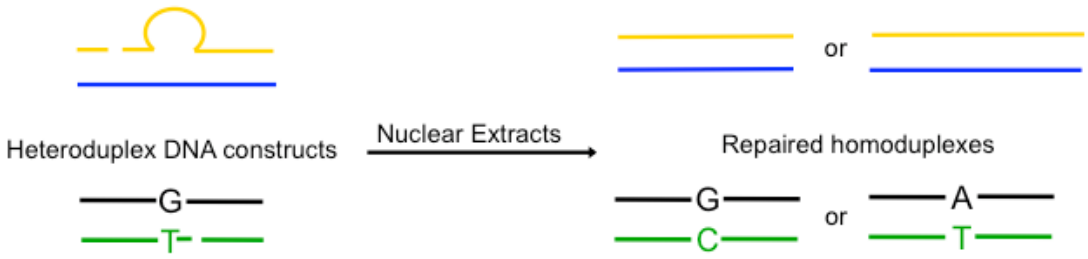


ssDNA fragments were isolated from homozygous PCR products with different CTG repeat sizes or different bases for the rs1981929 SNP. Complementary long (yellow) and short (blue) DNA strands were mixed together to form heteroduplex molecules containing IDLs. Complementary DNA strands with a G (black) and T (green) base at the SNP locus were mixed together to form heteroduplex molecules containing a G.T mismatch. A nick was introduced 5' to the mismatch by exposing the IDL or G.T heteroduplex DNA constructs to the Nt. *Bbv*CI or Nt. *Bst*NBI nicking endonucleases, respectively.

2.6.3 Exposure of heteroduplex constructs to nuclear extracts

The heteroduplex construct was exposed to nuclear extracts from HeLa S3 (MMR efficient) and LoVo (MSH2 deficient) cells (Active Motif, Belgium). Formation of homoduplex DNA molecules was indicative of repair (Figure 2.5). The procedure was based on the published method by Wang and Hays (2002a).

Figure 2.5: Formation of homoduplex DNA after exposure of the heteroduplex constructs to nuclear extracts

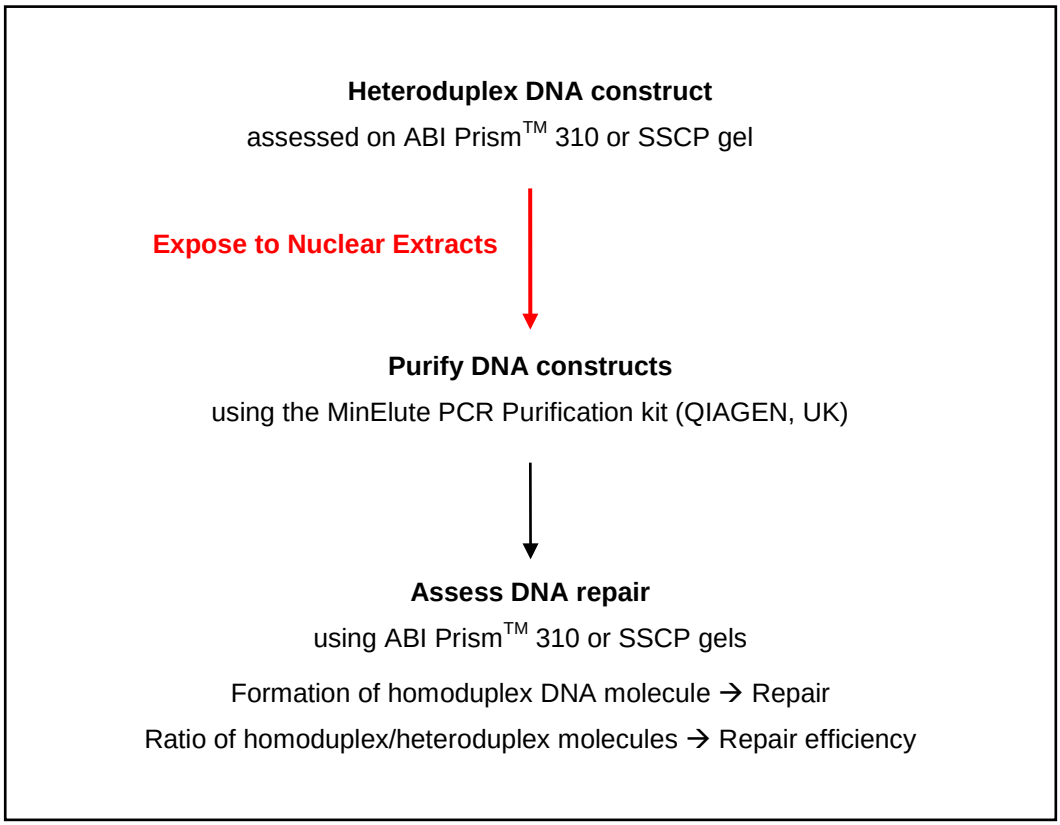


Exposure of the heteroduplex constructs to nuclear extracts would result in the formation of homoduplex DNA molecules if the extracts contained all necessary DNA repair proteins.

The standard MMR reaction mixture (15µl reaction volume) consisted of ~100ng of the DNA construct (substrate) to which 2-50µg (~0.3-10µl) of the nuclear extract was added along with 50ng/µl acetylated BSA and 3µl 5× solution 1 (Appendix, section B.7). The negative control tube was identically prepared except for the addition of nuclear extracts, which was replaced with nuclease free water. The reaction mixture was set-up on ice keeping all the reagents cool at all times to prevent loss of protein activity in the extracts. The mixture was incubated at 37ºC for 5-60 minutes (GeneAmp 9700 thermal cycler, Applied Biosystems, UK). The reaction was stopped by the addition of a stop solution containing Proteinase K (Appendix, section B.7) followed by an incubation at 37ºC for 30 minutes then at 75ºC for 15 minutes (enzyme deactivation step) (GeneAmp 9700 thermal cycler, Applied Biosystems, UK). The DNA constructs were immediately purified using the MinElute PCR Purification kit (QIAGEN, UK) and recovered in 10µl nuclease free water.

In order to assess repair, the products were directly analysed on the ABI Prism^{TM} 310 or heated at 95 ºC for 5 minutes followed by an overnight incubation at 37ºC (to allow duplex structures to reform) and analysed on SSCP gels. Figure 2.6 summarises the main steps involved in the *in vitro* assessment of DNA MMR and IDL repair.

Figure 2.6: Summary of experimental work flow for the assessment of MMR or IDL repair



2.6.4 ABI Prism™ genetic analysis

Fluorescent PCR products were analysed using an automated laser DNA analyzer, the ABI Prism™ 310 with the Genescan™ software (version 3.7.1 for large fragment analysis). The ABI Prism™ 3100 was used for sequencing analysis and both the ABI Prism™ 3100 and 3730 were used with the GeneMapper™ software for PGD workups and cases only.

2.6.4.1 Use of the ABI 310 Prism™ for heteroduplex formation and analysis

One microlitre of fluorescent PCR product (diluted to 1/10 when genomic DNA was amplified) was mixed with 12µl of deionised formamide (HiDi™ formamide, genetic analysis grade, Applied Biosystems, UK) and 0.5µl of Map Marker® 1000 TAMRA (50 to 1000 bases) (Bio Ventures Inc., USA) size standard. Samples were denatured at 95ºC for 5 minutes and loaded onto the genetic analyzer.

The denatured sample was subjected to capillary electrophoresis using Performance Optimised Polymer 4 (POP-4 C module; 5 second injection time, 15,000V, 60ºC, 20-40 minutes). The products were sized using the GeneScan™ analysis software (Applied Biosystems, UK).

2.6.4.1.1 Normalisation of peak heights obtained with FAM and HEX on the ABI Prism™ 310 genetic analyzer

The peak heights and peak areas obtained from GeneScan™ analysis reflect the relative amounts of labelled DNA detected. The fluorescence intensities for different colours/labels are detected with different sensitivities on the ABI Prism™ 310 genetic analyzer. Thus, equal amounts of DNA labelled with different fluorescent molecules may give different peak heights due to variation in absorbance efficiency. In order to normalise the peak heights, double labelled PCR products were produced (using the MMR1FAM and MMR2HEX primers together). Those products were run on the ABI Prism™ 310 and the ratio of the FAM peak height over the HEX peak height (FAM/HEX) was calculated.

2.6.4.1.2 Quantitative analysis using the ABI Prism™ 310 genetic analyzer

The HEX/TAMRA or FAM/TAMRA height ratios of same sized fragments were calculated from a serial dilution of HEX or FAM labelled products and the TAMRA marker. The ratios were plotted against the dilution factors. The linear regression slopes determined the TAMRA/HEX or TAMRA/FAM ratios, which were 0.42 and 0.63, respectively. A serial dilution of MapMarker® 1000 TAMRA (8fmol/band/µl) was run with each Prism™ run. This made it possible to plot the TAMRA peak heights or peak areas against the corresponding concentrations produced from the serial dilution, forming a standard dose response curve for that run. Since TAMRA was present in every individual sample that was run, the peak height of the sample (labelled with HEX or FAM) was multiplied by the TAMRA/HEX or TAMRA/FAM ratio in order to approximate the peak it would display were it labelled with TAMRA. The concentration of the sample was then extrapolated from the standard curve. This was done to approximate the concentrations of ssDNA and dsDNA (homoduplex or heteroduplex) samples.

2.6.4.1.3 Semi-quantitative assessment of DNA repair

Semi-quantitative analysis was used to measure relative repair efficiencies and helped overcome the issue of incomplete nicking of the heteroduplex constructs. The ratio of the peak heights from the nicked strand and the complementary non-nicked strand (R-value) was calculated for heteroduplex constructs containing IDLs. The R-values represented the ratio of fluorescence from the short strand and the complementary long strand when heteroduplexes were not nicked. (R = peak height of nicked or short strand/peak height of non-nicked or long strand).

IDL repair efficiency was assessed by comparing the R-value before and after exposure of the construct to nuclear extracts (N.E.): $\Delta R = R_{\text{negative control}} - R_{\text{sample exposed to N.E.}}$. Since the R-value of a given sample can change after processing by up to 0.1 in the absence of nuclear extracts, i.e. $R_{\text{negative control}} - R_{\text{unprocessed heteroduplex sample}} \leq 0.1$, a ΔR-value $\neq 0$ was considered indicative of repair only if $R_{\text{sample exposed to N.E.}} - R_{\text{unprocessed heteroduplex sample}} > 0.1$.

2.6.4.2 ABI Prism™ genetic analysis for PGD

One microlitre of fluorescent PCR product (diluted to 1/10 when genomic DNA was amplified) was mixed with 12µl deionised HiDi™ formamide and 0.5µl of size standard. GeneScan®-500 ROX and LIZ (50 to 500 bases) (Applied Biosystems, UK) were used for the ABI Prism™ 310/3100 and 3730, respectively. Samples were denatured at 95ºC for 5 minutes and loaded onto the genetic analyzer. The denatured sample was subjected to capillary electrophoresis using Performance Optimised Polymer (POP-4 G5 module; 5 second injection time, 15,000V, 60ºC, 18-25 minutes). The products were sized using the GeneScan™ and GeneMapper™ analysis software (Applied Biosystems, UK) for the ABI 310 and 3100/3730, respectively. Preparation of the Prism and loading of samples was carried out by Seema Dhanjal whenever the ABI Prism™ 3730 was used.

2.6.5 SSCP/Heteroduplex analysis using the PhastSystem™

2.6.5.1 Preparation of samples

All PCR products were initially checked on agarose gels to confirm DNA amplification. One microlitre of each PCR product was mixed with 2μl of formamide in a 0.5ml tube then denatured by heating at 95ºC for 10 minutes after mixing. The denatured samples were kept on ice (to avoid re-annealing of the single strands) until they were loaded on an ultra-thin pre-cast 12.5% or 20% non-denaturing polyacrylamide gel (PhastGel® Homogenous 12.5 or 20, GE Healthcare Life Sciences, UK).

Isolated ssDNA and duplex DNA constructs used for the functional assay were not denatured; they were directly loaded onto the pre-cast gels without the addition of formamide.

2.6.5.2 Sample loading and electrophoresis on the PhastSystem™

A well template for the placement of the samples was formed by rubbing NescoFilm™ against a 12-well template mould (GE Healthcare Life Sciences, UK). 1μl of each product was pipetted onto the NescoFilm™ wells and a 12-tooth sample applicator comb was used to simultaneously transfer approximately 0.3μl of each sample to the gel surface placed in the electrophoresis chamber of the PhastSystem™ separation and control unit (Pharmacia Biotech, UK). The running conditions of electrophoresis were accurately controlled in the PhastSystem™ by a microprocessor. Electrophoresis was achieved by use of solid buffer strips (PhastGel® Native buffer strips, GE Healthcare Life Sciences, UK) placed on either ends of the gel, in contact with both the gel surface and the electrodes. The gel was pre-run to generate a continuous buffer system before loading the samples. In order to find the optimal conditions for the detection of the amplified products by automated electrophoresis, the following parameters were changed: the temperature (4ºC, 10ºC, 15ºC or 20ºC), the duration of the pre-run step (9Vh or 49Vh) and the total run time of the separation (300Vh, 350Vh and 400Vh).

2.6.5.3 Staining of PhastGels®

After electrophoresis, the DNA bands were silver stained in the PhastSystem™ development unit chamber (Pharmacia Biotech, UK). This involved sixteen steps, which used nine solutions, following the manufacturers' instructions (Table 2.10). The staining solutions were freshly prepared just before starting the automated silver staining procedure. A hundred millilitres of each solution was required for each of the sixteen steps. The gels with significant results were scanned soon after completion of the silver staining as they may fade over a period of six to twelve months.

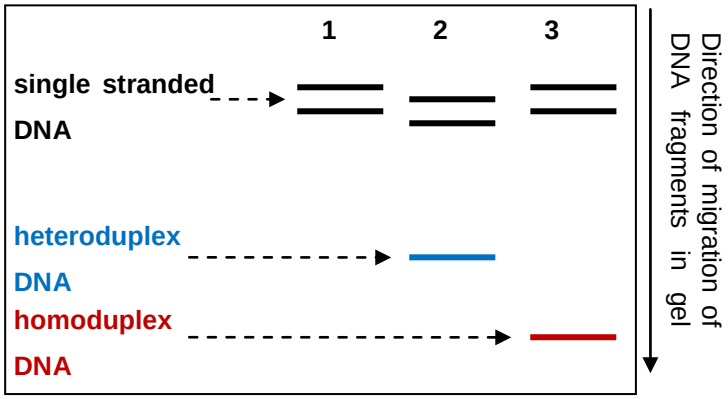
Table 2.10: Automated silver staining using the PhastSystem™ development unit
All reagents and salts used for the preparation of solutions were obtained from VWR International (UK)

Step	Solution	Temperature (ºC)	Time (min)	Purpose
1	Distilled deionised water	20ºC	0.5	Washing
2	50% (v/v) ethanol, 10% (v/v) acetic acid	50ºC	2	Fixation
3	10% (v/v) ethanol, 5% (v/v) acetic acid	50ºC	2	Fixation
4	10% (v/v) ethanol, 5% (v/v) acetic acid	50ºC	4	Fixation
5	10% glutaraldehyde	50ºC	6	Sensitisation
6	10% (v/v) ethanol, 5% (v/v) acetic acid	50ºC	2	Fixation
7	10% (v/v) ethanol, 5% (v/v) acetic acid	50ºC	5	Fixation
8	Distilled deionised water	50	2	Washing
9	Distilled deionised water	50	2	Washing
10	0.4% (w/v) silver nitrate	40	10	Staining
11	Distilled deionised water	30	0.5	Washing
12	Distilled deionised water	30	0.5	Washing
13	2.5% (w/v) sodium carbonate, 0.03% (v/v) formaldehyde	30	1	Developing
14	2.5% (w/v) sodium carbonate, 0.03% (v/v) formaldehyde	30	10	Developing
15	3.7% (w/v) Tris-HCl, 2.5% (w/v) sodium thiosulphate	30	2	Background reduction
16	10% (v/v) glycerol	50	5	Preservation

2.6.5.4 Analysis of SSCP gels

SSCP/heteroduplex analysis was used for the identification of single stranded, homoduplex and heteroduplex DNA molecules, which appeared as bands at different levels on the gel (Figure 2.7). The electrophoretic migration of samples in the gel was dependent on whether the DNA was single or double stranded, fragment size and DNA sequence.

Figure 2.7: Schematic diagram of SSCP gel showing the migration levels of single stranded (black bands), heteroduplex (blue band) and homoduplex (red band) DNA fragments



2.6.6 Sequencing analysis for the identification of SNP alleles and confirmation of mutation in PGD couples

Replicates of a PCR product were checked on an agarose gel and then pooled together to obtain 100µl of product which was filtered through a Microcon® centrifugal filter (Millipore, USA). The filter (in a 0.5ml tube) was centrifuged at 5000 rpm for 5 minutes (MSE Microcentaur, Sanyo, UK). 100µl of nuclease free water were added onto the filter and the tube was centrifuged again at 5000 rpm for 5 minutes. Another 100µl of nuclease free water were added and the centrifugation was repeated. 15µl of nuclease free water were added onto the filter and the filter was inverted (blue side down), placed in a new 0.5ml tube and centrifuged at 500rpm for 5 minutes. The last step was repeated and the flow-through was kept to set up the sequencing reaction using the Big Dye Terminator® v3.1 Cycle Sequencing kit (Applied Biosystems, UK).

Two reactions were set up on ice in 0.5ml tubes for each purified product (a forward and a reverse reaction). For each reaction, 8µl of the purified DNA were mixed with 4µl of cycle sequencing primer *BRCA1*c.3339T>G-F/R, rs1981929-F/R or MSH2_4SNP-F/R (Tables 2.5 and 2.6) at 2pmol/µl, 2µl big dye terminator cycle sequencing mix, 3µl big dye terminator cycle sequencing buffer and 3µl nuclease free water. The tube was placed on a thermal cycler (GeneAmp 9700, Applied Biosystems, UK) set to perform 25 cycles at 96ºC for 10 seconds, 50ºC for 5 seconds and 60ºC for 4 minutes. The product was purified by adding 2µl of EDTA (125mM), 2µl of sodium acetate pH4.6 (3M) and 50µl ethanol (96-100%) and mixing with a vortex (Autovortex SA6, Stuart Scientific, UK).

The sample was left at room temperature for 15 minutes then centrifuged at 13,000rpm for 25 minutes. The supernatant was discarded and the pellet was left to dry on a heating block at 37ºC for 5 minutes. 60µl of 70% ethanol were then added to the pellet, mixed on the vortex and centrifuged at 13,00rpm for 5 minutes. Once again, the supernatant was discarded and the pellet was left to dry at 37ºC for 5 minutes.

Ten microlitres of HiDi formamide were added directly to the pellet, mixed by pipetting up and down and transferred to a 96-well plate for loading on the ABI Prism™ 3100 genetic analyzer (using dye set Z and 3100-Avant Genetic 3100 BigDye® Terminator v3.1 Matrix Standard, Applied biosystems, UK). The data was analysed using the GeneMapper™ sequencing software (Applied Biosystems, UK). Loading of samples on the ABI Prism™ 3100 for sequencing was carried out by Seema Dhanjal.

2.6.7 Mini-sequencing for SNP analysis for the *MSH2* PGD protocol

Mini-sequencing reactions were carried out on amplified DNA at the selected region (PCR products for the rs1981929 SNP) using the SNaPshot™ multiplex kit (Applied Biosystems, UK) in order to identify the bases at the SNP site (A/T and/or G/C).

Samples were prepared by adding 5µl SAP (shrimp alkaline phosphatase) and 0.1µl ExoI (New England Biolabs, UK) to 15µl of PCR product. The mixture was incubated at 37ºC for one hour, 75ºC for 15 minutes and kept at 4ºC until used in the SNaPshot reaction.

Each SNaPshot™ reaction mix was set up on ice in a 0.2ml MicrAmp reaction tube and consisted of the following reagents: 5µl SNaPshot™ multiplex ready reaction mix, 3µl substrate, 1µl mini-sequencing primer or pooled forward and reverse primers (Table 2.11) at 10µM and 1µl nuclease free water (Promega, UK). The tube was loaded on the thermal cycler (GeneAmp 9700, Applied Biosystems, UK) set to perform 25 cycles at 96ºC for 10 seconds, 50ºC for 5 seconds and 60ºC for 30 seconds. This was followed by a post-extension treatment. One unit (1µl) of shrimp alkaline phosphatase (SAP) (New England BioLabs, UK) was added to the reaction and mixed thoroughly. The tube was incubated at 37ºC for one hour, 75ºC for 15 minutes (enzyme deactivation step) and cooled to 4ºC (GeneAmp 9700, Applied Biosystems, UK). The SNaPshot™ products were kept at -20ºC (for long term storage).

Table 2.11: Sequence, chromosome location and modifications of primers used for mini-sequencing experiments

Primer	Primer sequence	Locus	Product size
ForA/G (F)	5'-TAAATAGTAACTTTGGAGACCT-3'	2p22-p21	23 bases
RevT/C (R)	5'-GGTACAAATAGTACAG-3'	2p22-p21	17 bases

For: Forward primer; Rev: Reverse primer
These primers flank the rs1981929 SNP site located upstream of exon 8 in the *MSH2* gene.

The mini-sequencing analysis was completed with electrophoresis of the SNaPshot™ product on the ABI Prism™ 310 genetic analyzer (using the GS STR POP-4 E5 module and E5 matrix settings). Samples were prepared by mixing 0.5µl SNaPshot™ product with 0.5µl GeneScan®-120 LIZ and 12µl HiDi™ formamide (Applied Biosystems, UK). Base calling was done according to the colour and size of the fragments visualised after GeneScan™ analysis.

3 RESULTS

3.1 Expression analysis of DNA repair genes in human oocytes and embryos using microarrays

3.1.1 RNA Extraction

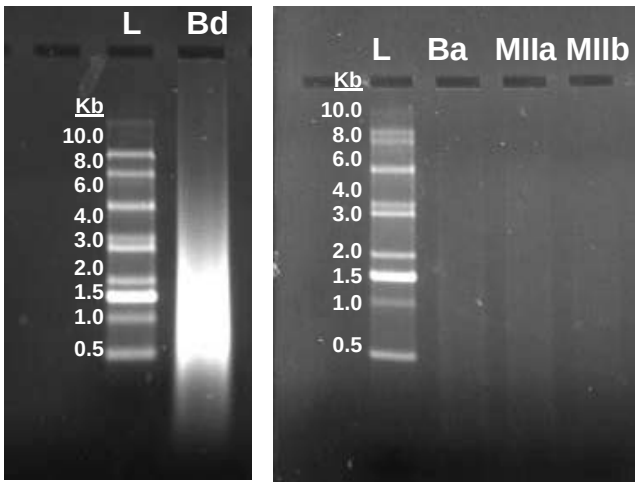
3.1.1.1 Samples collected and processed for microarray analysis

Total RNA was extracted from five sets of three oocytes (labelled MIIa-MIIe) and four sets of three blastocysts (labelled Ba-Bd) (section 2.3.1 in methods). The presence of a polar body was confirmed in all 15 oocytes collected for this study during tubing, indicating that all immature oocytes had matured into MII oocytes during incubation in the standard culture media, prior to tubing. The mean age of the oocyte donors was 35.77 years +/- 4.05. The average age per set ranged between 34 and 37 years (standard deviation = 1.36).

3.1.1.2 Quality and concentrations of RNA extracted from MII oocytes and blastocysts

Agarose gel electrophoresis and NanoDrop® analysis (section 2.5) were only used for the assessment of amplified RNA samples (section 2.4.1) as only ~2µl of the original total RNA could be spared; this was used on the Agilent 2100 Bioanalyzer, which is the most informative and sensitive of the three methods. Agarose gel electrophoresis analysis allowed a quick assessment of RNA quantity and quality (Figure 3.1).

Figure 3.1: Assessment of amplified human MII oocyte and blastocyst RNA samples on 1.5% agarose gels



The above gels show the RNA samples of highest quality, Bd (left image), and poorest quality, Ba, MIIa and MIIb (right image), which were not used for microarray analysis. The intensity of the smear was indicative of the quantity of amplified RNA.

The 18S and 28S rRNA are ~ 2 kilobases (Kb) and 5 Kb in size, respectively. Poly(A) selected and total RNA samples appear as a smear from ~ 0.5 to 6 Kb (resulting from the population of mRNAs) with the most intense area between 1.5 and 2 Kb. This smear was observed in the amplified samples with high concentrations, e.g. sample Bd (Figure 3.1). The three samples with the lowest RNA concentrations, Ba, MIIa and MIIb, were not used for microarray analysis.

Assessment of non-amplified RNA samples on the Agilent 2100 Bioanalyzer showed the 18S and 28S RNAs, which were clearly detected between 40-50 seconds of electrophoresis, corresponding to sizes just below 2 Kb and just above 4 Kb (Figure 3.2a). The amplified RNA samples showed one broad peak covering a wider size range (~0.2-3 Kb) rather than two sharp peaks (Figure 3.2b).

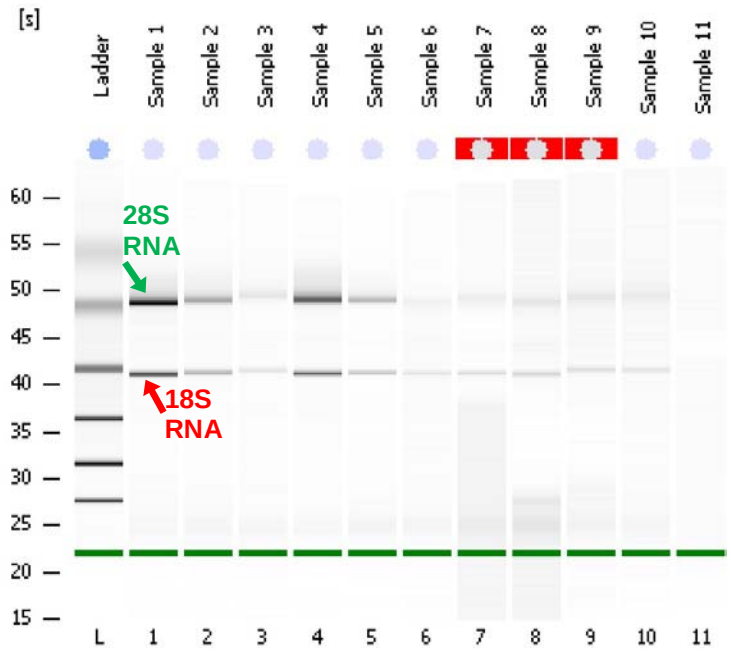
All samples selected to be used in this project were of good quality according to our agarose gel electrophoresis, NanoDrop® and Agilent 2100 Bioanalyzer analysis results. The samples with the highest concentrations were hybridized on the microarrays (Table 3.1).

The amount of total RNA extracted from each set of three blastocysts or MII oocytes ranged from 63 pg/µl to 393pg/µl and from 48pg/µl to 78pg/µl, respectively (Table 3.1). This was equivalent to a yield of ~ 250-1600pg of RNA per blastocyst and ~ 200-320pg per oocyte. The concentrations increased after two rounds of amplification to the following ranges: 91ng/µl to ~1µg/µl for the blastocyst sets and 134ng/µl to 292ng/µl for the oocyte sets. The best three RNA samples from each group were used for microarray analysis (samples Bb, Bc, Bd, MIIc, MIId and MIIe).

Figure 3.2: Assessment of RNA samples from human MII oocytes and blastocysts pre (a) and post (b) amplification using the Agilent 2100 Bioanalyzer

a) Pico chip electrophoresis results showing the non-amplified RNA samples represented as bands (i) and peaks in the individual electropherograms of every tested sample (ii)

(i)



(ii)

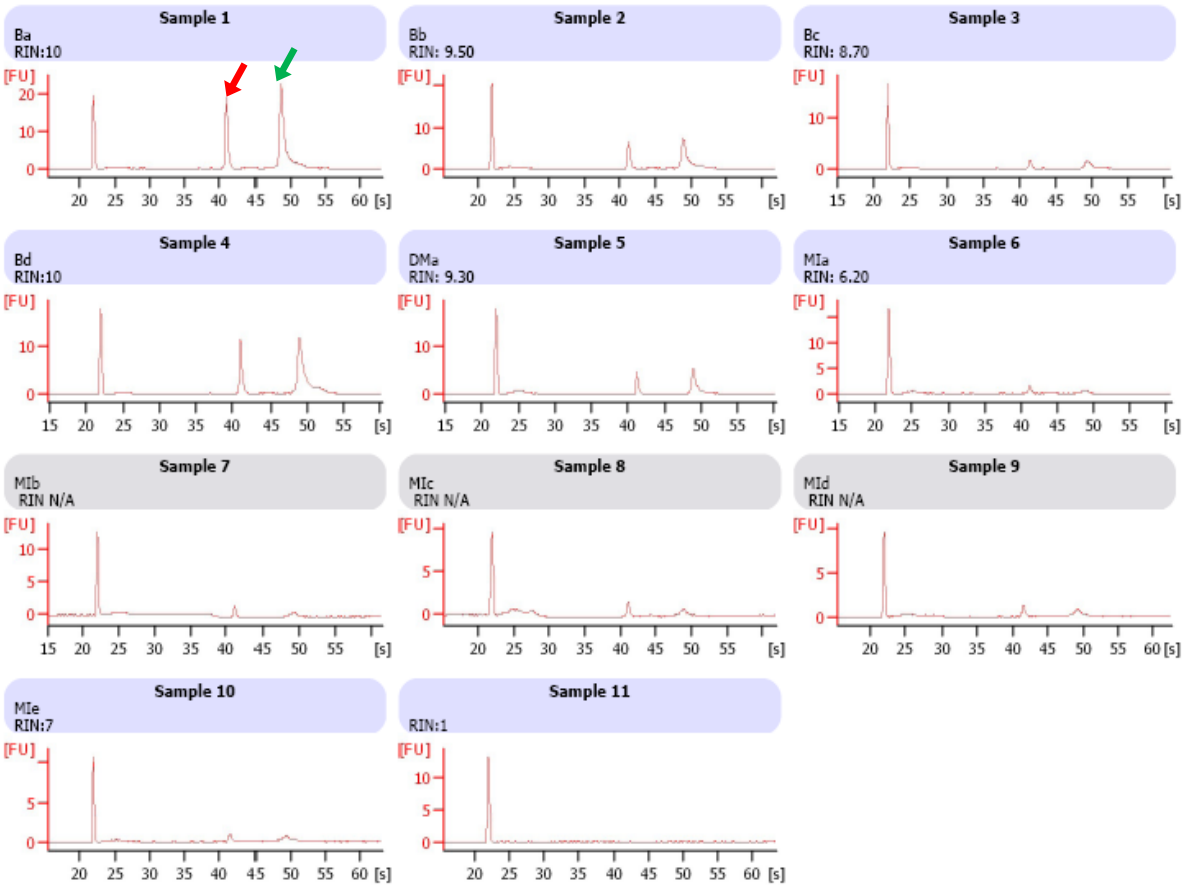
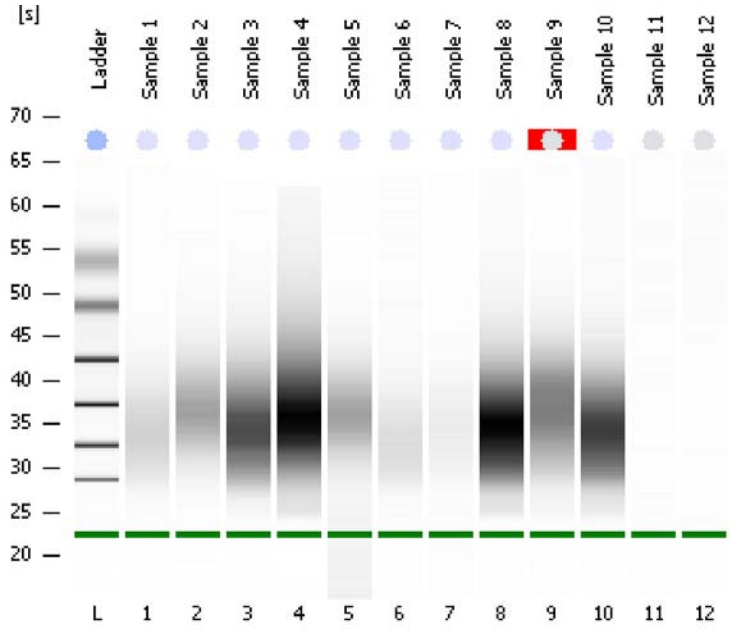


Figure 3.2 (continued): Assessment of RNA samples using the Agilent 2100 Bioanalyzer

b) Nano chip electrophoresis run showing the amplified RNA samples represented as bands (i) and peaks in the individual electropherograms of every tested sample (ii) or overlaid electropherograms of the triplicate samples sets selected for hybridisation on microarrays (iii)

(i)



(ii)

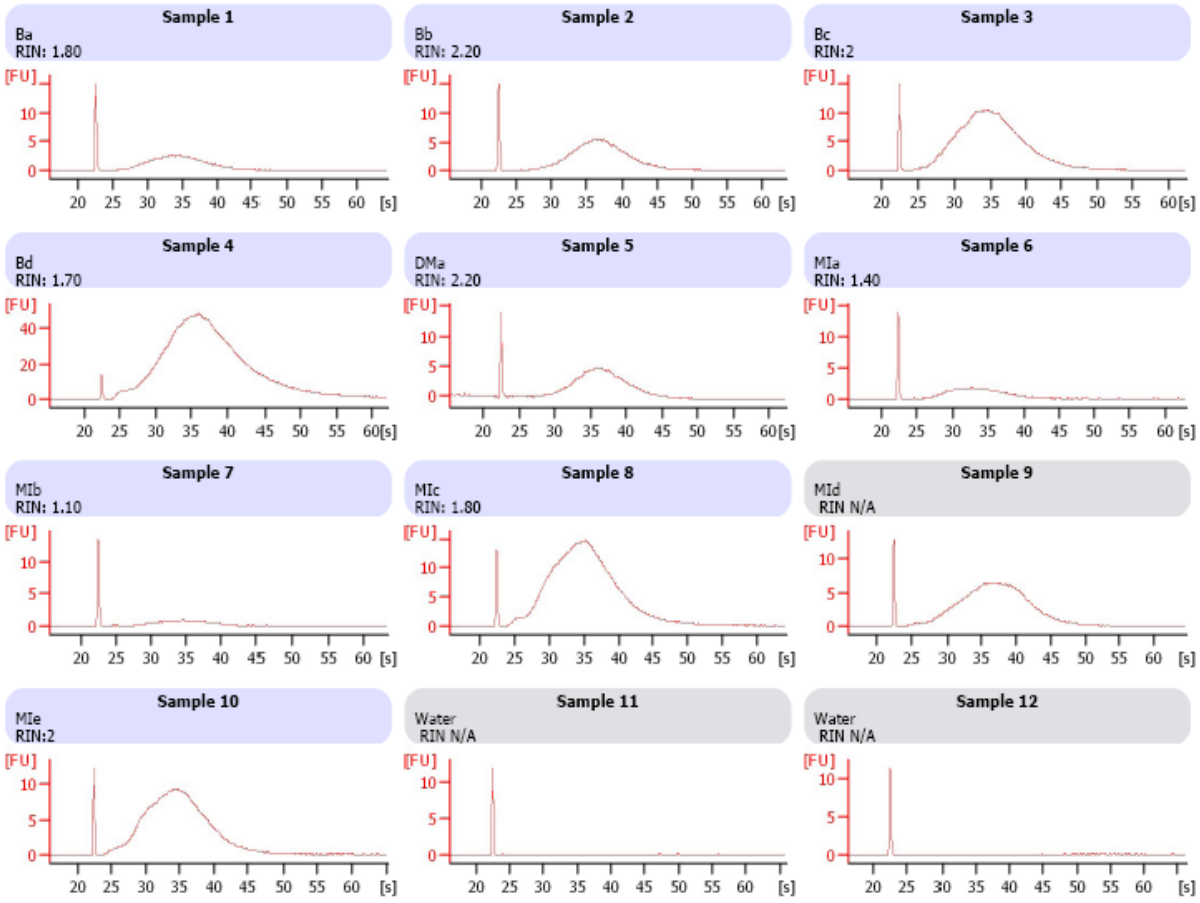
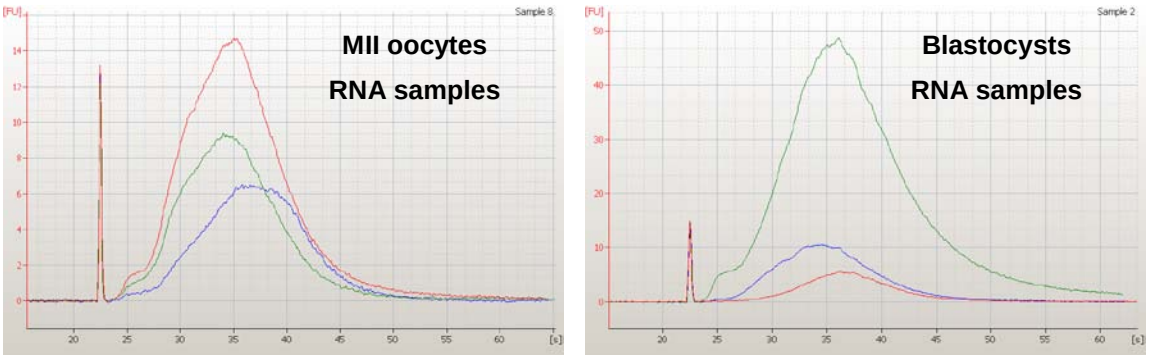


Figure 3.2 (continued): Assessment of RNA samples using the Agilent 2100 Bioanalyzer

(iii)



Samples were loaded on the chips in the following order (1-10): Ba, Bb, Bc, Bd, DMa (not used for this project), MIIa, MIIb, MIIc, MIId and MIIe. Samples 11 and 12 were blanks (water). The ladder peaks were: 25, 200, 500, 1000, 2000 and 4000 bases.

The 18S (red arrows) and 28S (green arrows) RNAs were clearly detected between 40-50 seconds of the electrophoresis run (corresponding to sizes just below 2Kb and just above 4Kb) in the non-amplified samples (a). The amplified samples showed one broad peak or smear covering a wider size range (~0.2-3 Kb) instead of the two sharp peaks (b).

(iii) The two electropherograms overlay the samples that were selected for hybridisation on *AB*'s Human Genome Survey Microarrays. The selected MII oocyte RNA samples were: c (red), d (blue) and e (green). The selected blastocyst RNA samples were: b (red), c (blue) and d (green).

Table 3.1: Assessment of RNA samples concentrations for microarray analysis

Sample	Bioanalyzer readings Pre-amplification (pg/µl)	Bioanalyzer readings Post-amplification (ng/µl)	NanoDrop® readings Post-amplification (ng/µl)	Visualisation on 1.5% agarose gels Post-amplification
Ba	393	51	59.46	✓ very faint
Bb	191	91	112.5	✓
Bc	63	207	232.83	✓
Bd	246	1027	1115.3	✓
MIIa	78	36	41.76	✓ very faint
MIIb	65	18	22.89	✓ very faint
MIIc	48	292	296.12	✓
MIId	50	134	130.3	✓
MIIe	52	184	206.48	✓

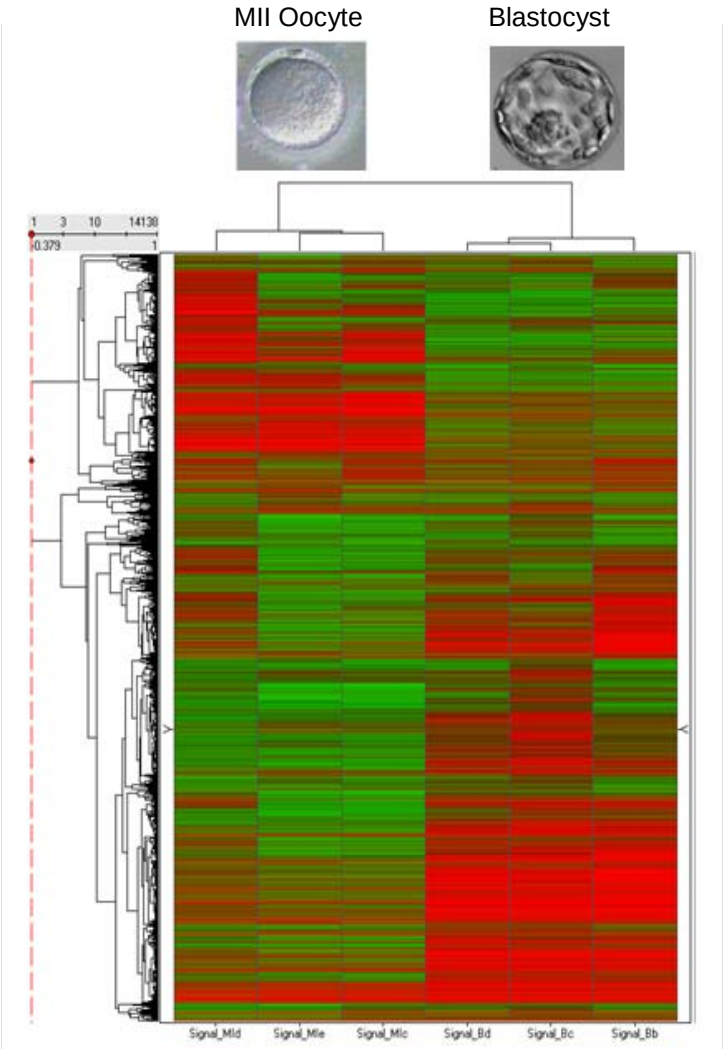
✓ : indicates the RNA sample was visualised on agarose gels. Samples Ba-Bd designate RNA extracted from blastocysts and MIIa-MIIe designate RNA exracted from MII oocytes.
All samples shown in grey were not considered of suitable concentration to be used for microarray analysis. The samples used were: Bb, Bc, Bd, MIIc, MIId and MIIe.

3.1.2 Microarray Results

3.1.2.1 General expression analysis

All six microarrays were hybridized and scanned successfully (section 2.6.1.1). Readings of all genes expressed across all samples were visualised together using hierarchical clustering, which grouped the three MII oocytes together and the three blastocysts together based on their distinct expression profiles (Figure 3.3).

Figure 3.3: Hierarchical clustering displaying the expression signatures of human MII oocytes (left) and blastocysts (right)



Red indicates an expression above median, green indicates an expression below median and black represents median expression**.** Hierarchical clustering grouped the three MII oocytes together and the three blastocysts together based on their expression profiles.

Out of the 32,878 oligo probes investigated, the total number detected was 13,118 and 11,734 in the blastocysts and oocytes respectively. None of the genes previously shown to be overexpressed in cumulus cells (Assou *et al.*, 2006; Hernandez-Gonzalez *et al.*, 2006) could be detected in any sample except for *BTC,* which was picked up at very low levels (near the limit of detection) in both oocytes and blastocysts, and *BMPR2,* which had moderate expression level in oocytes and low levels in blastocysts (Table 3.2).

Eleven out of the twelve housekeeping genes investigated were detected in both oocytes and blastocysts; *BMP7* was not detected in either sample sets. Three genes (*UBC, EEF1E1* and *TUBB4)* did not show significant differences in expression levels; however, five genes (*PPIA*, *HIST2H2AA*, *ACTB*, *GAPDH*, and *H2AFZ*) showed a greater than three fold increase in the blastocyst group compared to the oocyte group and three (*TBP, POLR2A* and *HPRT1*) showed a greater than three fold decrease in the blastocyst group versus the oocyte group (Table 3.3). Regarding genes associated with pluripotency, *NANOG* and *POU5F1* (*OCT3/4*) mRNAs were detected in high levels in the blastocysts. Only *POU5F1* was detected in MII oocytes with low expression levels. *SOX2* was not detected in either sample sets.

Global gene expression, across all chromosomes, was visualised by plotting the number of probes that detected high, medium and low levels (section 2.6.1.2) of mRNA expression on each individual chromosome in a bar chart (Figure 3.4). This allowed further assessment of the hybridisation of the samples on the microarrays, as gene rich (e.g. chromosomes 1, 2 and 19) and gene poor chromosomes (e.g. chromosomes 13, 21 and Y) showed high and low numbers of expressed genes, respectively. This analysis also showed that a greater number of genes had high expression levels in blastocysts compared to oocytes, as observed from the hierarchical cluster (Figure 3.3).

Table 3.2: Genes known to be overexpressed in cumulus cells and their expression levels in the MII oocyte and blastocyst

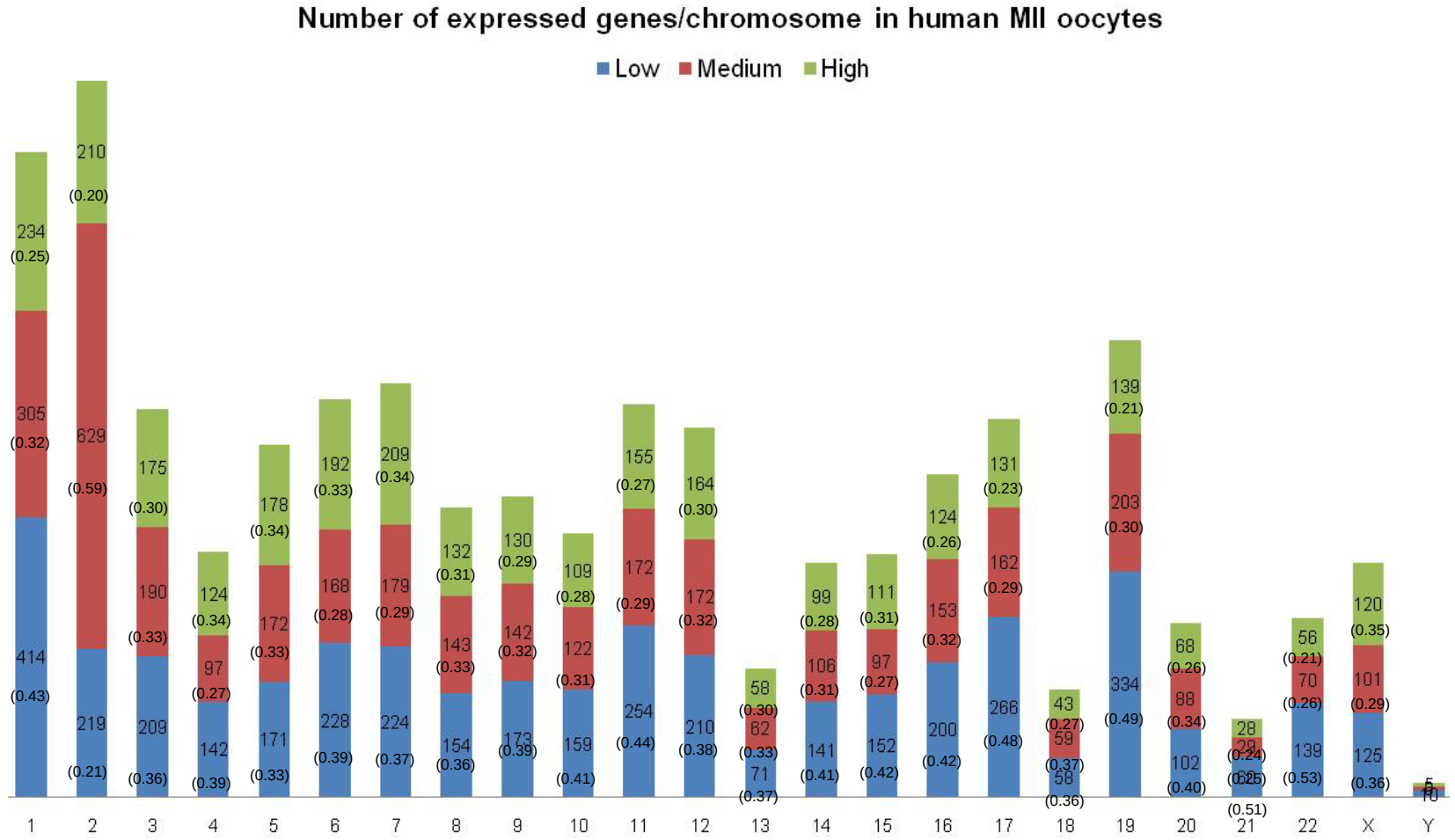
Gene symbol	MII oocytes expression level	Blastocysts expression level	Gene name	Cytoband	Reference sequence
BTC	Low	Low	Betacellulin	4q13-q21	NM_001729
AREG	Not detected	Not detected	amphiregulin (schwannoma-derived growth factor)	4q13-q21	NM_001657
EREG	Not detected	Not detected	Epiregulin	4q13.3	NM_001432
TNFSF11	Not detected	Not detected	tumor necrosis factor (ligand) superfamily, member 11	13q14	NM_033012
SEMA3A	Not detected	Not detected	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3A	7p12.1	NM_006080
C7	Not detected	Not detected	complement component 7	5p13	NM_000587
CD200	Not detected	Not detected	CD200 antigen	3q12-q13	NM_001004197
BMPR2	Medium	Low	bone morphogenetic protein receptor, type II (serine/threonine kinase)	2q33-q34	NM_001204
LHCGR	Not detected	Not detected	luteinizing hormone/choriogonadotropin receptor	2p21	NM_000233

Table 3.3: Housekeeping genes showing differential expression in the blastocyst versus the oocyte group ($p<0.05$) with their corresponding fold changes. *UBC, EEF1E1* and *TUBB4* (not shown here) did not show significant differences in expression levels ($p>0.05$) and *BMP7* was not detected in either sample sets.

Gene symbol	Fold change	P.value	Gene name	Cytoband	Reference sequence
H2AFZ	180.143	0.000	H2A histone family, member Z	4q24	NM_002106
GAPDH	17.582	0.001	glyceraldehyde-3-phosphate dehydrogenase	12p13	NM_001001303
ACTB	12.924	0.001	actin, beta	7p15-p12	NM_001101
HIST2H2AA (HIST2H2AC)	4.624	0.041	histone 2, H2aa|histone 2, H2ac	1q21-q23	NM_003516
PPIA	3.683	0.010	peptidylprolyl isomerase A (cyclophilin A)	7p13-p11.2	NM_008907
TBP	0.294	0.020	TATA box binding protein	6q27	NM_003194
POLR2A	0.291	0.031	polymerase (RNA) II (DNA directed) polypeptide A, 220kDa	17p13.1	NM_000937
HPRT1	0.265	0.008	hypoxanthine phosphoribosyltransferase 1 (Lesch-Nyhan syndrome)	Xq26.1	NM_000194

Figure 3.4: Bar chart displaying the levels of gene expression on each chromosome in human MII oocytes (a) and blastocysts (b)

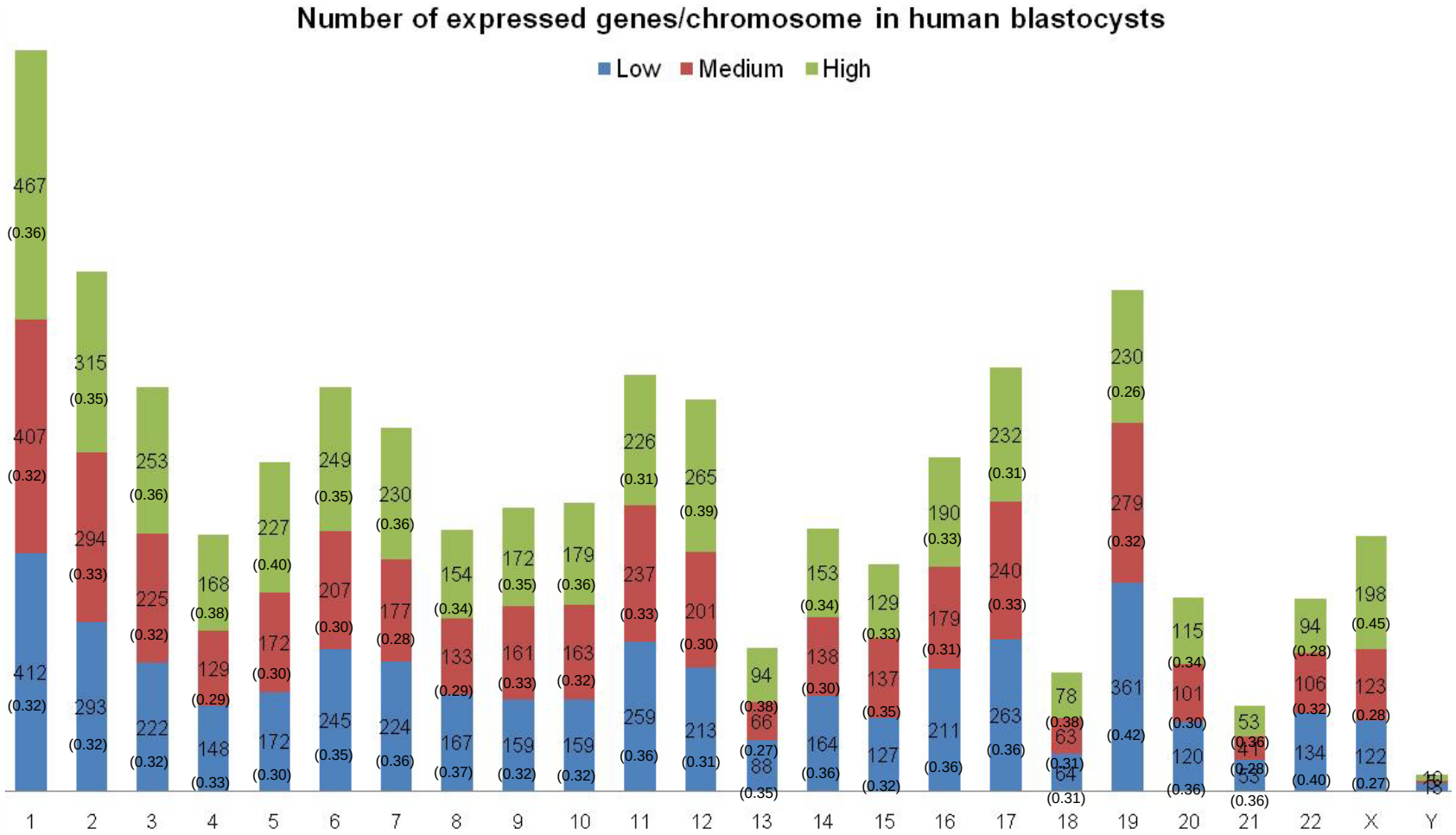
(a)



The numbers indicate the total number of probes representing genes expressed at low (blue), medium (red) or high (green) levels for each chromosome. The numbers in brackets show the proportion of genes expressed at low, medium or high levels out of the total number of genes on that chromosome.

Figure 3.4 (continued): Bar chart displaying the levels of gene expression on each chromosome in human MII oocytes (a) and blastocysts (b)

(b)



The numbers indicate the total number of probes representing genes expressed at low (blue), medium (red) or high (green) levels for each chromosome. The numbers in brackets show the proportion of genes expressed at low, medium or high levels out of the total number of genes on that chromosome.

3.1.2.2 Expression of DNA repair genes in MII oocytes and blastocysts

Out of the 154 DNA repair genes (Wood *et al.*, 2005) investigated, 153 were identified by probes on the *Applied Biosystems* Human Genome Survey Microarrays V2.0. mRNAs coding for 109 of those genes were detected in blastocysts (the remaining 43 had S.N. ratios < 3 in at least one of the replicates) and 107 of those genes were detected in oocytes (Tables 3.6-3.10). Eighty nine of these DNA repair genes were common to both the oocyte and blastocyst groups. No DNA repair gene probes were excluded when the top and bottom 5% signals were eliminated from the data.

Distinctive expression patterns were observed in the MII oocyte and blastocyst groups. Overall, 129 DNA repair genes were included in a t-test analysis and 55 were found to be differentially expressed in blastocysts compared to oocytes with a fold change > 3 ($p<0.05$). Amongst the differentially expressed DNA repair genes, 40 genes (73%) had significantly lower expression levels in the blastocyst group versus the oocyte group, whereas only 15 (27%) had significantly higher expression levels. All differentially expressed DNA repair genes are listed with their corresponding fold changes in Tables 3.4 and 3.5.

The mRNA expression levels of genes involved in the main repair pathways, relative to the structure specific flap endonuclease-1 (*FEN1*) gene, are illustrated in Figures 3.5-3.10. *FEN1* was selected as the reference gene as it was highly expressed in both the oocyte and blastocyst groups with similar signal levels. This allowed the analysis of the relative differences in gene expression within each group and between the two groups.

Table 3.4: Genes with significantly lower expression (p<0.05) in the blastocyst group versus the oocyte group (from greatest to lowest difference)

Gene symbol	Fold change	P.value	DNA repair pathway	Cytoband	Reference sequence
LIG1	3.01E-02	1.50E-04	NER	19q13.2-q13.3	NM_000234
DCLRE1A	3.48E-02	1.40E-04	Genes with suspected DNA repair function	10q25.1	NM_014881
ALKBH2 (ABH2)	4.10E-02	2.23E-04	Direct reversal of damage	12q24.11	NM_001001655
REV1L	4.21E-02	1.37E-04	DNA Polymerases	2q11.1-q11.2	NM_016316
FANCM	4.34E-02	1.07E-03	Genes defective in diseases associated with sensitivity to DNA damaging agents	14q21.3	NM_020937
RPA4	4.81E-02	2.59E-03	Genes with suspected DNA repair function	Xq21.33	NM_013347
FANCL	5.26E-02	9.30E-04	Genes defective in diseases associated with sensitivity to DNA damaging agents	2p16.1	NM_018062
BRCA2	5.68E-02	6.40E-03	DSBR - HR	13q12.3	NM_000059
TDP1	6.46E-02	7.85E-04	Repair of DNA-protein crosslinks	14q32.11	NM_018319
POLB	7.33E-02	2.61E-04	DNA Polymerases	8p11.2	NM_002690
CDK7	7.73E-02	5.86E-04	NER	5q12.1	NM_001799
FANCN (FLJ21816)	8.66E-02	4.45E-03	Genes defective in diseases associated with sensitivity to DNA damaging agents	16p12.1	NM_024675
RAD50	9.04E-02	3.88E-04	DSBR - HR	5q31	NM_005732
RPA1	9.31E-02	9.89E-03	NER	17p13.3	NM_002945
MSH5	9.60E-02	2.98E-03	MMR	6p21.3	NM_002441
RBBP8	1.03E-01	3.89E-03	DSBR - HR	18q11.2	NM_002894
MBD4	1.14E-01	1.36E-03	BER	3q21-q22	NM_003925
RAD54B	1.15E-01	4.28E-03	DSBR - HR	8q22.1	NM_006550
POLI	1.16E-01	5.94E-03	DNA Polymerases	18q21.1	NM_007195
DEPC-1	1.30E-01	4.28E-03	Direct reversal of damage	11p11.2	NM_139178
OGG1	1.33E-01	1.12E-03	BER	3p26.2	NM_016821
ATR	1.36E-01	3.86E-02	Other conserved DNA damage response genes	3q22-q24	NM_001184
MGMT	1.46E-01	3.59E-03	Direct reversal of damage	10q26	NM_002412
WRN	1.56E-01	1.66E-02	Genes defective in diseases associated with sensitivity to DNA damaging agents	8p12-p11.2	NM_000553
PCNA	1.59E-01	1.62E-03	DNA Polymerases	20pter-p12	NM_002592
KIAA1794	1.63E-01	1.41E-02	Genes defective in diseases associated with sensitivity to DNA damaging agents	15q25-q26	NM_018193
UNG	1.78E-01	3.30E-03	BER	12q23-q24.1	NM_080911
MSH3	1.78E-01	1.92E-02	MMR	5q11-q12	NM_002439
CCNH	1.86E-01	3.20E-02	NER	5q13.3-q14	NM_001239
POLQ	1.93E-01	3.31E-03	DNA Polymerases	3q13.33	NM_199420

Table 3.4 (continued): Genes with significantly lower expression (p<0.05) in the blastocyst group versus the oocyte group (from greatest to lowest difference)

Gene symbol	Fold change	P.value	DNA repair pathway	Cytoband	Reference sequence
PMS1	1.95E-01	5.02E-03	MMR	2q31-q33|2q31.1	NM_000534
RAD17	2.02E-01	4.40E-03	Other conserved DNA damage response genes	5q13	NM_002873
MSH2	2.19E-01	5.04E-03	MMR	2p22-p21	NM_000251
ERCC6	2.28E-01	2.21E-02	NER	10q11.23	NM_000124
MPG	2.47E-01	1.64E-02	BER	16p13.3	NM_002434
MNAT1	2.60E-01	4.99E-02	NER	14q23	NM_002431
FAAP24 (C19orf40)	2.73E-01	4.61E-02	Genes defective in diseases associated with sensitivity to DNA damaging agents	19q13.11	NM_152266
CHEK1	2.74E-01	1.20E-02	Other conserved DNA damage response genes	11q24-q24	NM_001274
UBE2A	2.76E-01	9.64E-03	Rad6 Pathway	Xq24-q25	NM_003336
RAD9A	2.98E-01	1.76E-02	Other conserved DNA damage response genes	11q13.1-q13.2	NM_004584
DCLRE1B	3.69E-01	2.73E-02	Genes with suspected DNA repair function	1p13.2	NM_022836

Table 3.5: Genes with significantly higher expression (p<0.05) in the blastocyst group versus the oocyte group (from greatest to lowest difference)

Gene symbol	Fold change	P.value	DNA repair pathway	Cytoband	Reference sequence
XRCC5 (Ku70)	36.946	1.90E-03	NHEJ	2q35	NM_021141
SMUG1	15.840	1.04E-02	BER	12q13.11-q13.3	NM_014311
TDG	12.848	9.45E-03	BER	12q24.1	NM_003211
RAD51L3	11.441	1.39E-02	HR	17q11	NM_002878
SHFM1	11.156	1.11E-03	HR	7q21.3-q22.1	NM_006304
RAD23B	9.150	6.87E-03	NER	9q31.2	NM_002874
MDC1	9.124	8.07E-04	Other conserved DNA damage response genes	6pter-p21.31	NM_014641
FLJ35220	6.418	6.57E-03	Editing & processing nucleases	17q25.3	NM_173627
APTX	6.309	6.33E-03	Genes with suspected DNA repair function	9p13.3	NM_175073
FANCG	4.971	1.60E-02	Genes defective in diseases associated with sensitivity to DNA damaging agents	9p13	NM_004629
FANCD2	4.314	4.42E-02	Genes defective in diseases associated with sensitivity to DNA damaging agents	3p26	NM_033084
XRCC1	3.493	2.83E-02	BER	19q13.2	NM_006297
NUDT1	3.201	2.54E-02	Modulation of nucleotide pools	7p22	NM_002452
UBE2V2	3.045	2.12E-02	Rad6 Pathway	8q11.21	NM_003350
POLD1	2.765	3.53E-02	DNA Polymerases	19q13.3	NM_002691

All DNA repair pathways were represented in human oocytes and blastocysts. However, most differentially expressed DNA repair genes were detected in lower levels in the blastocyst compared to the oocyte.

3.1.2.2.1 Base excision repair (BER)

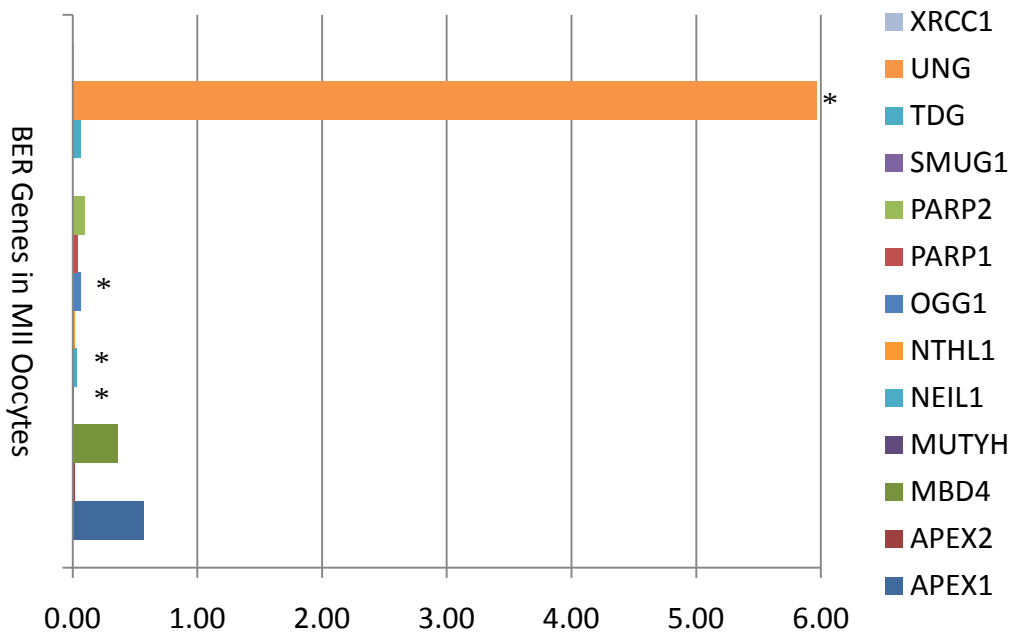
Most of the BER genes examined were expressed in both MII oocytes and blastocysts (Table 3.6). *UNG*, *APEX1* and *POLB* mRNAs were detected at high levels in both human oocytes and blastocysts and *OGG1* mRNA was detected at medium levels in the MII oocytes and low levels in the blastocysts. Figure 3.5 displays the different expression patterns of BER genes (relative to *FEN1)* in human MII oocytes and blastocysts.

Furthermore, the mRNA expression of four enzymes involved in the protection against free radicals (Cu-Zn-superoxide dismutase (*Cu-Zn-SOD* or *SOD1*), Mn-superoxide dismutase (*Mn-SOD* or *SOD2*), glutathione peroxidase (*GPX*) and γ-glutamylcysteine synthetase (*GCS*)) was investigated in our data. *SOD1*, *SOD2* and *GPX* were highly expressed in human MII oocytes and blastocysts; *GCS*, however, was not detected in MII oocytes and had low expression levels in blastocysts.

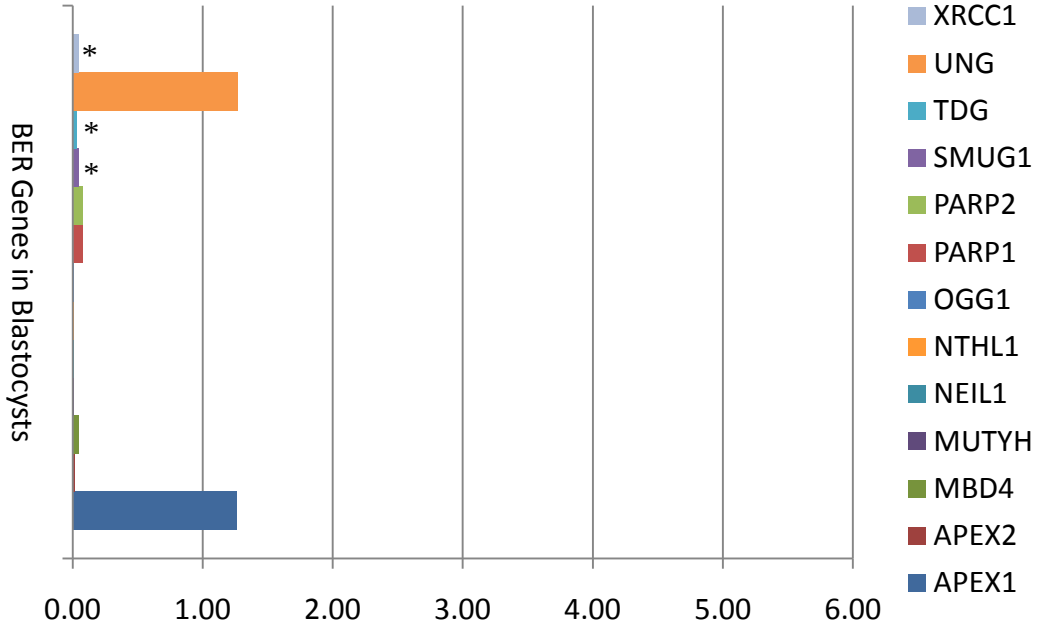
Table 3.6: Expression levels of mRNAs coding for BER genes in human MII oocytes and blastocysts

Gene symbol	MII oocytes signal level	Blastocysts signal level	Cytoband	Reference sequence
APEX1	High	High	14q11.2-q12	NM_001641
APEX2	Low	Low	Xp11.22	NM_014481
LIG3	Not detected	Not detected	17q11.2-q12	NM_013975
MBD4	High	Medium	3q21-q22	NM_003925
MPG	Medium	Not detected	16p13.3	NM_002434
MUTYH	Low	Low	1p34.3-p32.1	NM_012222
NEIL1	Medium	Low	15q23	NM_024608
NEIL2	Not Detected	Not detected	8p23.1	NM_145043
NTHL1	Low	Low	16p13.3	NM_002528
OGG1	Medium	Low	3p26.2	NM_016821
PARP1	Medium	High	1q41-q42	NM_001618
PARP2	Medium	High	14q11.2-q12	NM_005484
PNKP	Not detected	Not detected	19q13.3-q13.4	NM_007254
SMUG1	Not detected	Medium	12q13.11-q13.3	NM_014311
TDG	Not detected	High	12q24.1	NM_003211
UNG	High	High	12q23-q24.1	NM_080911
XRCC1	Low	Medium	19q13.2	NM_006297

Figure 3.5: BER genes expressed in human MII oocytes (top) and blastocysts (bottom) with mRNA signal levels relative to *FEN1*



The asterisk denotes the genes that were expressed at significantly higher levels in MII oocytes compared to blastocysts ($p<0.05$).



The asterisk denotes the genes that were expressed at significantly higher levels in blastocysts compared to MII oocytes ($p<0.05$).

3.1.2.2.2 Double Strand Break Repair (DSBR)

Overall, mRNA templates coding for 12/19 genes involved in HR repair and 3/6 genes involved in NHEJ were detected in both the MII oocyte and blastocyst groups (Table 3.7).

MRE11A and *EXO1* mRNAs were detected at medium levels in human MII oocytes and blastocysts and *RAD50* mRNA was detected at high levels in MII oocytes but low levels in blastocysts ($p<0.05$, 11 fold increase in oocytes). *NBS1* mRNA was detected at low levels in MII oocytes but was not detected in the blastocyst group (low signal detected for only two out of three replicates). *RAD51* and *RAD52* were expressed at high levels in both groups (Table 3.7). mRNA expression of *BRCA2* was only detected in the MII oocytes (medium level). Figure 3.6 displays the different expression patterns of HR repair genes (relative to *FEN1)* in human MII oocytes and blastocysts.

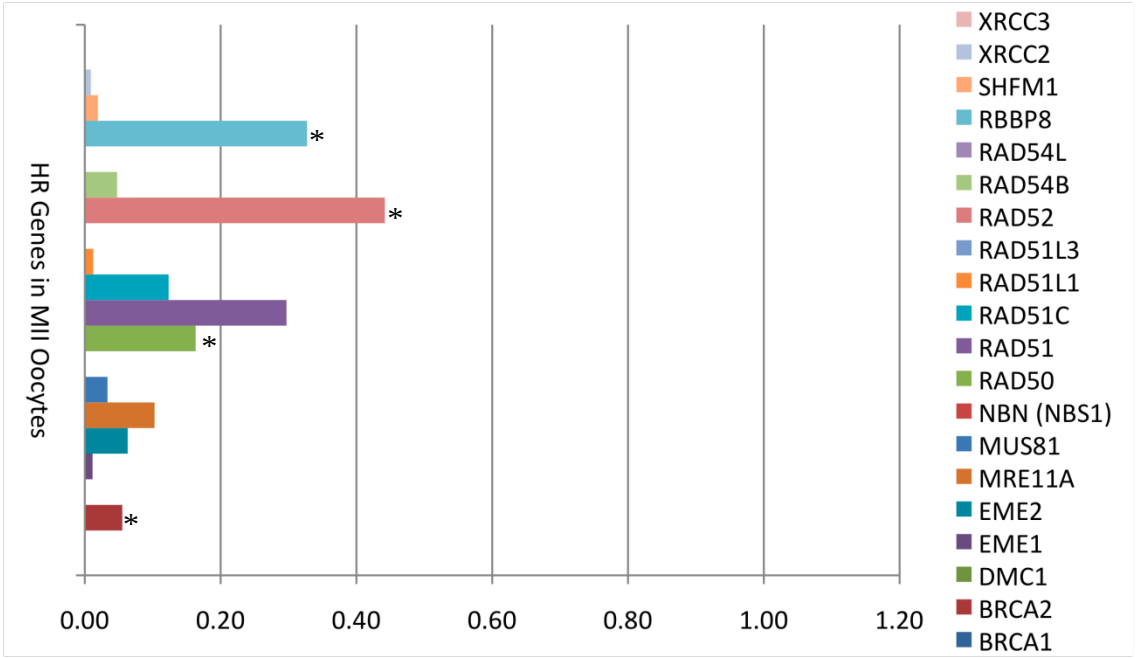
Among the genes that influence the choice of pathway for the repair of DSBs, *RBBP8* (or *CtIP*) was expressed at high and medium levels in MII oocytes and blastocysts, respectively, *FANCD2* expression levels were significantly higher in the blastocyst versus the oocyte group ($p<0.05$ and 4.3 fold increase) (Table 3.5) and *BRCA1* mRNA expression was detected at medium levels in the blastocysts but was not detected in MII oocytes. *ATR* and *CHEK1* had medium expression levels in both groups with significantly higher levels in MII oocytes versus blastocysts (Table 3.4). *53BP1* mRNAs were detected at high and medium levels in MII oocytes and blastocysts, respectively. *MDC1* mRNAs were detected at low and medium levels in MII oocytes and blastocysts, respectively. *CHEK2* and *ATM* mRNAs were not detected in either sample groups. *PARP1* showed higher mRNA expression levels in blastocyts versus MII oocytes (but the difference was not statistically significant) and *RAD18* mRNA was only detected in the blastocysts.

Table 3.7: Expression levels of mRNAs coding for DSBR genes via homologous recombination (HR) and non-homologous end joining (NHEJ) in human MII oocytes and blastocysts

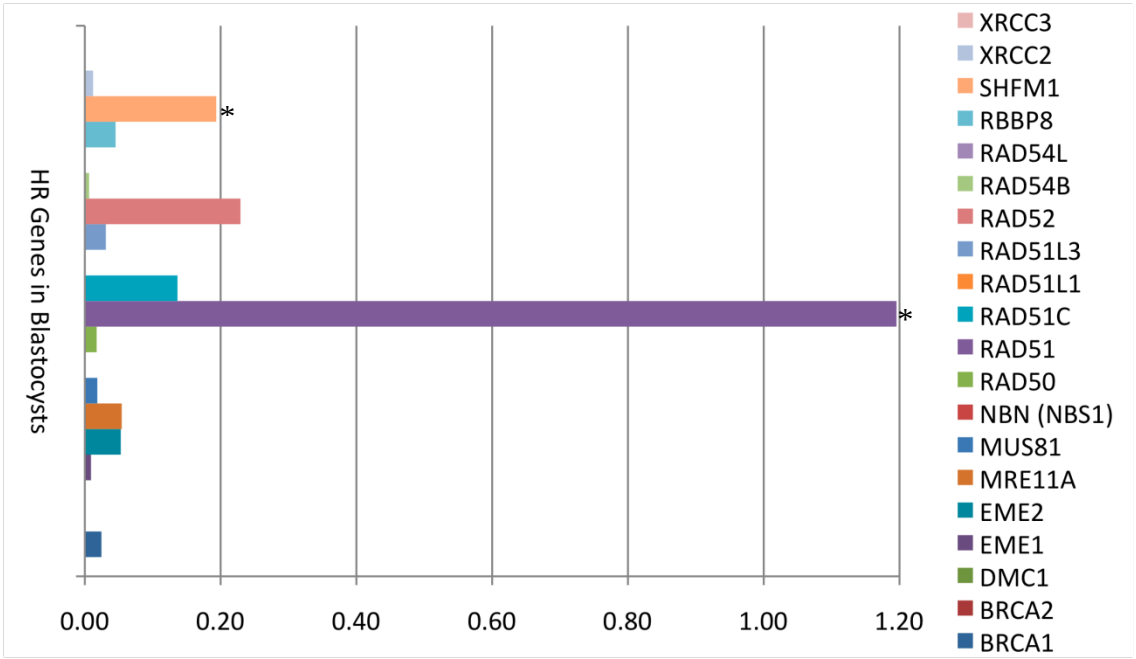
Pathway	Gene symbol	MII oocytes signal level	Blastocysts signal level	Cytoband	Reference sequence
HR	*BRCA1*	Not Detected	Medium	17q21	NM_007295
	BRCA2	Medium	Not detected	13q12.3	NM_000059
	DMC1	Not detected	Not detected	22q13.1	NM_007068
	EME1	Low	Low	17q21.33	NM_152463
	EME2	Medium	Medium	16p13.3	NM_001010865
	MRE11A	Medium	Medium	11q21	NM_005590
	MUS81	Medium	Medium	11q13	NM_025128
	NBN (NBS1)	Not detected	Not detected	8q21	NM_002485
	RAD50	High	Low	5q31	NM_005732
	RAD51	High	High	15q15.1	NM_002875
	RAD51C	High	High	17q22-q23	NM_002876
	RAD51L1	Low	Not detected	14q23-q24.2	NM_002877
	RAD51L3	Not detected	Medium	17q11	NM_002878
	RAD52	High	High	12p13-p12.2	NM_002879
	RAD54B	Medium	Low	8q22.1	NM_006550
	RAD54L	Not Detected	Not detected	1p32	NM_003579
	RBBP8	High	Medium	18q11.2	NM_002894
	SHFM1	Low	High	7q21.3-q22.1	NM_006304
	XRCC2	Low	Low	7q36.1	NM_005431
	XRCC3	Not detected	Not detected	14q32.3	NM_005432
NHEJ	*DCLRE1C*	Not Detected	Medium	10p13	NM_022487
	LIG4	Not Detected	Not detected	13q33-q34	NM_002312
	PRKDC	Not detected	Not detected	8q11	NM_006904
	XRCC4	Medium	Medium	5q13-q14	NM_003401
	XRCC5 (Ku 80)	Medium	High	2q35	NM_021141
	XRCC6 (Ku70)	High	High	22q13.2-q13.31	NM_001469
Other genes influencing DSBR pathway	*ATM*	Not Detected	Not detected	11q22-q23	NM_000051
	FANCD2	Low	Low	3p26	NM_033084
	ATR	Medium	N.D./Medium (2/3)	3q22-q24	NM_001184
	CHEK1	Medium	Medium	11q24-q24	NM_001274
	CHEK2	Not detected	Not detected	22q11|22q12.1	NM_007194
	MDC1	N.D./Low (2/3)	Medium	6pter-p21.31	NM_014641
	53BP1	High	Medium	15q15-q21	NM_005657

N.D.: Not detected

Figure 3.6: HR repair genes expressed in human MII oocytes (top) and blastocysts (bottom) with mRNA signal levels relative to *FEN1*



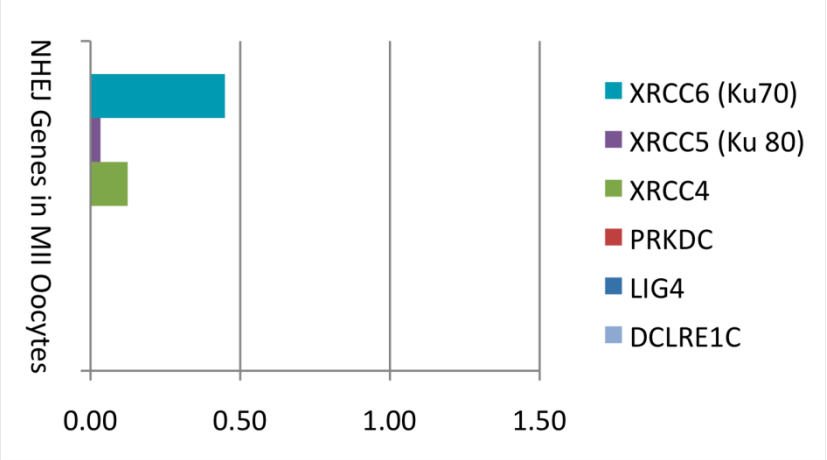
The asterisk denotes the genes that were expressed at significantly higher levels in MII oocytes compared to blastocysts ($p<0.05$).



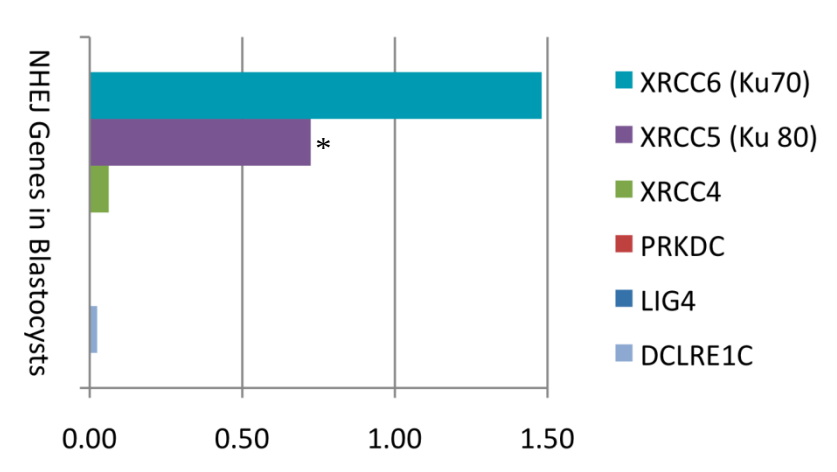
The asterisk denotes the genes that were expressed at significantly higher levels in blastocysts compared to MII oocytes ($p<0.05$).

The NHEJ genes *XRCC5* (*Ku70*) and *XRCC6* (*Ku80*) had medium to high expression levels in the MII oocytes and blastocysts, with significantly higher *XRCC5* (*Ku70*) mRNA levels ($p<0.05$ and 37 fold greater) in the blastocyst group versus the oocyte group (Table 3.5). *XRCC4* had high and medium mRNA expression levels in MII oocytes and blastocysts, respectively. *LIG4* expression, however, could not be detected in either group. (*LIG4* was detected in low levels in only two out of the three MII oocyte replicates). *DCLRE1C* (Artemis) was only expressed in the blastocysts (Table 3.7). Figure 3.7 displays the different expression patterns of NHEJ repair genes (relative to *FEN1)* in human MII oocytes and blastocysts.

Figure 3.7: NHEJ repair genes expressed in human MII oocytes (top) and blastocysts (bottom) with mRNA signal levels relative to *FEN1*



The asterisk denotes the genes that were expressed at significantly higher levels in MII oocytes compared to blastocysts ($p<0.05$).



The asterisk denotes the genes that were expressed at significantly higher levels in blastocysts compared to MII oocytes ($p<0.05$).

3.1.2.2.3 Mismatch Repair (MMR)

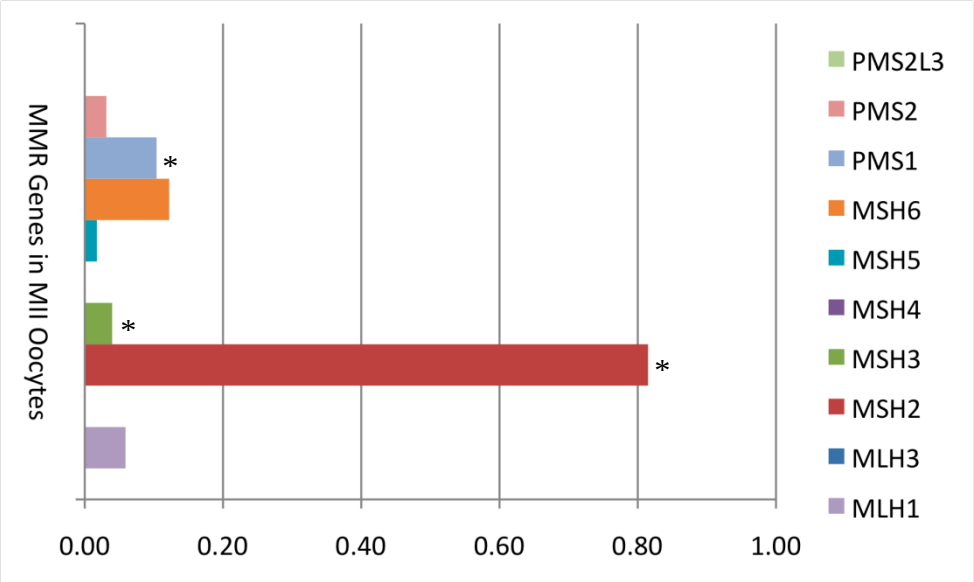
The main MMR genes (*MLH1*, *MSH2*, *MSH3*, *MSH6*, *PMS1* and *PMS2*) were expressed in human MII oocytes and blastocysts (Table 3.8). *MSH2* and *MSH6* mRNAs were both detected with high signals in oocytes and blastocysts. Three of the key MMR genes (*MSH2*, *MSH3* and *PMS1*) were expressed at significantly higher levels in the human MII oocytes versus the blastocysts (Table 3.4). *MSH5*, which is involved in meiosis, was also overexpressed in oocytes compared to blastocysts (which showed very low *MSH5* mRNA levels). Figure 3.8 displays the different expression patterns of MMR genes (relative to *FEN1)* in human MII oocytes and blastocysts.

Table 3.8: Expression levels of mRNAs coding for MMR genes in human MII oocytes and blastocysts

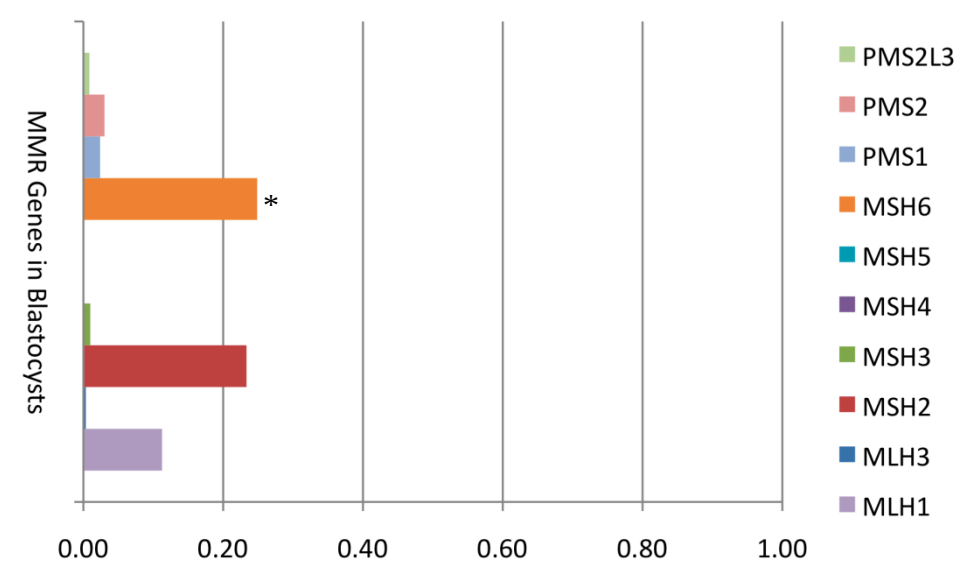
Gene symbol	MII oocytes signal level	Blastocysts signal level	Cytoband	Reference sequence
MLH1	Medium	High	3p21.3	NM_000249
MLH3	Not detected	N.D./Low (2/3)	14q24.3	NM_014381
MSH2	High	High	2p22-p21	NM_000251
MSH3	Medium	Low	5q11-q12	NM_002439
MSH4	Not detected	Not detected	1p31	NM_002440
MSH5	Low	Not detected	6p21.3	NM_002441
MSH6	High	High	2p16	NM_000179
PMS1	Medium	Medium	2q31-q33|2q31.1	NM_000534
PMS2	Medium	Medium	7p22.2	NM_000535
PMS2L3	Not detected	Low	7q11.23	NM_005395

N.D.: Not detected

Figure 3.8: MMR genes expressed in human MII oocytes (top) and blastocysts (bottom) with mRNA signal levels relative to *FEN1*



The asterisk denotes the genes that were expressed at significantly higher levels in MII oocytes compared to blastocysts ($p<0.05$).



The asterisk denotes the genes that were expressed at significantly higher levels in blastocysts compared to MII oocytes ($p<0.05$).

3.1.2.2.4 Nucleotide Excision Repair (NER)

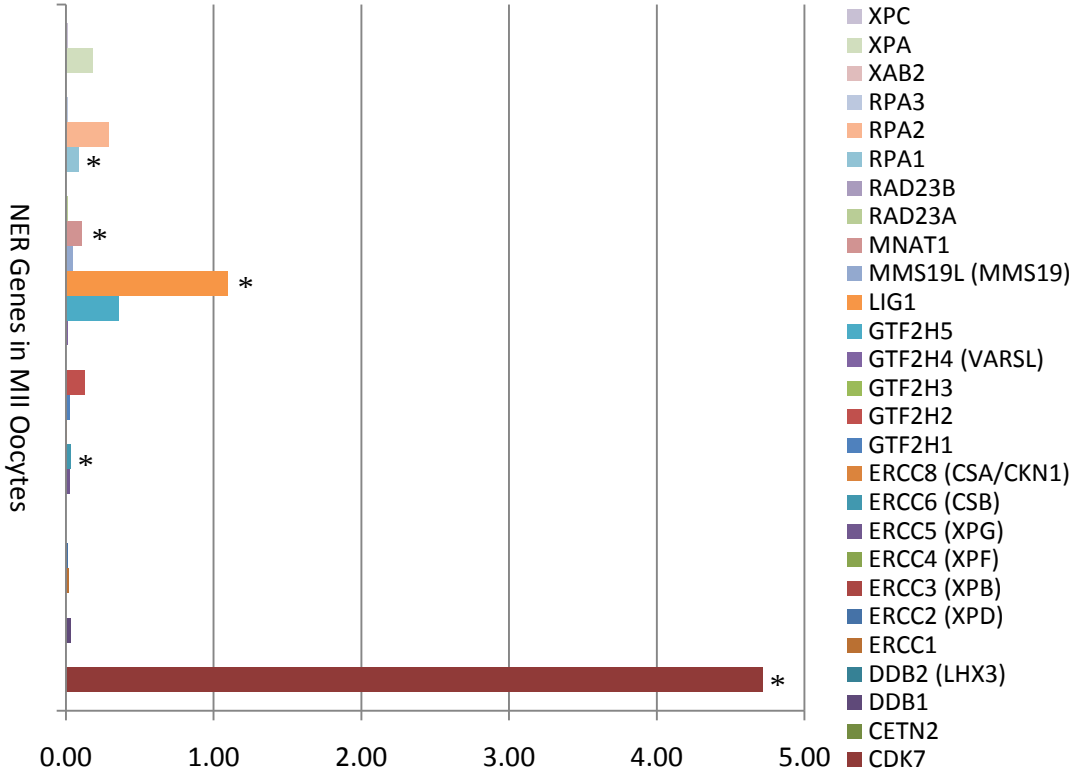
mRNA transcripts coding for 18 NER genes were detected in both MII oocytes and blastocysts (see Table 3.9 for full list of genes) some of which showed differential expression between the two groups. Most were significantly higher in the MII oocytes compared to the blastocysts, specifically the kinase subunits of TFIIH *CDK7*, *CCNH*, *MNAT1* as well as *RPA1*, *ERCC6* (*CSB)* and *LIG1* (Table 3.4). Only *RAD23B* had significantly higher mRNA levels in the blastocyst group (Table 3.5). The transcription-coupled repair (TCR) genes, *ERCC6* (*CSB*), *GTF2H1*, *2* & *5* and *MMS19L*, had medium or high mRNA expression levels in the human MII oocyte. Only *MMS19L* had a higher signal detected in the blastocyst group versus the oocyte group and the difference was not significant. Figure 3.9 displays the expression patterns of all NER genes (relative to *FEN1)* in human MII oocytes and blastocysts.

Table 3.9: Expression levels of mRNAs coding for NER genes in human MII oocytes and blastocysts

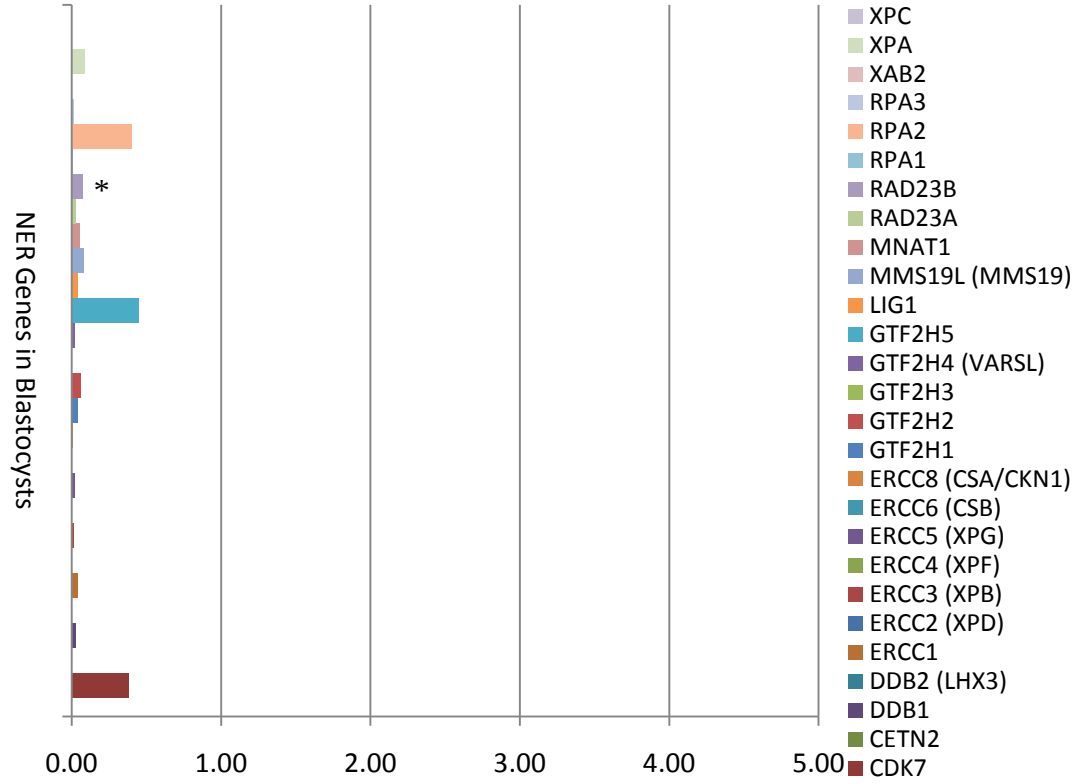
Gene symbol	MII oocytes signal level	Blastocysts signal level	Cytoband	Reference sequence
CCNH	High	Medium	5q13.3-q14	NM_001239
CDK7	High	High	5q12.1	NM_001799
CETN2	Not Detected	Not detected	Xq28	NM_004344
DDB1	Medium	Medium	11q12-q13	NM_001923
DDB2 (LHX3)	Not detected	Not detected	11p12-p11	NM_000107
ERCC1	Low	Medium	19q13.2-q13.3	NM_001983
ERCC2 (XPD)	Low	Not detected	19q13.3	NM_000400
ERCC3 (XPB)	Not detected	Low	2q21	NM_000122
ERCC4 (XPF)	Not detected	Not detected	16p13.3-p13.11	NM_005236
ERCC5 (XPG)	Medium	Low	13q22|13q33	NM_000123
ERCC6 (CSB)	Medium	Not detected	10q11.23	NM_000124
ERCC8 (CSA)	Low	Low	5q12.1	NM_000082
GTF2H1	Medium	Medium	11p15.1-p14	NM_005316
GTF2H2	High	Medium	5q12.2-q13.3	NM_001515
GTF2H3	Not detected	Not detected	12q24.31	NM_001516
GTF2H4 (VARSL)	Low	Medium	6p21.3	NM_001517
GTF2H5	High	High	6q25.3	NM_207118
LIG1	High	Medium	19q13.2-q13.3	NM_000234
MMS19L (MMS19)	Medium	High	10q24-q25	NM_022362
MNAT1	Medium	Medium	14q23	NM_002431
RAD23A	Low	Medium	19p13.2	NM_005053
RAD23B	Not Detected	High	9q31.2	NM_002874
RPA1	Medium	Low	17p13.3	NM_002945
RPA2	High	High	1p35	NM_002946
RPA3	N.D./Low (2/3)	Low	7p22	NM_002947
XAB2	Not detected	Not detected	19p13.2	NM_020196
XPA	High	High	9q22.3	NM_000380
XPC	Low	Not detected	3p25	NM_004628

N.D.: Not detected

Figure 3.9: NER genes expressed in human MII oocytes (top) and blastocysts (bottom) with mRNA signal levels relative to *FEN1*



The asterisk denotes the genes that were expressed at significantly higher levels in MII oocytes compared to blastocysts ($p<0.05$).



The asterisk denotes the genes that were expressed at significantly higher levels in blastocysts compared to MII oocytes ($p<0.05$).

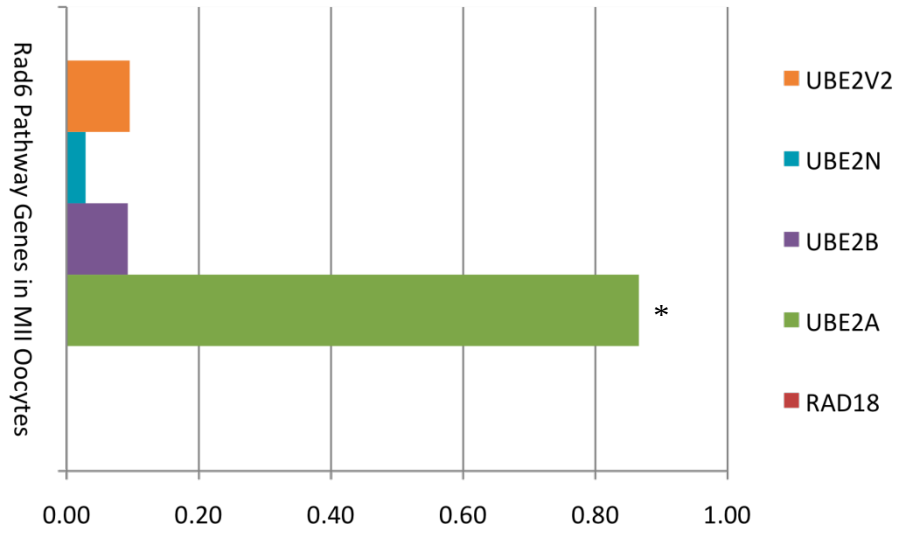
3.1.2.2.5 Other DNA repair pathways

The three genes involved in direct reversal of DNA damage *ALKBH2* (*ABH2*), *DEPC-1* and *MGMT* were all expressed in MII oocytes and blastocysts (Table 3.10). mRNA levels were significantly higher ($p<0.05$) in the oocytes (by 24, 7.7 and 6.8 folds repectively) (Table 3.4). The DNA-protein crosslinks repair gene *TDP1* had high mRNA expression levels picked up in the MII oocytes and low levels in the blastocysts (Table 3.10). The post-replication repair (RAD6 pathway) gene *UBE2A* had significantly higher expression levels in the human MII oocyte than the blastocyst ($p<0.05$ and 3.6 fold) (Table 3.4). *UBE2V2* and *UBE2B* showed higher expression levels in blastocysts versus MII oocytes (Table 3.5) though the differences in expression levels were not significant for *UBE2B*. Figure 3.10 displays the different expression patterns of the RAD6 pathway genes (relative to *FEN1)* in human MII oocytes and blastocysts.

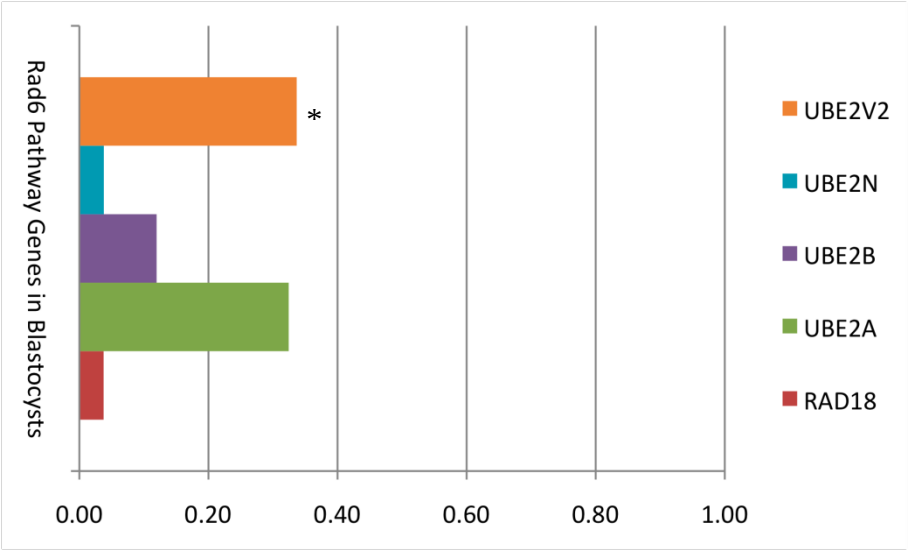
3.1.2.2.6 Genes associated with apoptosis

Analysis of few apoptosis related genes showed that the anti-apoptotic genes *BCL-2* and *BCL-W* were both expressed at low levels in human MII oocytes and blastocysts and *BCL-XL* was not detected in either of the two groups. The pro-apoptotic genes *BAX* and *BAK* were expressed at low levels in MII oocytes and medium levels in the blastocysts. *BAD*, another pro-apoptotic gene from the BCL-2 family, was only detected in the blastocysts.

Figure 3.10: RAD6 pathway genes expressed in human MII oocytes (top) and blastocysts (bottom) with mRNA signal levels relative to *FEN1*



The asterisk denotes the genes that were expressed at significantly higher levels in MII oocytes compared to blastocysts ($p<0.05$).



The asterisk denotes the genes that were expressed at significantly higher levels in blastocysts compared to MII oocytes ($p<0.05$).

Table 3.10: Expression levels of mRNAs coding for genes involved in DNA repair pathways in human MII oocytes and blastocysts

DNA repair pathway	Gene symbol	MII oocytes signal level	Blastocysts signal level	Cytoband	Reference sequence
Chromatin Structure	*CHAF1A*	High	High	19p13.3	NM_005483
	H2AFX (*H2AX*)	High	High	11q23.2-q23.3	NM_002105
Direct reversal of damage	*ALKBH2* (*ABH2*)	High	High	12q24.11	NM_001001655
	DEPC-1	Medium	Low	11p11.2	NM_139178
	MGMT	High	Medium	10q26	NM_002412
Rad6 Pathway	*RAD18*	Not Detected	Medium	3p25-p24	NM_020165
	UBE2A	High	High	Xq24-q25	NM_003336
	UBE2B	Medium	High	5q23-q31	NM_003337
	UBE2N	Medium	Medium	12q22	NM_003348
	UBE2V2	Medium	High	8q11.21	NM_003350
Modulation of nucleotide pools	*DUT*	High	High	15q15-q21.1	NM_001948
	NUDT1	Medium	High	7p22	NM_002452
	RRM2B	Not detected	Low	8q23.1	NM_015713
Repair of DNA-protein crosslinks	*TDP1*	High	Low	14q32.11	NM_018319
DNA Polymerases	*MAD2L2*	Medium	Medium	1p36	NM_006341
	PCNA	High	High	20pter-p12	NM_002592
	POLB	High	High	8p11.2	NM_002690
	POLD1	Not detected	Low	19q13.3	NM_002691
	POLE	Low	Low	12q24.3	NM_006231
	POLG	Low	Medium	15q25	NM_002693
	POLH	Not Detected	Not detected	6p21.1	NM_006502
	POLI	Low	Not detected	18q21.1	NM_007195
	POLK	Low	Not detected	5q13	NM_016218
	POLL	Low	Low	10q23	NM_013274
	POLM	Not detected	Low	7p13	NM_013284
	POLN	Not detected	Not detected	4p16.3	NM_181808
	POLQ	Medium	Low	3q13.33	NM_199420
	REV1L	High	Medium	2q11.1-q11.2	NM_016316
	REV3L	Not Detected	Not detected	6q21	NM_002912
Editing and processing nucleases	*EXO1*	Medium	Medium	1q42-q43	NM_003686
	FEN1	High	High	11q12	NM_004111
	FLJ35220	Not detected	Medium	17q25.3	NM_173627
	SPO11	Not Detected	Not detected	20q13.2-q13.3	NM_012444
	TREX1	Not Detected	Not detected	3p21.3-p21.2	NM_033629
	TREX2	Medium	Medium	Xq28	NM_080701
Genes defective in diseases associated with sensitivity to DNA damaging agents	*ATM*	Not Detected	Not detected	11q22-q23	NM_000051
	BLM	High	High	15q26.1	NM_000057
	FAAP24	Low	Low	19q13.11	NM_152266
	FANCA	Not detected	Not detected	16q24.3	NM_000135
	FANCB	Medium	Low	Xp22.2	NM_152633
	FANCC	Medium	Not detected	9q22.3	NM_000136
	FANCD2	Low	Low	3p26	NM_033084
	FANCE	Not Detected	Low	6p22-p21	NM_021922
	FANCF	Low	Not detected	11p15	NM_022725
	FANCG	Not detected	Medium	9p13	NM_004629
	FANCL	High	Medium	2p16.1	NM_018062
	FANCM	Medium	Low	14q21.3	NM_020937

Table 3.10 (continued): Expression levels of mRNAs coding for genes involved in DNA repair pathways in human MII oocytes and blastocysts

DNA repair pathway	Gene symbol	MII oocytes signal level	Blastocysts signal level	Cytoband	Reference sequence
Genes defective in diseases associated with sensitivity to DNA damaging agents	*FANCN*	Medium	Not detected	16p12.1	NM_024675
	KIAA1794	High	Medium	15q25-q26	NM_018193
	RECQL4	Low	Low	8q24.3	NM_004260
	WRN	Medium	Low	8p12-p11.2	NM_000553
Other conserved DNA damage response genes	*ATR*	Medium	N.D. /Medium (2/3)	3q22-q24	NM_001184
	CHEK1	Medium	Medium	11q24-q24	NM_001274
	CHEK2	Not detected	Not detected	22q11|22q12.1	NM_007194
	CLK2	Not detected	Not detected	1q21	NM_003993
	HUS1	Not detected	Not detected	7p13-p12	NM_004507
	MDC1	N.D./Low (2/3)	Medium	6pter-p21.31	NM_014641
	PER1	Medium	Medium	17p13.1-17p12	NM_002616
	RAD1	Medium	Medium	5p13.2	NM_002853
	RAD17	High	High	5q13	NM_002873
	RAD9A	Medium	Low	11q13.1-q13.2	NM_004584
	TP53	N.D./Low (2/3)	N.D./Low (2/3)	17p13.1	NM_000546
Other identified genes with suspected DNA repair function	*APTX*	Medium	High	9p13.3	NM_175073
	DCLRE1A	High	Medium	10q25.1	NM_014881
	DCLRE1B	Low	Not detected	1p13.2	NM_022836
	HEL308	Low	Low	4q21.23	NM_133636
	NEIL3	Not Detected	Not detected	4q34.3	NM_018248
	RDM1	Not detected	Not detected	17q11.2	NM_145654
	RECQL	Low	Medium	12p12	NM_002907
	RECQL5	Not detected	Not detected	17q25.2-q25.3	NM_001003715
	RPA4	Low	Not detected	Xq21.33	NM_013347

N.D.: Not detected

3.1.3 Summary of microarray analysis results

- The triplicate sets of extracted RNA from three pooled human MII oocytes or blastocysts were successfully amplified and hybridised on the Applied Biosystems' microarrays.

- Global gene expression analysis across all chromosomes and the hierarchical cluster showed that a greater proportion of genes were highly expressed in the blastocysts compared to the oocytes.

- Among the 129 DNA repair genes that were included in a t-test, 55 were found to be differentially expressed in blastocysts compared to oocytes. Most (73%) were detected in lower levels in the blastocyst group compared to the oocyte group ($p<0.05$, fold change > 3).

- While all DNA repair pathways were represented in human oocytes and blastocysts, the expression profiles of DNA repair genes were different in the two sample groups.

- Most BER (11/17), MMR (6/10) and NER (18/28) genes were expressed in both MII oocytes and blastocysts. For DSBR, 12/19 genes involved in HR and 3/6 genes involved in NHEJ were expressed in both groups. The mRNA expression levels for the different genes were distinct in the two sample groups.

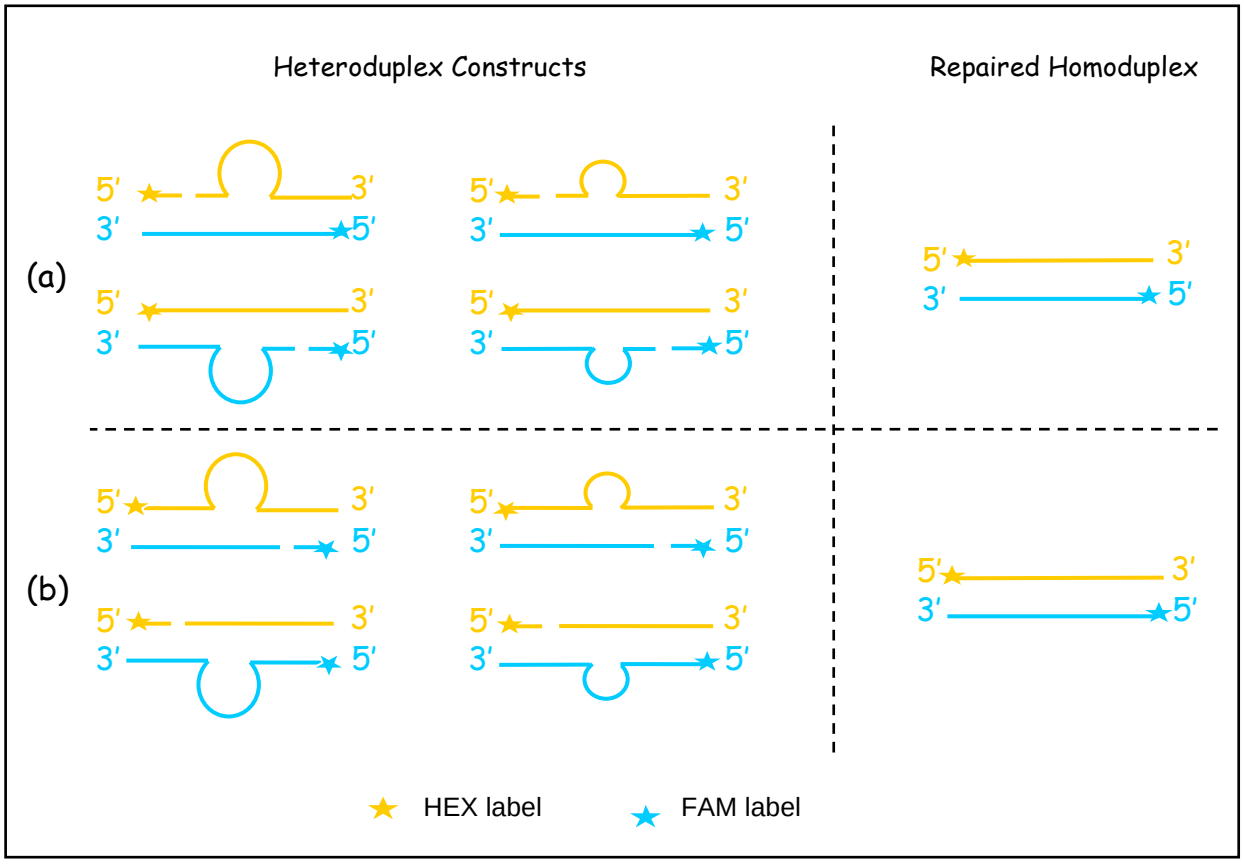
3.2 Development of a cell free in vitro functional assay for the assessment of IDL and mismatch repair

3.2.1 Formation of heteroduplex DNA constructs containing insertion/deletion loops and base-base mismatches

3.2.1.1 Design of DNA templates

Heteroduplex DNA molecules were made to include either an insertion/deletion loop (IDL) of 3, 21 or 24 residues in size or a G.T or A.C mismatch using the method described in section 2.6.2. The constructs were prepared both with and without a 5'nick, which directs the repair to a specific strand. Figure 3.11 illustrates the use of a nick to direct the repair of IDLs. The strands of the heteroduplex DNAs were either fluorescently labelled for accurate sizing on the ABI Prism™ 310 automated genetic analyzer (section 2.6.4.1) or non-labelled for SSCP/heteroduplex analysis (section 2.6.5).

Figure 3.11: Diagram showing the flexible structure of the heteroduplex construct for IDLs



The size of the IDL could be varied and the nick placed on either the long or short strand, giving a wide variety of heteroduplex molecules. The MMR machinery should recognise the presence of the loops as insertions in section (a) and deletions in section (b) and repair the nicked strand by omitting the extra bases or inserting additional bases, respectively.

3.2.1.1.1 Selection of the DNA sequence and PCR primers

The primers were selected so that the amplified products would be around 300 or 600 base pairs (bp) in size and included the selected CTG triplet repeat region involved in DM1 or the rs1981929 SNP site as well as a single recognition sequence for a nicking enzyme (section 2.4.2.1). Only one nicking enzyme had a single recognition site in each of the two sequences. The nicking endonucleases Nt. BbvCI and Nt. BstNBI were used to create a single nick on one strand of the duplex DNA construct 5' to the CTG repeat and A/G SNP site, respectively. Figure 3.12 summarises the sequence structures and the full sequences are shown in Figure 3.13.

Figure 3.12: Diagrams illustrating the design of the DNA sequence used for the formation of heteroduplex DNA constructs with IDLs (top) and single base mismatches (bottom)

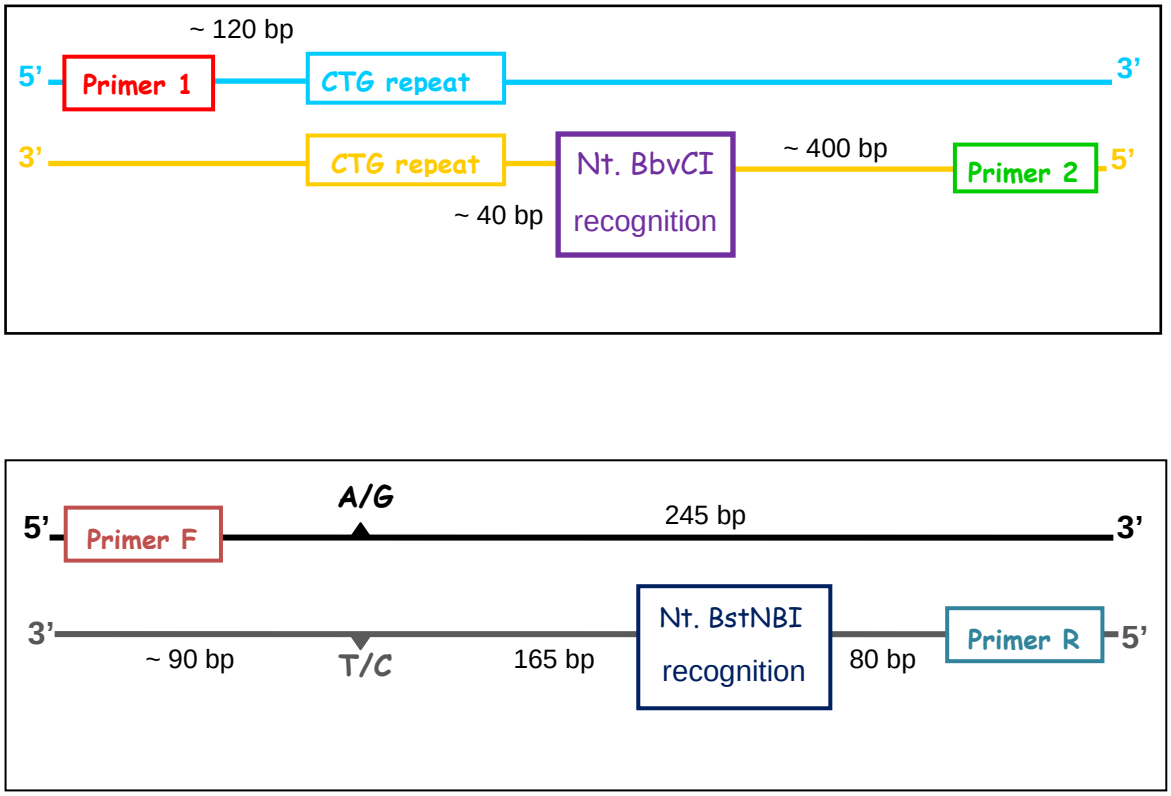
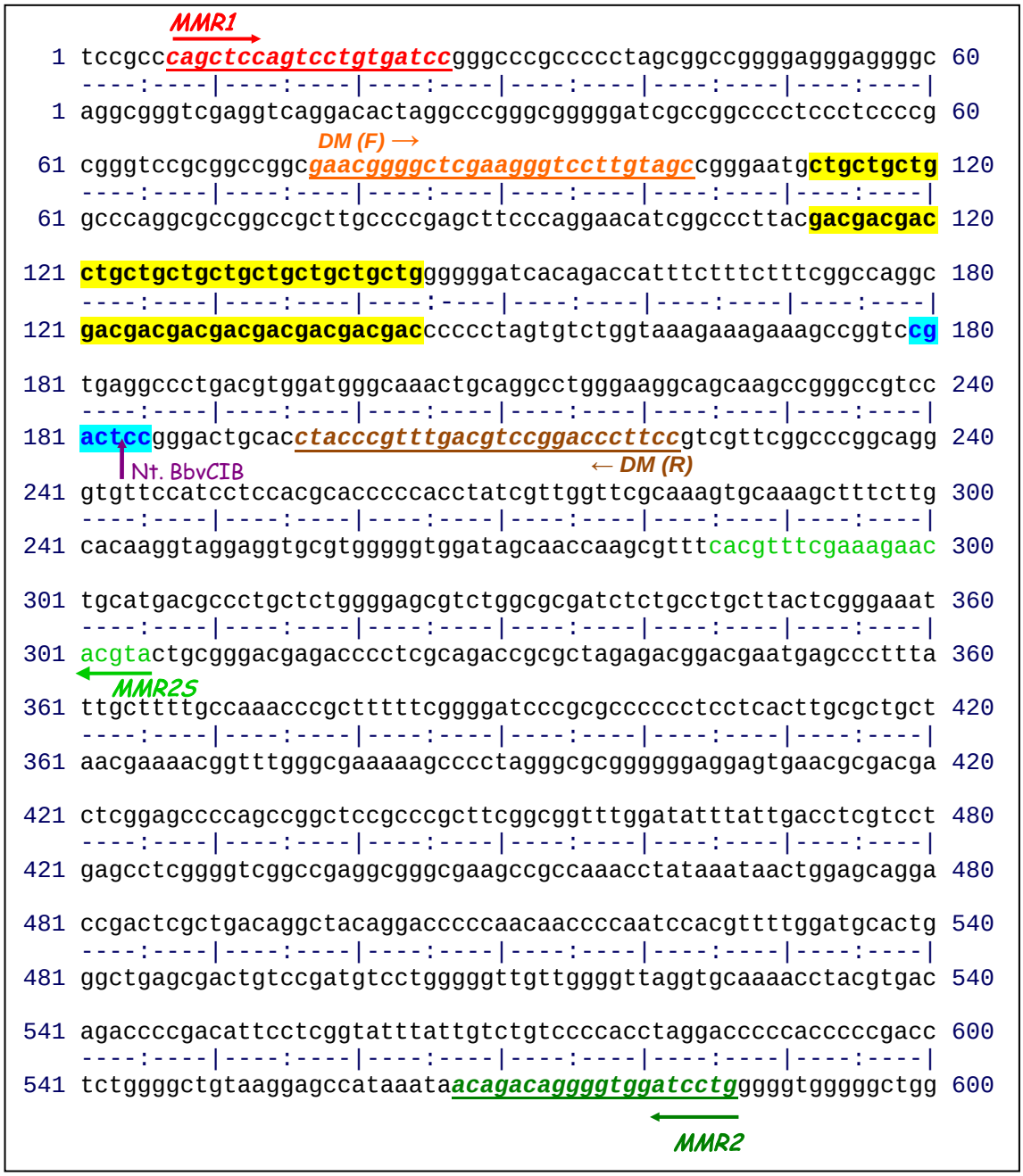


Figure 3.13: Genomic DNA sequences used for the formation of heteroduplex DNA constructs

1) DNA sequence used for the formation of heteroduplex constructs with IDLs located on 19q13.2-q13.3 (accession number: L00727, selection 10561-11161), upstream of the *DMPK* gene




```
          MMR1
          ---->
  1 tccgccCAGCTCCAGTCCTGTGATCCgggcccgccccctagcggccggggagggaggggc 60
    ----:----|----:----|----:----|----:----|----:----|----:----|
  1 aggcgggtcgaggtcaggacactaggcccgggcgggggatcgccggcccctccctccccg 60
                        DM (F) ---->
 61 cgggtccgcggccggcGAACGGGGCTCGAAGGGTCCTTGTAGCcgggaatgCTGCTGCTG 120
    ----:----|----:----|----:----|----:----|----:----|----:----|
 61 gcccaggcgccggccgcttgccccgagcttcccaggaacatcggcccttacGACGACGAC 120

121 CTGCTGCTGCTGCTGCTGCTGggggggatcacagaccatttctttctttcggccaggc 180
    ----:----|----:----|----:----|----:----|----:----|----:----|
121 GACGACGACGACGACGACGACccccctagtgtctggtaaagaaagaaagccggtccg 180

181 tgaggccctgacgtggatgggcaaactgcaggcctgggaaggcagcaagccgggccgtcc 240
    ----:----|----:----|----:----|----:----|----:----|----:----|
181 actccgggactgcacCTACCCGTTTGACGTCCGGACCCTTCCgtcgttcggcccggcagg 240
      Nt. BbvCIB                          <---- DM (R)
241 gtgttccatcctccacgcacccccacctatcgttggttcgcaaagtgcaaagctttcttg 300
    ----:----|----:----|----:----|----:----|----:----|----:----|
241 cacaaggtaggaggtgcgtgggggtggatagcaaccaagcgtttcacgtttcgaaagaac 300

301 tgcatgacgccctgctctggggagcgtctggcgcgatctctgcctgcttactcgggaaat 360
    ----:----|----:----|----:----|----:----|----:----|----:----|
301 acgtactgcgggacgagacccctcgcagaccgcgctagagacggacgaatgagcccttta 360
    <----
     MMR2S
361 ttgcttttgccaaacccgctttttcggggatcccgcgcccccctcctcacttgcgctgct 420
    ----:----|----:----|----:----|----:----|----:----|----:----|
361 aacgaaaacggtttgggcgaaaaagcccctagggcgcgggggggaggagtgaacgcgacga 420

421 ctcggagccccagccggctccgcccgcttcggcggtttggatatttattgacctcgtcct 480
    ----:----|----:----|----:----|----:----|----:----|----:----|
421 gagcctcggggtcggccgaggcgggcgaagccgccaaacctataaataactggagcagga 480

481 ccgactcgctgacaggctacaggacccccaacaaccccaatccacgttttggatgcactg 540
    ----:----|----:----|----:----|----:----|----:----|----:----|
481 ggctgagcgactgtccgatgtcctgggggttgttggggttaggtgcaaaacctacgtgac 540

541 agaccccgacattcctcggtatttattgtctgtccccacctaggacccccacccccgacc 600
    ----:----|----:----|----:----|----:----|----:----|----:----|
541 tctggggctgtaaggagccataaataACAGACAGGGGTGGATCCTGggggtgggggctgg 600
                                          <----
                                           MMR2
```

The DNA sequence shows the CTG triplet repeat (highlighted in yellow), the recognition site for the nicking enzyme (highlighted in blue), the PCR primer sequences (forward in red and reverse in green) and the DM primer sequences (forward in orange and reverse in brown).

The Sixpack display tool, available on the EMBL-EBI website (http://srs.ebi.ac.uk), was used to display the DNA sequence in both directions.

Figure 3.13 (continued): Genomic DNA sequences used for the formation of heteroduplex DNA constructs

2) DNA sequence used for the formation of heteroduplex constructs with single base mismatches including the rs1981929 SNP site located within intron 8 of the *MSH2* gene on 2p21 (Chromosome 2: 47,525,573–47,526,053)

```
  1 TAATGTCATGTTTCTGCATCTATATTACTTGTTGGGTTTACAGACGAGGTAGTGTATTAT 60
    ----:----|----:----|----:----|----:----|----:----|----:----|
  1 ATTACAGTACAAAGACGTAGATATAATGAACAACCCAAATGTCTGCTCCATCACATAATA 60
                  rs1981929 F
                 ------------>
 61 TAGTGGGAAGCTTTGAGTGCTACATCATCTCCCTTTCTATAAAATAAATTGAGTACGAAA 120
    ----:----|----:----|----:----|----:----|----:----|----:----|
 61 ATCACCCTTCGAAACTCACGATGTAGTAGAGGGAAAGATATTTTATTTAACTCATGCTTT 120

121 CAATTTGAATTAAAACACCTGAGTAAATAGTAACTTTGGAGACCTGIACTGTACTATTTGTA 180
    ----:----|----:----|----:----|----:----|----:- ---|----:----|
121 GTTAAACTTAATTTTGTGGACTCATTTATCATTGAAACCTCTGGACITGACATGATAAACAT 180

181 CCTTTTGGATCAAATGATGCTTGTTTATCTCAGTCAAAATTTTATGATTTGTATTCTGTA 240
    ----:----|----:----|----:----|----:----|----:----|----:----|
181 GGAAAACCTAGTTTACTACGAACAAATAGAGTCAGTTTTAAAATACTAAACATAAGACAT 240

241 AAATGAGATCTTTTTATTTGTTTGTTTTACTACTTTCTTTTAGGAAAACACCAGAAATTA 300
    ----:----|----:----|----:----|----:----|----:----|----:----|
241 TTTACTCTAGAAAAATAAACAAACAAAATGATGAAAGAAAATCCTTTTGTGGTCTTTAAT 300

301 TTGTTGGCAGTTTTTGTGACTCCTCTTACTGATCTTCGTTCTGACTTCTCCAAGTTTCAG 360
    ----:----|----:----|----:----|----:----|----:----|----:----|
301 AACAACCGTCAAAAACACTGAGGAGAATGACTAGAAGCAAGACTGAAGAGGTTCAAAGTC 360
                  | Nt. BstNBI

361 GAAATGATAGAAACAACTTTAGATATGGATCAGGTATGCAATATACTTTTTAATTTAAGC 420
    ----:----|----:----|----:----|----:----|----:----|----:----|
361 CTTTACTATCTTTGTTGAAATCTATACCTAGTCCATACGTTATATGAAAAATTAAATTCG 420
                                      <------------
                                      rs1981929 R

421 AGTAGTTATTTTTAAAAAGCAAAGGCCACTTTAAGAAAGTTTGTAGATTTTTCTTTTTAG 480
    ----:----|----:----|----:----|----:----|----:----|----:----|
421 TCATCAATAAAAATTTTTCGTTTCCGGTGAAATTCTTTCAAACATCTAAAAAGAAAAATC 480
```

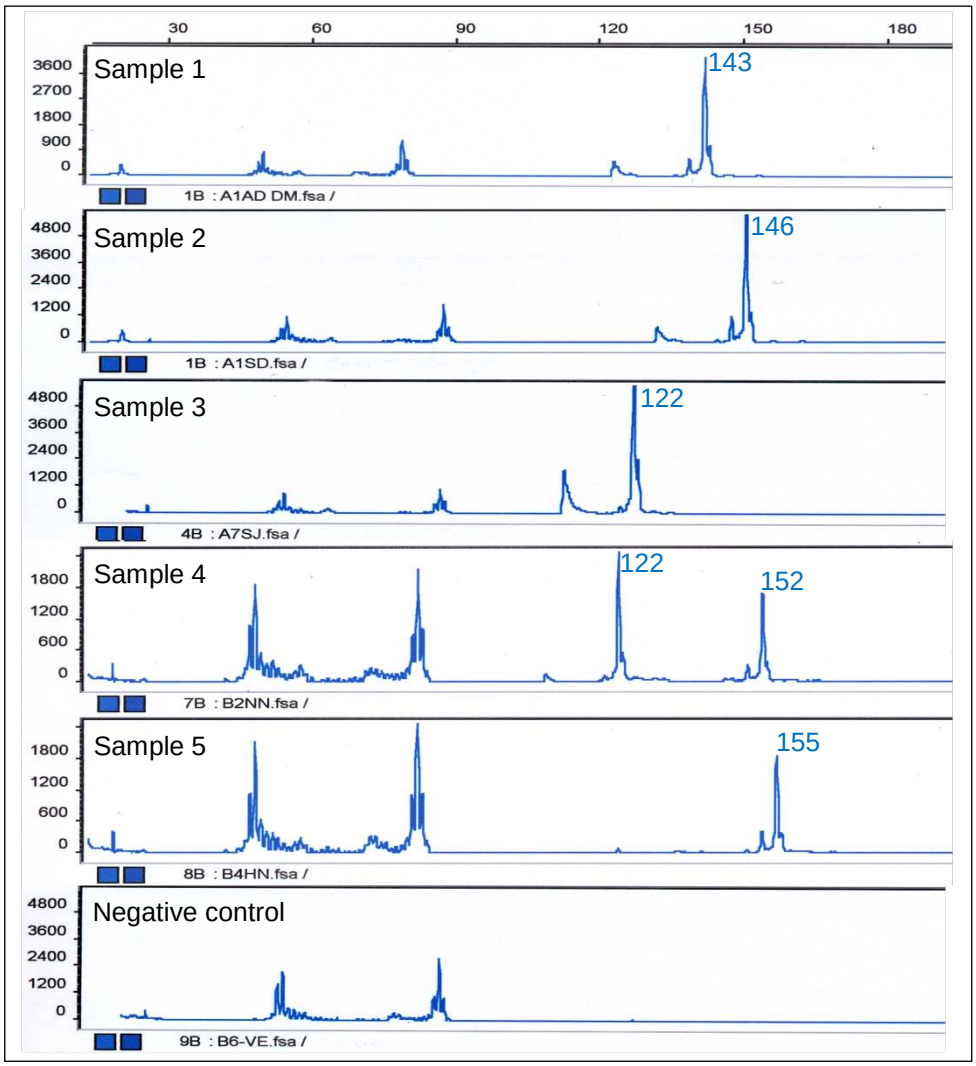
The DNA sequence shows the A/G SNP site (highlighted in yellow), the recognition site for the nicking enzyme (highlighted in blue) and the PCR primer sequences (forward in red and reverse in green).

The Sixpack display tool, available on the EMBL-EBI website (http://srs.ebi.ac.uk), was used to display the DNA sequence in both directions.

3.2.1.1.2 Selection of DNA samples

Genomic DNAs from homozygous individuals with different alleles were amplified and used for the formation of heteroduplex DNA constructs (section 2.6.2.1). The DNA samples used for the formation of IDL constructs were selected based on allele size differences. Twelve control genomic DNA samples were amplified by fluorescent PCR (F-PCR) using the DM primer set labelled with FAM and were visualised as blue peaks on the ABI Prism™ 310. Some of the results panels obtained from GeneScan™ fragment size analysis are shown in Figure 3.14 and the allele sizes are listed in Table 3.11.

Figure 3.14: GeneScan™ fragment size analysis result panels showing five amplified products for the DM1 triplet repeat locus and a negative control (missing genomic DNA)



The peaks represent the amplified DNA fragments at the DM repeat locus (120-152bp long) obtained from control DNA samples. The first two peaks that are under 90bp are unspecific products resulting from primer dimers and were observed in the negative control sample as well. Samples 1, 2, 3 and 5 were homozygous and sample 4 was heterozygous for the DM repeat locus. Samples 1-3 were used for the formation of heteroduplex constructs with IDLs.

Table 3.11: Allele sizes obtained from twelve genomic DNA samples amplified by F-PCR using the DM primers

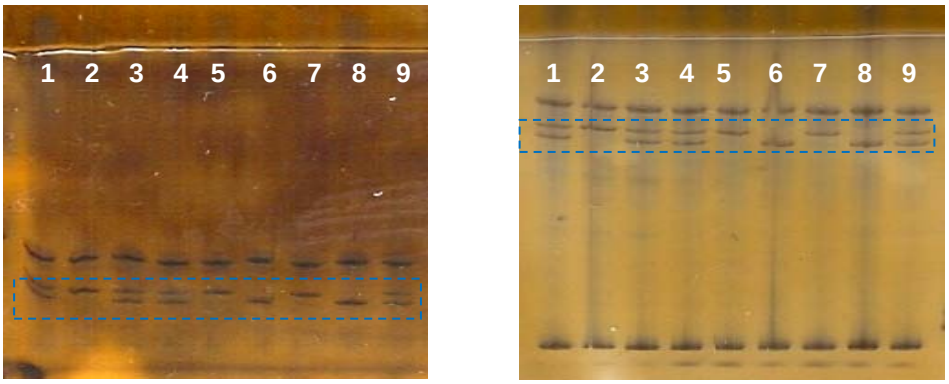
DNA sample	1	2	3	4	5	6	7	8	9	10	11	12
Allele sizes (bp)	143	146	122	122 152	155	129 149	122 168	122	122	144 148	142 152	122 149

Samples 1, 2 and 3 were selected and used for the formation of heteroduplex DNA constructs with 3, 21 and 24-nucleotide IDLs.

As only homozygous samples were considered; this eliminated samples 4, 6, 7, 10, 11 and 12. Amongst the homozygous samples, samples 1, 2 and 3 with allele sizes of 146, 143 and 122, respectively, were selected and used in the formation of heteroduplex DNA contructs with 3, 21 and 24-nucleotide IDLs. The IDL could be 9 nucleotides in size if samples 2 and 5 were used or 33 nucleotides if samples 3 and 5 were chosen (Figure 3.14 and Table 3.11).

Samples that were homozygous A/A and G/G at the rs1981929 SNP locus were needed for the formation of constructs with a G.T or A.C mismatch. Nine control genomic DNA samples were amplified using the rs1981929 primer set and analysed on SSCP Homogenous 12.5 and 20 gels (Figure 3.15).

Figure 3.15: SSCP analysis on PhastGel® Homogeneous 12.5 (left) and 20 (right) run at 10ºC with long pre-run showing nine genomic DNA samples amplified with the rs1981929-F/R primers



The top band was common to all samples. The lower bands (in the blue boxes) were representative of the A or G alleles at the rs1981929 locus. Samples 1, 3, 4 and 9 were heterozygous as they show both bands/alleles. Samples 2, 5 and 7 were homozygous for one allele and samples 6 and 8 were homozygous for the other allele.

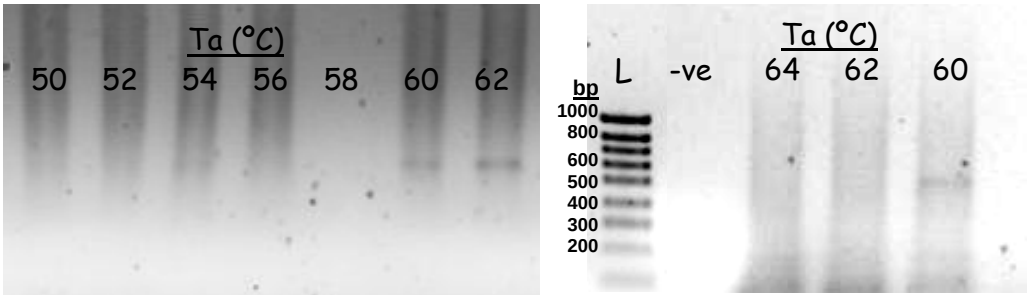
Homozygous samples could be grouped into two groups depending on the genotype (A/A or G/G for the rs1981929 SNP). Samples 2, 5 and 7 seemed to have the same allele and samples 6 and 8 seemed to have the other; while samples 1, 3, 4 and 9 were heterozygous. Samples 7 and 8 were sequenced to confirm homozygosity and determine the nucleotide base at the SNP site (method described in section 2.6.6). It was important to confirm that no other change in sequence (mutation or SNP) had caused the conformational change observed on the SSCP gels. The full sequences obtained on the ABI Prism™ 3100 genetic analyzer are shown in the Appendix (Figure C.1). Samples 7 and 8 were identified as G/G and A/A, respectively; the two samples were used for the formation G.T and A.C heteroduplex DNA constructs.

3.2.1.2 Optimisation of PCR conditions

The PCR using the DM primers had been previously optimised (Piyamongkol *et al.*, 2001). The PCR using the MMR and MSH2 primers, however, needed to be optimised to avoid problems such as non-specific annealing of primers. This was done empirically by varying the annealing temperature, DNA template amount, primer, $MgCl_2$ and DNA polymerase concentrations (section 2.4.2.2.1).

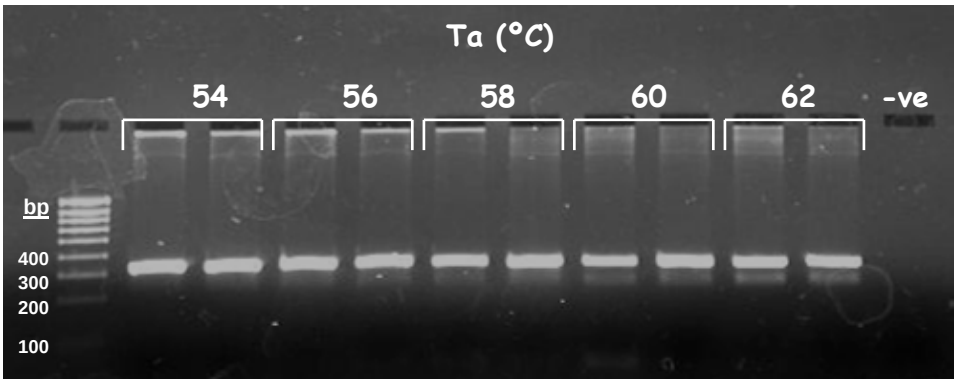
The amount of genomic DNA used as a template for the amplification reaction was generally 1µl (equivalent to 100ng DNA) for 25µl reactions. The calculated melting temperatures (Tm) were 64ºC and 64-66ºC for the MMR and rs1981929 primers, respectively. The annealing temperature (Ta) for these PCRs was expected to be at ~5 to 10ºC below the Tm; thus 55-60ºC were estimated to be reasonable annealing temperatures for both primer sets. In order to determine the temperature that would give the most stringent PCR conditions, different annealing temperatures were tested in parallel using the Mastercycler Gradient® (Eppendorf, UK). The temperatures ranged between 50ºC and 64ºC for the MMR primers and 54ºC and 62ºC for the rs1981929 primers. The PCR products were assessed by comparing their band intensities on agarose gels (Figures 3.16 and 3.17). The optimal annealing temperature was found to be 60ºC for the MMR and 56ºC for the rs1981929 primers.

Figure 3.16: 1% agarose gel (run at 55V) showing annealing temperature (Ta) optimisations for the MMR primer set



PCR products using the MMR primers at different annealing temperatures: 50-62ºC (left) and 60-64ºC (right); Expected product size: ~600bp
L: 1Kb ladder (1000, 800, 700, 600, 500, 400, 300, 200 and 100bp); -ve:negative control
The optimal annealing temperature for the MMR primers was 60ºC, as the product amplified at this temperature had the strongest band intensity.

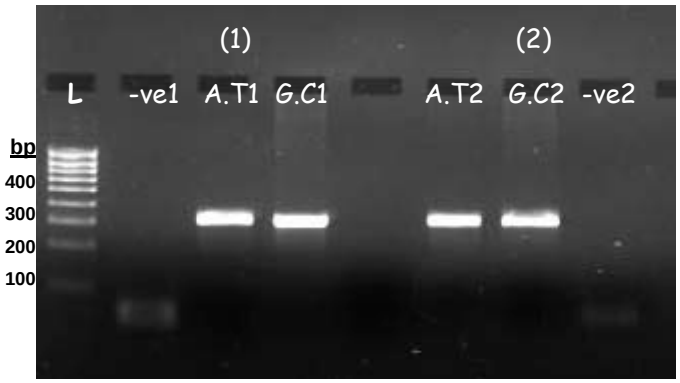
Figure 3.17: 2% agarose gel (run at 70V) showing annealing temperature (Ta) optimisation for the rs1981929 primer set



PCR products using the rs1981929 primers at different annealing temperatures: 54-62ºC;
Expected product size: 328bp; L: 1Kb ladder; -ve:negative control
The optimal annealing temperature for the rs1981929 primers was 56ºC. Unspecific products (~250bp) were detected at higher annealing temperatures, especially at 60 and 62ºC.

The initial PCR conditions used for the rs1981929 primers (0.2µM primers, HiFi DNA polymerase and 56ºC annealing temperature) were tested again with one primer biotinylated in each set. The products were checked on a 2% agarose gels (Figure 3.18).

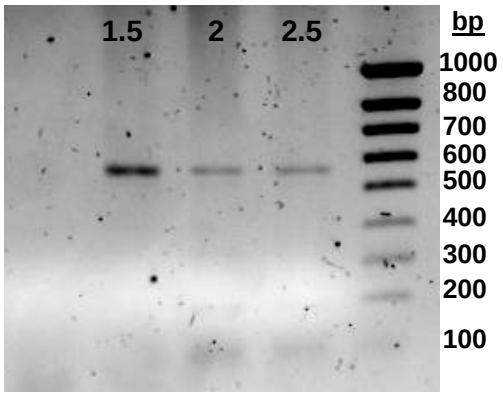
Figure 3.18: 2% agarose gel (run at 75V) showing the amplified products obtained from the two selected homozygous DNA samples using the rs1981929-B-F+ rs1981929-R (1) & rs1981929-F+ rs1981929-B-R (2) primer sets



L: 1Kb ladder; -ve:negative control
Both rs1981929 primer sets 1 and 2 were amplified efficiently (strong band intensities) at 56ºC using 0.2μM primers and HiFi DNA polymerase; no unspecific products were observed.

As the amplification was satisfactory, no further optimisation was necessary for the rs1981929 primers. The MMR primers, however, required further optimisation. In order to this, a magnesium titration was performed. The amplification was assessed at 1.5, 2.0 and 2.5mM $MgCl_2$ concentrations. The optimal concentration was 1.5mM (Figure 3.19).

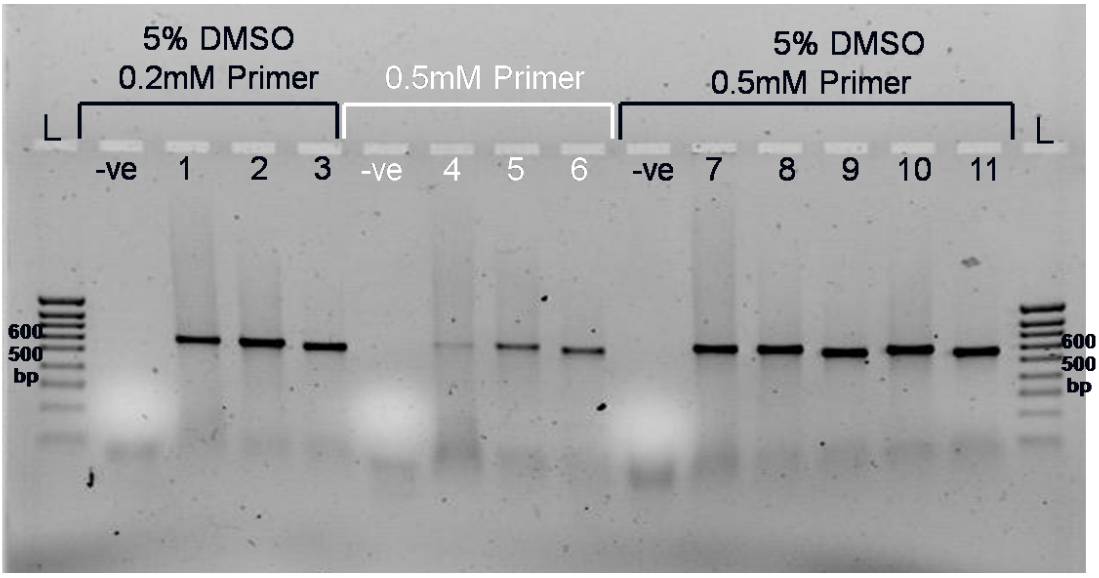
Figure 3.19: 1% agarose gel for the titration of $MgCl_2$ for PCR using the MMR primers



Three PCR products amplified at different $MgCl_2$ concentrations (1.5, 2.0 and 2.5mM) were compared on this gel. The optimal $MgCl_2$ concentration was 1.5mM as its corresponding band had the strongest intensity.

The PCR was further optimised by varying the DNA polymerase and primer concentrations used. AmpliTaq Gold® and Expand High Fidelity PCR system (HiFi) gave better results than AmpliTaq® DNA polymerase. The primer concentration was also increased from 0.2μM to 0.5μM. However, the PCR amplification was still not efficient, particularly for the production of heteroduplex constructs. The addition of 5% DMSO to the PCR mix significantly improved the amplification efficiency (Figure 3.20). The same PCR conditions were used for the MMR1+MMR2/MMR2S primer sets, which targeted the same region but produced long/short products, repectively. A summary of the optimised PCR conditions for the MMR primers is listed in Table 3.12.

Figure 3.20: 1% agarose gel showing increased amplification efficiency with increased primer concentration in the presence of 5% DMSO



L: 1Kb ladder; -ve: negative control
The strongest bands were obtained for the PCR products amplified with 0.5μM of MMR primers and 5% DMSO. These conditions were better than 0.2μM primers with DMSO or 0.5μM primers without DMSO.

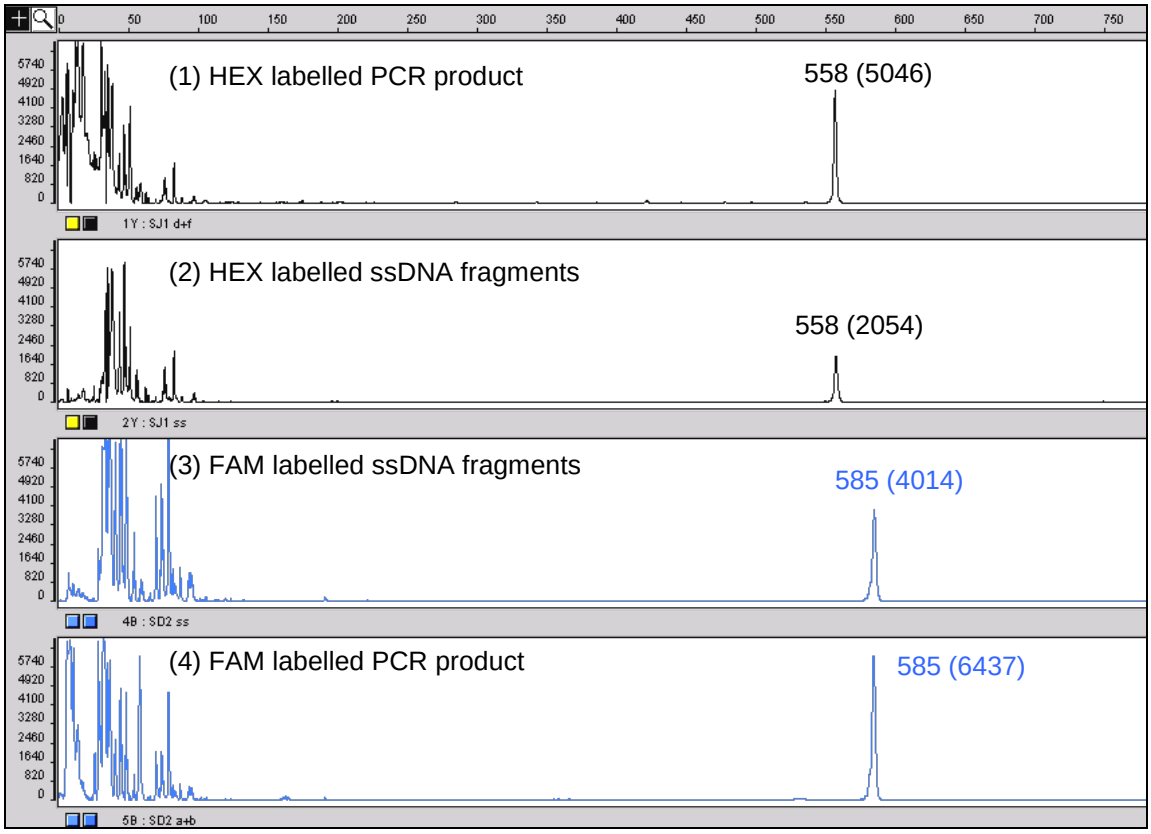
Table 3.12: Optimal PCR reaction conditions for primers used for the formation of heteroduplex DNA constructs with G.T/A.C mismatches (top row) and IDLs (lower rows)

Primer set	Product size (bp)	Annealing temperature	Primer concentration	DNA polymerase	Other changes
rs1981929-F & rs1981929-R	328	56ºC	0.2μM	HiFi	-
DM-F & DM- R	122-200	60ºC	0.2μM	AmpliTaq®/ HiFi	-
MMR1 & MMR2/MMR2S	558-639/ 277-358	60ºC	0.5μM	AmpliTaq Gold®/ HiFi	5% DMSO

3.2.1.3 Separation of DNA strands

Single strands were isolated using Dynabeads® and purified using the MinElute PCR Purification kit (QIAGEN, UK) (section 2.6.2.2). The labelled single strands prepared for the IDL constructs were detected on the ABI Prism™ 310 genetic analyzer. Figure 3.21 displays four panels from the ABI GeneScan™ analysis showing a PCR product labelled with HEX (black peak in panel 1), the HEX-labelled single strands isolated from that product (black peak in panel 2), a PCR product labelled with FAM (blue peak in panel 4) and the FAM-labelled single strands isolated from that product (blue peak in panel 3).

Figure 3.21: GeneScan™ fragment size analysis panels showing the dsDNA PCR products and the isolated ssDNA fragments for the formation of heteroduplex DNA constructs containing IDLs



The peaks represent the sized DNA fragments and are shown in blue/black for the FAM/HEX labelled samples, respectively. The sizes of the DNA fragments are displayed above the corresponding peaks. The peak heights, which relate to the product quantity, are given in parentheses.

Panels 2 and 3 show the ssDNA fragments separated from the HEX labelled (panel one) and FAM labelled (panel 4) PCR products, respectively. The peak heights were lower in the ssDNA samples compared to the starting PCR products.

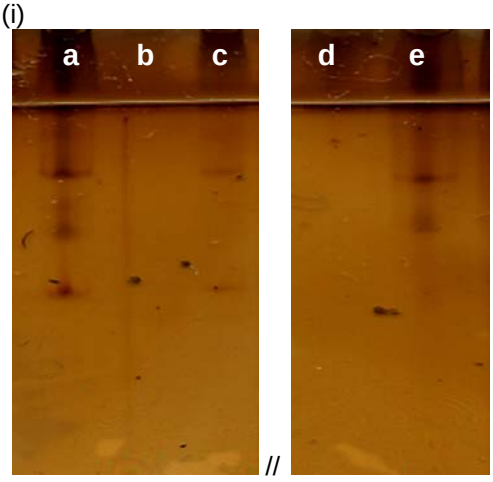
Note: The sizing of labelled DNA fragments on the ABI Prism™ shifts by up to 3bp depending on the DNA sequence (i.e. forward & reverse strands migrate differently).

The single stranded DNA was clearly visible on the genetic analyzer confirming the success of the separation of the DNA strands using Dynabeads®. Clearly, some loss of product took place at this stage as the peak heights, which correspond to the concentration of the sample, were lower for the single strands compared to the starting PCR products.

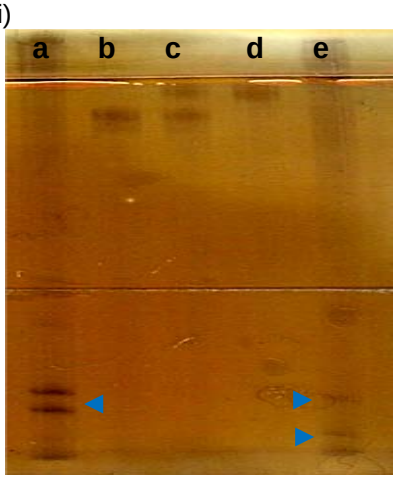
The labelled single strands could not be observed on SSCP gels. In fact, both the amplified products and their corresponding single strands were difficult to visualise on SSCP. To improve detection, the PhastSystem™ SSCP protocol was empirically optimised by changing the run temperature (4ºC, 7ºC, 10ºC, 15ºC and 20ºC), the polyacrylamide concentration of the gel (PhastGel® Homogeneous 12.5 or 20), the length of the pre-run (an extended pre-run gives a continuous buffer system and allows better separation) and the total separation time (in volthours or Vh). The best conditions for the 561 and 585bp PCR products were at a total run time of 350Vh on PhastGel® Homogeneous 12.5 gels at 4ºC with an extended pre-run. The same procedure was tested with non-labelled primers which improved the visibility of the DNA fragments on SSCP gels, especially for duplex DNA structures that appeared at the lower part of the gel (Figure 3.22).

Figure 3.22: SSCP gels (PhastGel® Homogeneous 12.5) run at 4ºC showing labelled (i) and non-labelled (ii) double stranded and single stranded DNA fragments used for the formation of heteroduplex DNA constructs containing IDLs

(i)



(ii)



i) Labelled DNA fragments:
a- Short PCR product (561bp)
b- Mixture of short + long strands (d & e)
c- Long PCR product (585bp)
d- Short single strands (561 bases)
e- Long single strands (585 bases)

ii) Non-labelled DNA fragments:
a- Short PCR product (561bp)
b- Short single strands (561 bases)
c- Mixture of short + long strands (b & d)
d- Long single strands (585 bases)
e- Long PCR product (585bp)

The blue arrows point at the duplex DNA structures that were only visible on gel (ii) for the non-labelled PCR products.

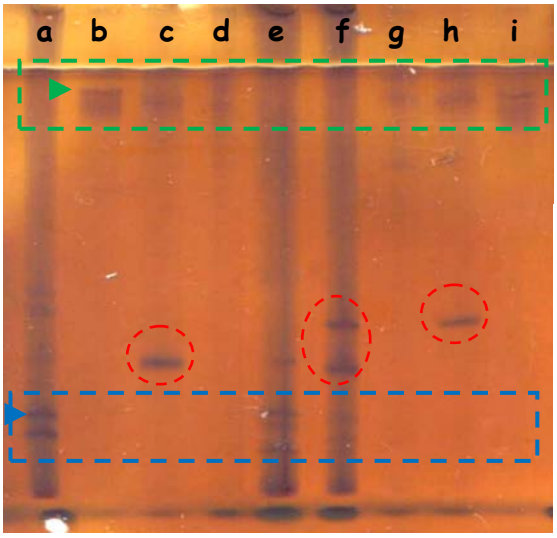
Detection of the non-labelled 328bp A.T and G.C PCR products and their individual single strands (A, T, G and C) was achieved at a total run time of 350Vh on the PhastGel® Homogeneous 20 gels at 4ºC with an extended pre-run (Figure 3.25 below).

3.2.1.4 Hybridisation of single strands to form heteroduplex DNA molecules

Hybridisation of ssDNA for the formation of heteroduplex constructs is described in section 2.6.2.3. The heteroduplexes were detected on SSCP gels and the ABI Prism™ 310. Visualising the hetroduplexes on SSCP gels was crucial to confirm the formation of IDL (secondary structures) when the ssDNA fragments were mixed together. The duplex DNA molecules (in PCR products) and the single strands with a 24bp size difference were detected on the Homogeneous 12.5 PhastGel® at 4ºC. However, the heteroduplexes formed after mixing the ssDNA fragments could not be seen initially, when the single strands were mixed at room tempeature. An overnight incubation at 37ºC proved to be necessary for the heteroduplex molecules to form.

The observation of heteroduplex molecules on SSCP gels confirmed the hybridisation of the ssDNA fragments and the formation of a 24-nucleotide (24-nt) IDL structure. Figure 3.23 displays an SSCP gel showing distinctive electrophoretic migrations of ssDNA fragments, homoduplex and heteroduplex DNA molecules. The electrophoretic migration of mismatched heteroduplex DNA was retarded relatively to homoduplex DNA due to the distorted DNA conformation.

Figure 3.23: PhastGel® Homogeneous 12.5 run at 4ºC showing the ssDNA fragments, homoduplex DNA molecules (PCR products) and heteroduplex DNA molecules with 24-nt IDLs

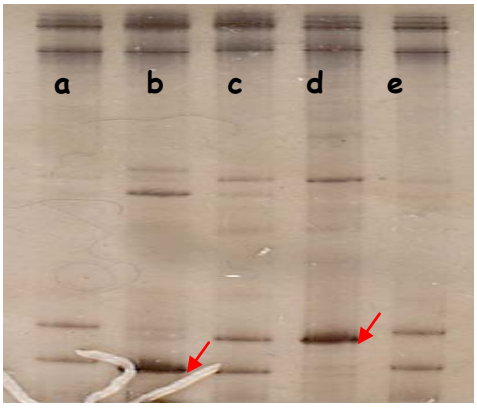


a- Homoduplex (585bp PCR product)
b- ssDNA 585 bases
c- 24-nt IDL heteroduplex (585/561 b+d)
d- ssDNA 561 bases
e- Homoduplex (561bp PCR product)
f- Mix of 2 PCR products (a + e)
g- ssDNA 585 bases (complementary to b)
h- 24-nt IDL heteroduplex (585/561 g+i)
i- ssDNA 561 bases (complementary to d)

The green and blue boxes show the ssDNA fragments and homoduplex DNA molecules (PCR products), respectively. The heteroduplex DNA molecules with a 24-nt IDL on one of the strands are circled in red. The heteroduplex DNA constructs were prepared from isolated single strands, which were mixed and left to hybridise overnight at 37ºC.

The electrophoretic migration of mismatched heteroduplex DNA constructs with a smaller IDL of only 3 nucleotides could not be easily distinguished from the homoduplex DNAs on neither Homogeneous 12.5 nor 20 gels, even after SSCP optimisations. Figure 3.24 shows the best conditions obtained with these samples (Homogeneous 20 run at 15ºC with an extended pre-run). The heteroduplex structures seemed to migrate to a level between the two levels reached by the homoduplex molecules.

Figure 3.24: PhastGel® Homogeneous 20 run at 15ºC showing the electrophoretic migrations of the heteroduplex DNA constructs with a 3-nt IDL compared to homoduplex molecules

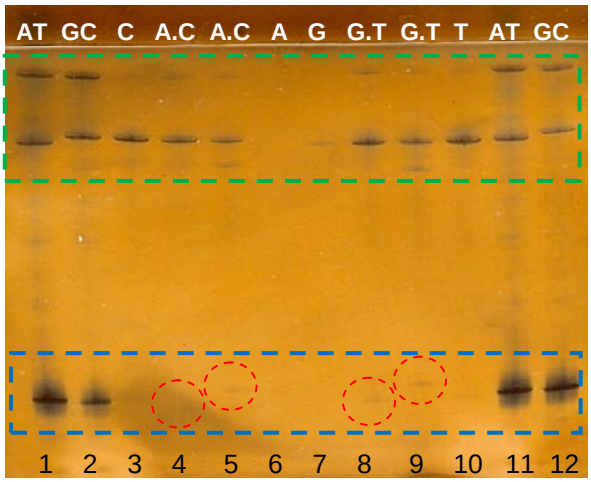


a- Homoduplex (582bp)
b- Heteroduplex (582/585)
c- Homoduplex (585bp)
d- Heteroduplex (582/585) made from opposite strands to ones mixed in sample b
e- Homoduplex (582bp) (same as sample a)

The red arrows designate the bands corresponding to the heteroduplex molecules (b & d). The heteroduplex structures migrated to a level between the two levels reached by the homoduplex molecules.

Similarly, the G.T heteroduplexes were formed from ssDNA fragments isolated from PCR products (prepared from primers listed in Table 2.3) or from commercially synthesised oligonucleotide strands (180 bases long; Table 2.9). The G.T heteroduplex constructs were visualised on SSCP gels after an overnight incubation at 37ºC (Figure 3.25). The G.T or A.C heteroduplex molecule travelled slightly further than the homoduplex molecules and the nicked heteroduplex molecule migrated slower.

Figure 3.25: PhastGel® Homogeneous 20 run at 4ºC showing ssDNA fragments, homoduplex and heteroduplex DNA molecules with an A.C or G.T mismatch (at rs1981929 SNP site)

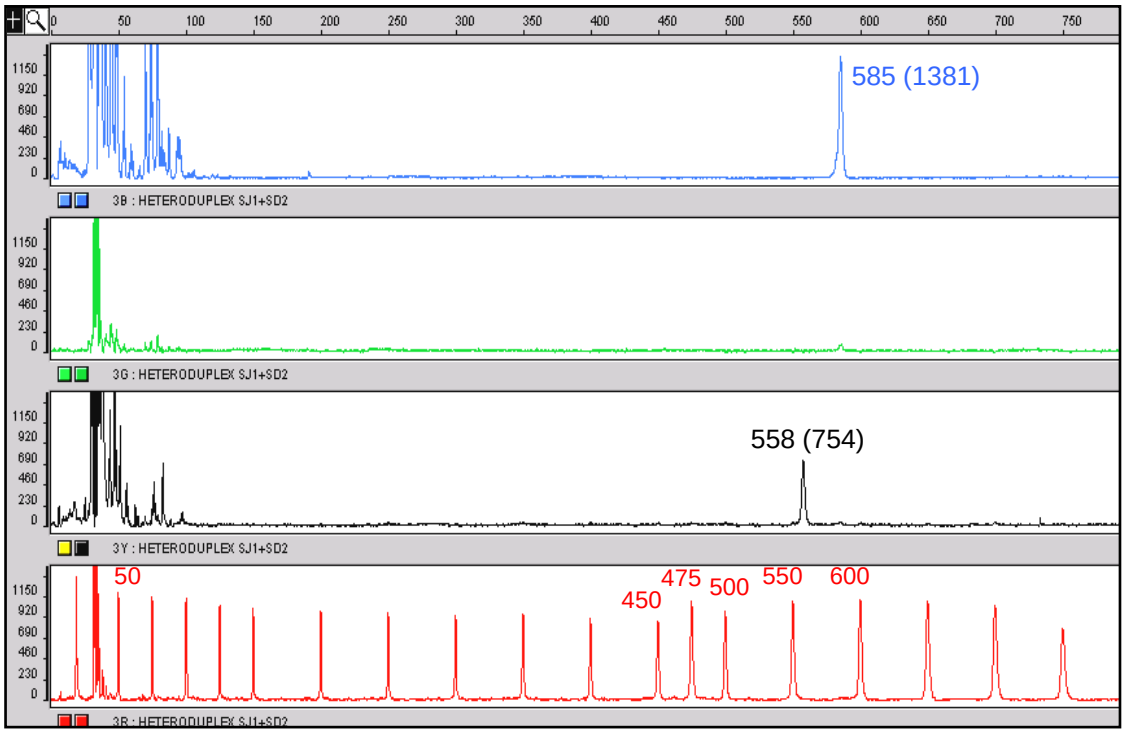


1- A/T homoduplex
2- G/C homoduplex
3- C strand
4- A.C heteroduplex (mixed C & A strands)
5- Nicked A.C heteroduplex
6- A strand
7- G strand
8- Nicked G.T heteroduplex
9- G.T heteroduplex
10- T strand
11- A/T homoduplex
12- G/C homoduplex

The green and blue boxes show the ssDNA fragments and homoduplex DNA molecules (328bp PCR products), respectively. The heteroduplex DNA molecules with A.C or G.T mismatches are circled in red. The heteroduplex constructs were prepared from isolated single strands, which were left to hybridise overnight at 37ºC. The heteroduplex constructs were nicked by exposing the samples to the Nt.BstNBI endonuclease.

The heteroduplex molecules with both strands labelled with the fluorescent dyes FAM and HEX were observed and quantified using the ABI Prism™ 310 genetic analyzer (Figure 3.26).

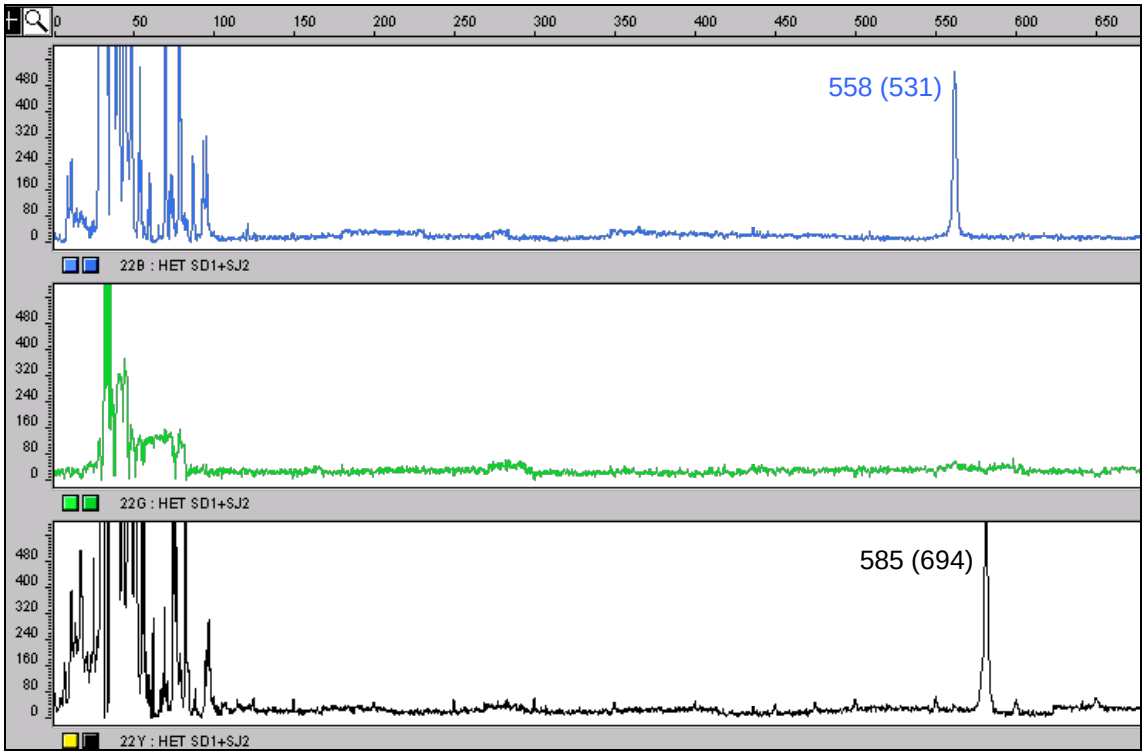
Figure 3.26: GeneScan™ analysis panels showing the double labelled heteroduplex construct



The bottom panel shows the TAMRA labelled size standard peaks (red). In the top panel, the blue peak represents the FAM labelled long strand of the heteroduplex molecule. In the third panel, the black peak represents the HEX labelled short strand.

In order to eliminate the variable absorbance sensitivities of different fluorescent molecules, the peak heights were normalised using double labelled PCR products produced by using the MMR1FAM and MMR2HEX primers together (section 2.6.4.1.1 in methods). The FAM/HEX peak height ratio was less than one and seemed to be constant between the different samples run together or on separate days. The average value of FAM/HEX was 0.80 (the values ranged between 0.76 and 0.84). This means that in the sample shown in Figure 3.26 there was an excess of FAM labelled ssDNA, which was not in heteroduplex form. In order to remove the free single strands, the mixture of the ssDNA was sequentially exposed to the two combinations of Dynabeads® with the biotinylated strands attached to them. This method eliminated all excess ssDNA leaving only heteroduplex DNA molecules in the sample (Figure 3.27). This was confirmed by the changed FAM/HEX peaks ratio of the sample, which was equal to 0.77.

Figure 3.27: GeneScan™ analysis panels showing the double labelled heteroduplex DNA construct exposed to Dynabeads® to remove the excess single strands

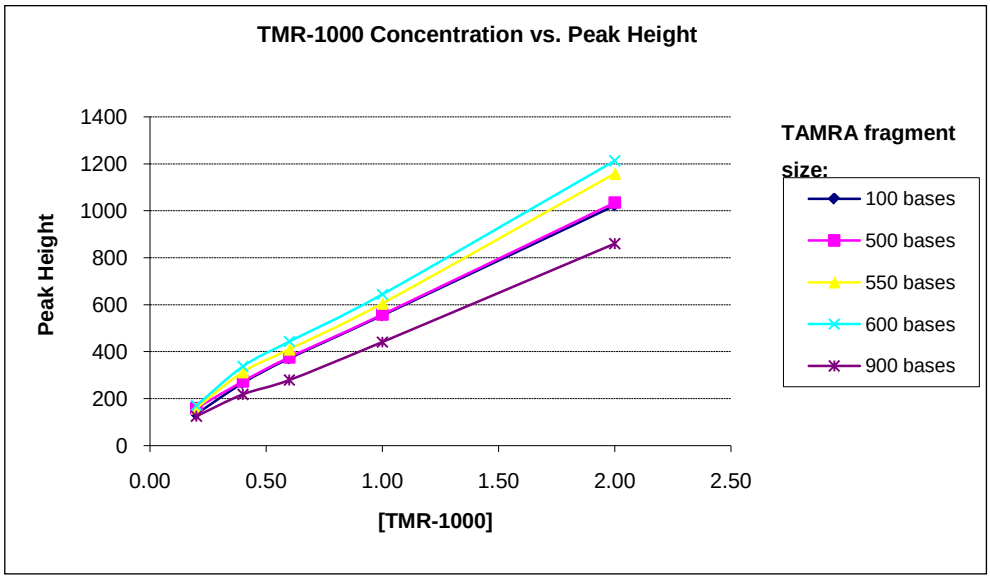


All DNA fragments in the sample shown were in heteroduplex form, as the amount of ssDNA labelled with FAM was equal to that labelled with HEX. This was determined from the FAM/HEX peaks ratio that was equal to 0.77 and matched the absorbance ratio of the two dyes on the ABI Prism™ 310.

The use of the FAM/HEX absorbance ratio allowed the selection of the ssDNA samples to mix together and the determination of the necessary dilution factor for the sample of higher concentration in order to avoid having a large excess of one strand during the hybridisation step and encourage heteroduplex formation. During the formation of FAM double labelled constructs, the peak heights alone were sufficient to match the concentrations of the two strands. If the ssDNA samples were processed on different dates the FAM/TAMRA ratio allowed the comparison of the sample concentrations, as a specific concentration of the TAMRA size standard was present in every sample.

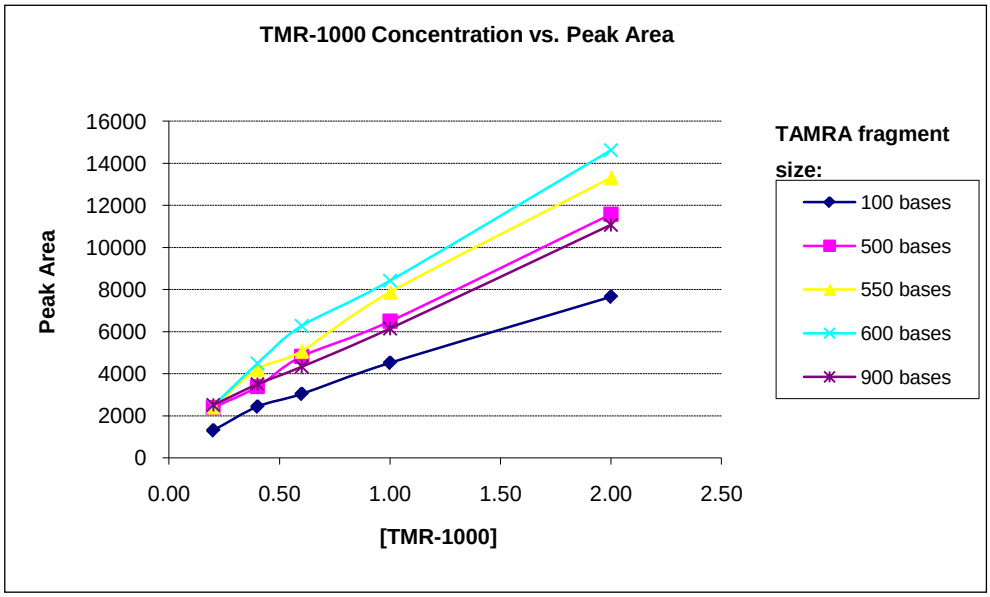
The Prism™ analysis was made to be fully quantitative by running a serial dilution of MapMarker® 1000 TAMRA with each Prism™ run and plotting a standard dose response curve for that run (Section 2.6.4.1.2 in methods). The peak height of the FAM or HEX labelled sample was then multiplied by the TAMRA/FAM or TAMRA/HEX absorbance ratio, which were 0.63 and 0.42, respectively, and the concentration of the sample was extrapolated from the standard curve (Figures 3.28 and 3.29). This method was used to estimate the concentration of heteroduplex DNA samples before exposure to the nicking enzyme or nuclear extracts.

Figure 3.28: Standard curve relating DNA fragment peak height to TAMRA concentration



A serial dilution of the TAMRA-1000 size standard was run on the ABI Prism™ 310. The peak heights obtained for each fragment size were plotted against the concentrations. The resulting standard curves were used for the estimation of samples that were run on the prism. A TAMRA-1000 concentration of 1 was equivalent to 0.3fmol (for each fragment size).

Figure 3.29: Standard curve relating DNA fragment peak area to TAMRA concentration



A serial dilution of the TAMRA-1000 size standard was run on the ABI Prism™ 310. The peak areas obtained for each fragment size were plotted against the concentrations. The resulting standard curves were used for the estimation of samples that were run on the prism. A TAMRA-1000 concentration of 1 was equivalent to 0.3fmol (for each fragment size).

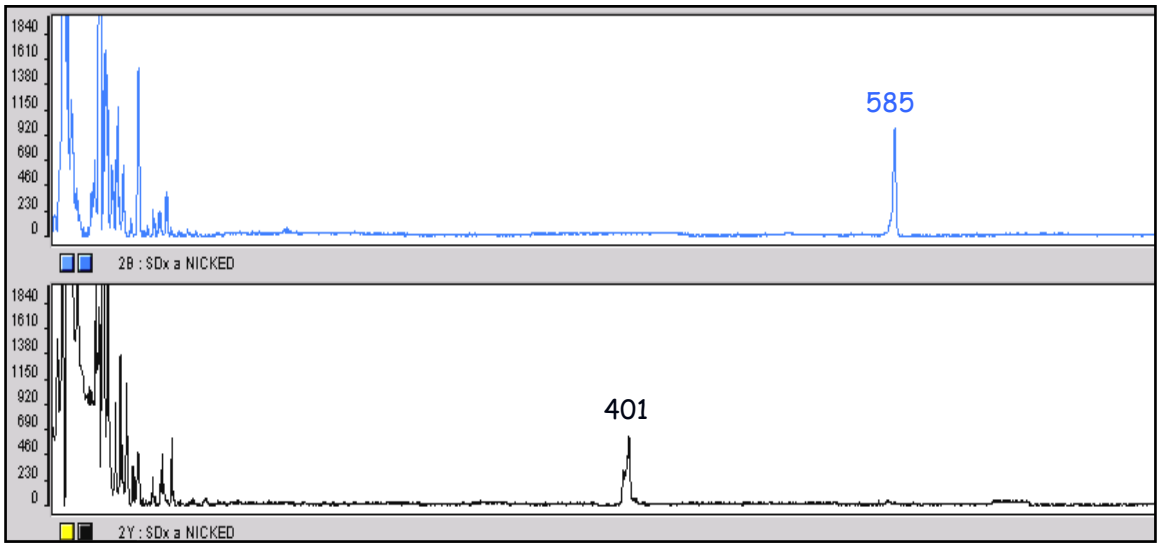
Peak heights ratios were adopted for quantitative analysis (not peak areas) as they were better correlated with DNA concentration and the fragment size had little effect on the standard curve.

3.2.1.5 Nicking of the heteroduplex DNA constructs

The heteroduplex DNA molecules were exposed to the nicking enzymes Nt. Bbv CI (for IDL constructs) and Nt.BstNBI (for constructs with G.T mismatches). This created a nick on one strand, leaving the other unchanged. The creation of a nick was visible on both the ABI Prism™ 310 (for double labelled heteroduplexes) and SSCP analysis (for non-labelled heteroduplexes), confirming the formation of the nicked heteroduplex construct.

Figure 3.30 shows the nicked heteroduplex molecule. The FAM labelled strand (shown in blue) was not changed. The HEX labelled strand (shown in black) was nicked and separated into two fragments upon denaturation of the DNA molecule. The fragment at the 5' end (401bp in size) remained labelled and was detected on the ABI Prism™ 310. The other fragment (at the 3' end) could not be detected on the Prism™, as it was not labelled anymore.

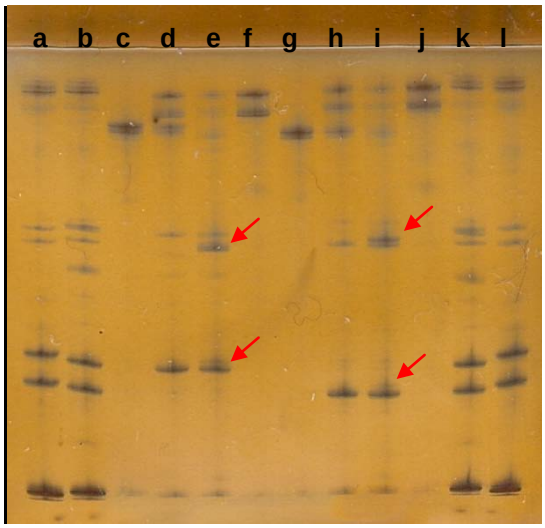
Figure 3.30: GeneScan™ analysis panels showing the nicked heteroduplex construct



The HEX labelled strand was nicked and resulted in two fragments after the duplex DNA molecule was denatured. The fragment at the 5' end was 401bp in size (black peak) and was detected on the ABI Prism™ 310. The other fragment could not be detected as it had lost the 5' HEX label. The FAM labelled strand was unaffected by the nicking endonuclease as shown by the blue peak (585 bases).

Figure 3.31 shows the nicked heteroduplex molecules with the 3-nucleotide IDLs on a Homogeneous 20 gel (long pre-run, 350Vh, 4ºC). The nicked heteroduplex DNA molecules, in lanes e and i, were distinguished from the homoduplex DNA molecules and non-nicked heteroduplex molecules. A similar observation was obtained with nicked A.C and G.T heteroduplex molecules (Figure 3.25).

Figure 3.31: PhastGel® Homogeneous 20 run at 4ºC showing the nicked and non-nicked heteroduplex DNA molecules with 3-nt IDL



a- Homoduplex (582bp)
b- Homoduplex (585bp)
c- ssDNA 585 bases
d- Heteroduplex construct (582/585 bases)
e- Nicked heteroduplex
f- ssDNA 582 bases
g- ssDNA 582 bases (complementary to f)
h- Heteroduplex construct (582/585 bases) from g & i
i- Nicked heteroduplex construct
j- ssDNA 585 bases (complementary to c)
k- Homoduplex (585bp)
l- Homoduplex (582bp)

The heteroduplex constructs were nicked with the Nt. BbvCIB endonuclease.
The red arrows indicate the nicked heteroduplex DNA constructs with 3-nt IDL.

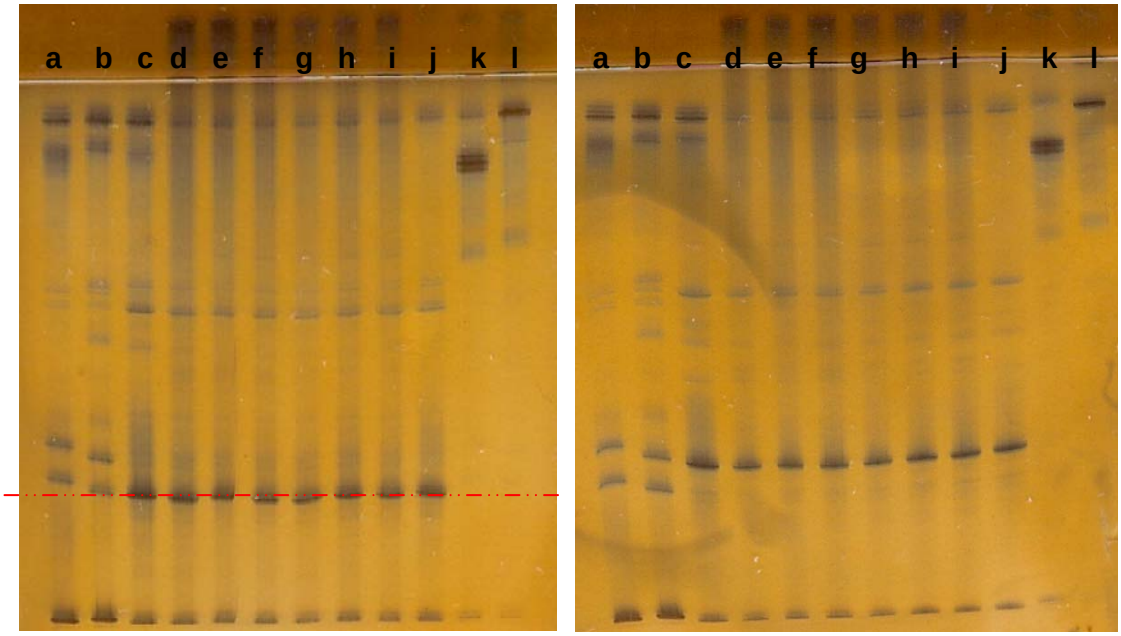
3.2.2 Exposure of heteroduplex constructs to nuclear extracts and repair assessment

In order to avoid loss of DNA templates affecting assessment of repair, all heteroduplex constructs were checked on the ABI Prism^{TM} 310 or by SSCP/Heteroduplex analysis before exposure to nuclear extracts (N.E.). Additionally, the samples were incubated at 37ºC overnight to make sure all DNA molecules were in duplex form rather than fragments of ssDNA prior to repair assessment.

Different conditions were used when exposing the heteroduplex DNA constructs with a 3, 21 or 24-nucleotide IDL or a G.T/C.A mismatch to nuclear extracts from HeLa S3 and LoVo cells (section 2.6.3). The different parameters involved the amount of N.E. (2µg and 20µg) and exposure time (30 seconds, 60 seconds, 5, 15, 30 and 60 minutes and overnight). Exposure to 2µg/20µg HeLa S3 or LoVo nuclear extracts for 60 minutes or overnight did not degrade any of the DNA constructs (Figures 3.32, 3.33, 3.35 and 3.36).

Changes in electrophoretic migration were consistently observed for the 3-nt heteroduplex samples exposed to HeLa S3 N.E. compared to the original heteroduplex sample and the reaction's negative control (Figure 3.32). This was indicative of repair as the shift in migration matched the homoduplex counterpart (Figure 3.33).

Figure 3.32: PhastGel® Homogeneous 20 run at 4ºC showing 3-nt IDL heteroduplex constructs with the insertion loop on opposite strands (for left and right gel) exposed to 2µg HeLa S3 and LoVo N.E. for 15, 30 and 60 minutes

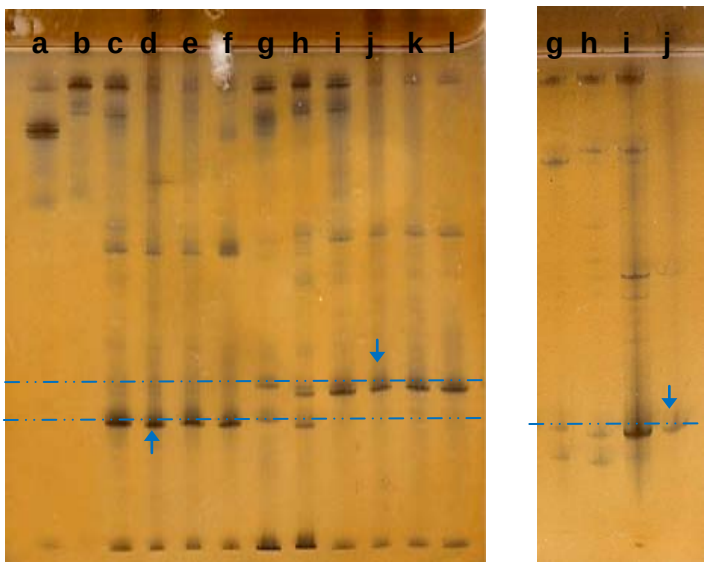


a- Homoduplex (298bp)
b- Homoduplex (301bp)
c- Heteroduplex (298/301) with 3-nt IDL
d- Heteroduplex + HeLa S3 N.E. for 15 min
e- Heteroduplex + HeLa S3 N.E. for 30 min
f- Heteroduplex + HeLa S3 N.E. for 60 min
g- Heteroduplex + LoVo N.E. for 15 min
h- Heteroduplex + LoVo N.E. for 30min
i- Heteroduplex + LoVo N.E. for 60min
j- Heteroduplex - N.E. (negative control: exposed to all reagents except N.E.)
k- 298 base ssDNA
l- 301 base ssDNA
The red line indicates level of heteroduplex DNA structures.
The heteroduplex sample exposed to 2µg HeLa S3 N.E. for 60min (lane f) showed the most noticeable shift in migration. Samples c and j, representing the pure heteroduplex construct and the construct exposed to the control solution without nuclear extracts, had the same migration levels. The 3-nt IDL heteroduplex constructs run on the left and right gels had the insertion loop on opposite strands. The 3-nt IDL was on strand 1 (forward strand) for the right gel and on strand 2 (reverse strand) for the left gel. Due to electrophoresis problems (smiling), it was not possible to assess repair on this gel.

When the IDL was on strand 1 (forward strand), exposure to N.E. resulted in the formation of a homoduplex that was equal to the size of the shorter strand (i.e. the 298bp homoduplex resulted from the 298/301 bases heteroduplex molecules). When the IDL was on the opposite strand, exposure to N.E. resulted in the formation of a homoduplex that was equal to the size of the longer strand (301bp) (Figure 3.33).

Therefore, IDL repair could target either the long or short strand in the absence of a nick (Figure 3.34).

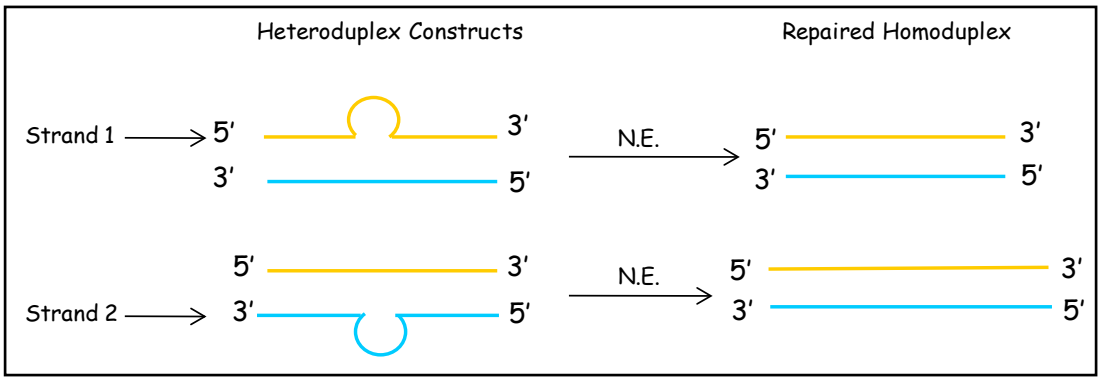
Figure 3.33: PhastGel® Homogeneous 20 run at 4ºC showing 3-nt IDL heteroduplex constructs exposed to 2µg HeLa S3 and LoVo N.E. for 60 minutes



a- 298 base ssDNA
b- 301 base ssDNA
c- Heteroduplex with 3-nt IDL on strand 2
d- Heteroduplex + HeLa S3 N.E.
e- Heteroduplex + LoVo N.E.
f- Heteroduplex - N.E. (negative control: exposed to all reagents except N.E.)
g- Homoduplex (298bp)
h- Homoduplex (301bp)
i- Heteroduplex with 3-nt IDL on strand 1
j- Heteroduplex + HeLa S3 N.E. for 60 min
k- Heteroduplex + LoVo N.E. for 60 min
l- Heteroduplex - N.E. (negative control: exposed to all reagents except N.E.)

The blue lines designate level of homoduplex DNA structures and the blue arrows indicate repair. The heteroduplex sample (d), which contained the IDL loop on strand 2 and was exposed to HeLa S3 N.E., travelled slightly further down the gel and had an electrophoretic migration that matched that of the 301bp homoduplex sample (h). The heteroduplexes (j and k), which contained the IDL loop on strand 1 and were exposed to N.E., travelled less far down the gel than the unexposed corresponding heteroduplex samples and had an electrophoretic migration that matched that of the 298bp homoduplex sample (g).

Figure 3.34: Diagram showing the direction of repair of heteroduplex constructs with IDLs

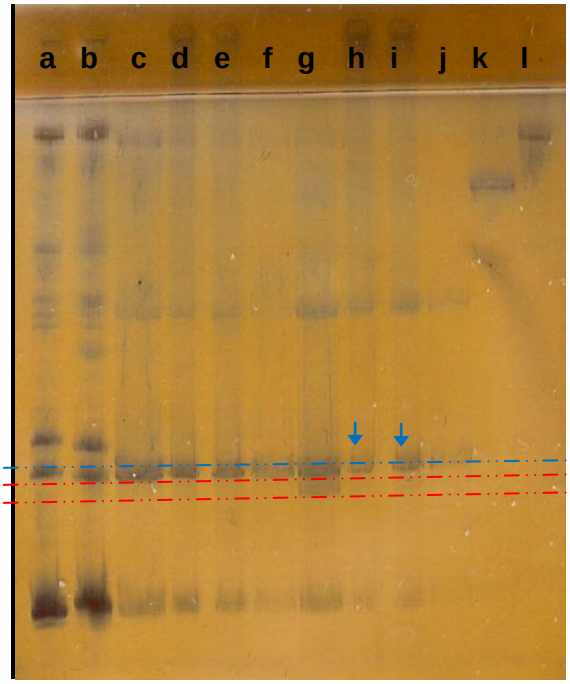


Strand 1 was shortened to match the size of strand 2 when the IDL was on strand 1.
strand 1 was elongated to match the size of strand 2 when the IDL was on strand 2.

Post exposure to N.E., the nicked heteroduplex constructs with a 3-nt IDL loop showed a shift in electrophoretic migration that matched that of the shorter homoduplex sample (298bp). The nicked heteroduplex constructs could not be seen post exposure to N.E.. From the results shown in Figure 3.35, it was deduced that repair (detected in samples h and i) was directed towards the long (301 bases) strand to give the shorter (298bp) homoduplex as samples h and i had a migration pattern that was similar to the 298bp PCR product. The nick was on the 301-base strand; therefore, in presence of a 5' nick repair was nick directed.

There was a consistent observation in two sets of repair reactions which did not show repair of 3-nt or 24-nt IDLs on heteroduplex analysis. A new band that did not correspond to either homoduplex or heteroduplex DNA migration levels was detected (Figure 3.36). These DNA fragments may be repair intermediates in which excision of the loop or part of the opposite strand took place but resynthesis was not completed, presumably due to insufficient amounts of dNTPs or lack of necessary enzymes.

Figure 3.35: PhastGel® Homogeneous 20 run at 4ºC showing nicked and non-nicked 3-nt IDL heteroduplex constructs exposed to 2µg HeLa S3 and LoVo N.E. for 15 minutes



a- Homoduplex (298bp)
b- Homoduplex (301bp)
c- Heteroduplex construct with 3-nt IDL on strand 2
d- Heteroduplex + HeLa S3 N.E.
e- Heteroduplex + LoVo N.E.
f- Heteroduplex - N.E.
g- Nicked heteroduplex (nick 5' to IDL on strand 2)
h- Nicked heteroduplex + HeLa S3 N.E.
i- Nicked heteroduplex + LoVo N.E.
j- Nicked heteroduplex - N.E.
k- 298 base ssDNA
l- 301 base ssDNA

The red lines indicate levels of heteroduplex DNA structures (top: non-nicked; lower: nicked). The blue line designates the level of the 298bp homoduplex DNA structures and the blue arrows indicate possible repair. The nicked heteroduplex samples (h and i), which contained the IDL loop and nick on strand 2 and were exposed to N.E., travelled less far down the gel compared to the unexposed heteroduplex sample and had an electrophoretic migration that matched that of the 298bp homoduplex sample (a). Therefore, repair was directed towards the 301-base strand in samples h and i.

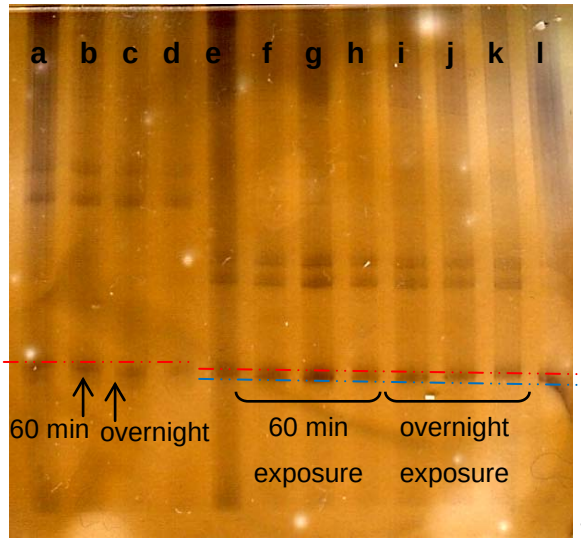
Figure 3.36: PhastGel® Homogeneous 20 run at 4ºC showing non-nicked 3 and 24-nt IDL heteroduplex constructs exposed to 2µg HeLa S3 and LoVo N.E. for 60 minutes

a- Heteroduplex (582/585 bases) with 3-nt IDL on strand 1
b- 3-nt IDL heteroduplex + HeLa S3 N.E.
c- 3-nt IDL heteroduplex - N.E.
d- Heteroduplex (561/585 bases) with 24-nt IDL on strand 1
e- 24-nt IDL heteroduplex + HeLa S3 N.E.
f- 24-nt IDL heteroduplex - N.E.
g- Heteroduplex (561/585 bases) with 24-nt IDL on strand 2
h- 24-nt IDL heteroduplex + HeLa S3 N.E.
i- 24-nt IDL heteroduplex + LoVo S3 N.E.
j- 24-nt IDL heteroduplex - N.E.
k- Mix of complementary ssDNA fragments 561bases
l- Mix of complementary ssDNA fragments 585bases

The purple circles indicate bands that appeared in samples exposed to N.E. and did not correspond to either homoduplex or heteroduplex DNA band levels.

Increasing the amount of nuclear extract from 2µg to 20µg did not improve IDL repair detection. Furthermore, repair seemed to be proportional to exposure time to nuclear extract, possibly up to 60 minutes. The repair observed after overnight exposures to N.E. was similar to the results obtained for 60 minute exposures (Figure 3.37).

Figure 3.37: PhastGel® Homogeneous 20 run at 4ºC showing non-nicked 3 and 24-nt IDL heteroduplex constructs exposed to 2µg HeLa S3 and LoVo N.E. for 60 minutes and overnight

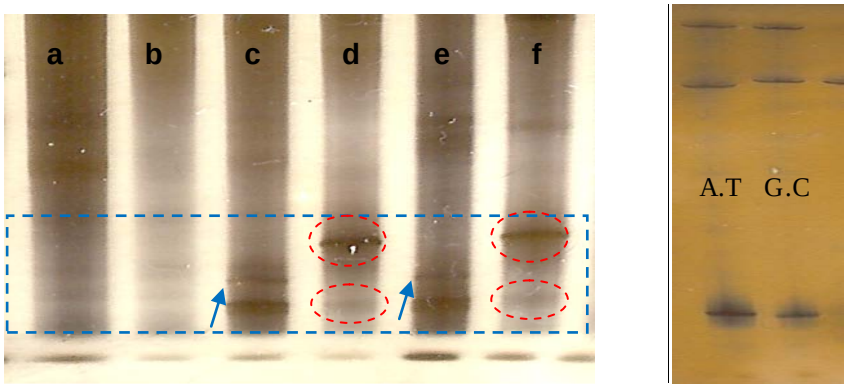


a- Heteroduplex construct (561/585 bases) with 24-nt IDL on strand 2
b- 24-nt IDL heteroduplex + HeLa S3 N.E. overnight
c- 24-nt IDL heteroduplex + LoVo N.E. overnight
d- 24-nt IDL heteroduplex - N.E. overnight
e- Heteroduplex construct (582/585bases) with 3-nt IDL on strand 2
f- 3-nt IDL heteroduplex + HeLa S3 N.E. for 60 min
g- 3-nt IDL heteroduplex + LoVo N.E. for 60 min
h- 3-nt IDL heteroduplex - N.E. for 60 min
i- 3-nt IDL heteroduplex + HeLa S3 N.E. overnight
j- 3-nt IDL heteroduplex + LoVo overnight
k- 3-nt IDL heteroduplex - N.E. overnight
l- Mix of complementary ssDNA fragments 582 bases

The red lines indicate levels of heteroduplex molecules and the blue line designates the expected level of the 582bp homoduplex DNA structures. Repair was detected in all samples exposed to nuclear extracts.

Following exposure to nuclear extracts, repair of single base mismatches was assessed by heteroduplex analysis. Repair was detected after exposure of nicked G.T heteroduplex constructs to 2µg HeLa S3 N.E. for 60 minutes (Figure 3.38). As the T strand was nicked, MMR should have resulted in G.C homoduplexes. No G.C homoduplex samples were run on the same gel; however, another gel run at the same conditions was used to compare the migration levels of G.C and A.T homoduplex samples. The G.C homoduplexes seemed to run slightly further than A.T homoduplexes. It was thus deduced that the exposed heteroduplex constructs (samples d and f in Figure 3.38) resulted in G.C homoduplexes (samples c and e in Figure 3.38); this means that in this situation repair of the G.T mismatch was nick directed. Sample c showed both A.T and C.G homoduplexes; this could have resulted from repair of the non-nicked substrates that can target either strands. As seen from the band intensities, the substrate (d) contained more nicked than non-nicked constructs and more G.C than A.T homoduplexes were produced as a result of nick directed repair.

Figure 3.38: PhastGel® Homogeneous 20 run at 4ºC (with extended pre-run) showing nicked G.T heteroduplex constructs exposed to 2µg HeLa S3 N.E. for 60 minutes



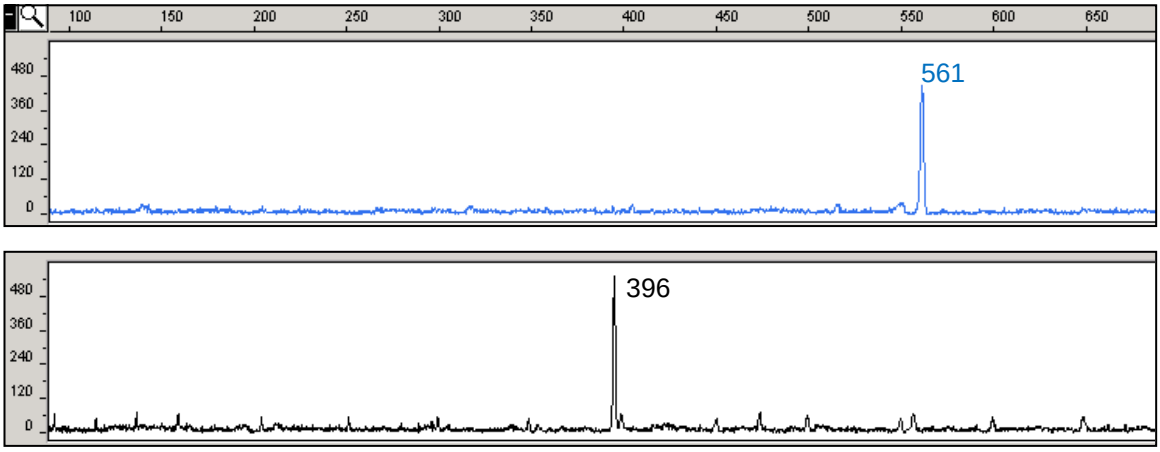
a- A/T homoduplex + HeLa S3 N.E.
b- A/T homoduplex - N.E.
c- Nicked G.T heteroduplex + HeLa S3 N.E.
d- Nicked G.T heteroduplex - N.E.
e- Nicked G.T heteroduplex + HeLa S3 N.E.
f- Nicked G.T heteroduplex - N.E.
The blue box highlights the region containing the dsDNA molecules and the heteroduplex constructs are circled in red. The top heteroduplex constructs are nicked and the lower ones are non-nicked (nicking reaction efficiency <100%). Bands at the top region of the gel represent single strands. Samples in lanes a & b, which are homoduplex A.T constructs, had weak bands which was probably due to suboptimal duplex formation because of unequal concentrations of individual DNA strands. The gel on the right shows the homoduplex A.T and G.C samples (328bp PCR products). The blue arrows indicate repair. Sealing of the nick was ~100% efficient as in both nicked G.T heteroduplex samples exposed to N.E. (c and e) the band intensity of the non-nicked heteroduplex molecules was increased compared to the starting substrates (d and f) and none of the nicked templates could be detected post exposure to nuclear extract. Exposure of the heteroduplex substrates to N.E. resulted in the formation of both A.T and G.C homoduplexes in sample c, showing that repair can be nick-independent.

Sealing of the nick as a result of ligation without repair was detected in samples c and e in Figure 3.38, as the intensities of the bands corresponding to the non-nicked heteroduplexes (samples d and f) were increased after exposure to extracts. Heteroduplex analysis on SSCP gels could only show the presence or absence of a formed homoduplex, which indicated repair. In order to assess and compare repair efficiency, semi-quantitative analysis was used on the ABI Prism™ 310.

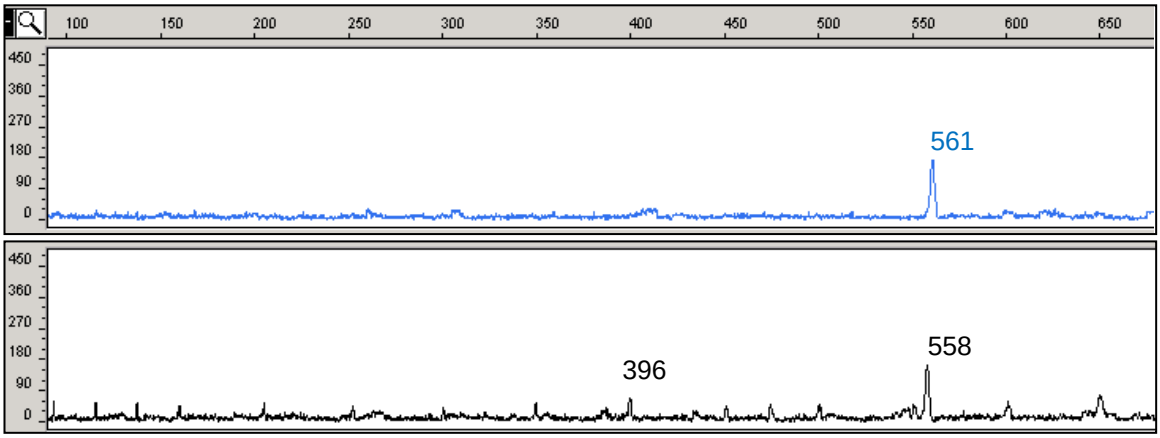
DNA repair was initially detected on the ABI Prism™ 310 for double-labelled nicked heteroduplexes containing a 24-nt IDL after exposure to 2µg of HeLa S3 N.E. for 15 minutes. The formation of a HEX-labelled DNA fragment, which was the size of the complementary non-nicked strand, confirmed repair (Figure 3.39).

Figure 3.39: GeneScan™ analysis panels showing the nicked heteroduplex DNA before and after exposure to nuclear extracts for 15 minutes

1) Nicked heteroduplex construct with 24-nt IDL (nick on long strand)



2) Nicked heteroduplex construct after exposure to 2µg HeLa S3 N.E. for 15 minutes

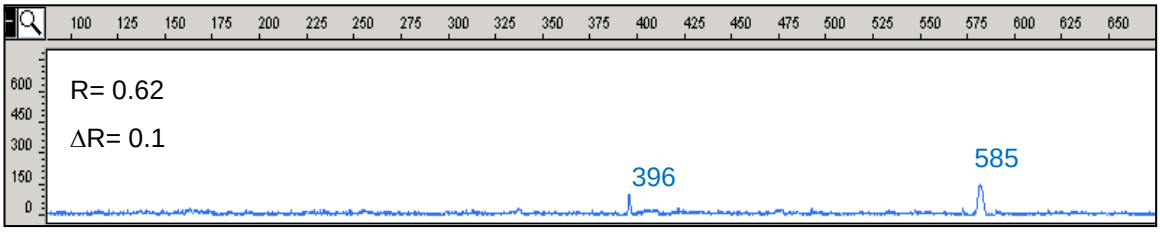


The appearance of a new Hex-labelled (black) peak corresponding to 558 bases (which is equivalent to 561 on the opposite strand) was indicative of IDL repair. Some of the nicked fragments (396 base peak) were converted to 561 bases sized fragments using the non-nicked FAM-labelled (blue) strand as a template.

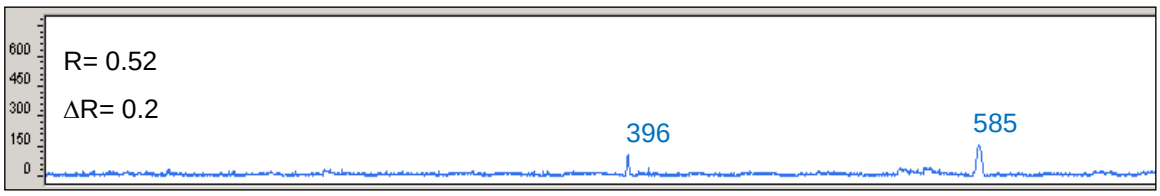
Semi-quantitative fluorescence analysis on the ABI Prism™ 310 (section 2.6.4.1.3) was used to determine the ratios of fluorescence from the nicked or short strand and the complementary non-nicked or long strand (R = peak height of nicked or short strand/peak height of non-nicked or long strand). IDL repair efficiency was assessed by comparing the R-value before and after exposure of the construct to nuclear extracts, $\Delta R = R_{\text{negative control}} - R_{\text{sample exposed to N.E.}}$ (Figure 3.40). This type of analysis helped overcome the issue of incomplete nicking of the heteroduplex constructs. In order to eliminate the issue of the genetic analyzer's sensitivities to different labels and to simplify the repair assessment on the ABI Prism™ 310, the heteroduplex DNA molecules were prepared with both strands FAM-labelled.

Figure 3.40: GeneScan™ analysis showing the nicked heteroduplex constructs containing a 24-nt IDL after exposure to nuclear extracts for 30 and 60 seconds

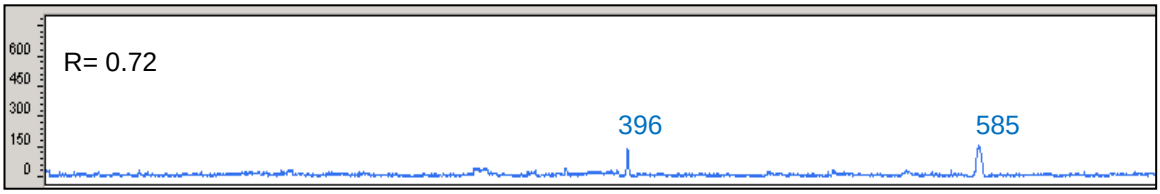
1) Nicked heteroduplex construct after 30s exposure to HeLa S3 nuclear extracts



2) Nicked heteroduplex construct after 60s exposure to HeLa S3 nuclear extracts



3) Negative Control



R = nicked strand peak height/complementary strand peak height. The R-value was calculated for each sample along with the ΔR-values ($R_{negative\ control} - R_{sample\ exposed\ to\ N.E.}$), which showed repair efficiency. Exposure of heteroduplexes to 60 seconds resulted in greater repair efficiency compared to 30 seconds.

Exposure of the heteroduplex constructs to nuclear extracts resulted in changes in the R-value. This change was considered indicative of repair if two conditions were met:

1) $\Delta R\text{-value} \neq 0$

2) $R_{sample\ exposed\ to\ N.E.} - R_{unprocessed\ heteroduplex\ sample}$ was > 0.1, as the R-value can change after processing of the sample in the absence of nuclear extracts ($R_{negative\ control} - R_{unprocessed\ heteroduplex\ sample}$ was ≤ 0.1).

Table 3.13 lists the results obtained using the R-value system of analysis under all tested experimental conditions and Table 3.14 summarises the main findings.

Table 3.13: ΔR values following exposure of heteroduplex constructs with 3, 21 and 24-nt IDLs to HeLa S3 and LoVo nuclear extracts

Sample number	IDL size (nt)	Nicked strand	N.E. (µg)	Type of N.E.	Exposure time (min)	ΔR post N.E. exposure	Repair	Targeted strand
1	3	-	2	HeLa S3	0.5	0.01	No	-
2	3	long	2	HeLa S3	0.5	-0.04	No	-
3	3	long	20	HeLa S3	0.5	0.34	Yes	long
4	3	long	20	HeLa S3	0.5	0.11	Yes	long
5	3	short	20	HeLa S3	0.5	0.1	No	-
6	3	long	2	HeLa S3	1	0	No	-
7	3	short	2	HeLa S3	1	0	No	-
8	3	short	2	HeLa S3	1	0.03	No	-
9	3	short	20	HeLa S3	1	0.22	Yes	short
10	3	short	20	HeLa S3	1	0.2	Yes	short
11	3	short	2	HeLa S3	5	0.09	No	-
12	3	short	2	HeLa S3	5	0.05	No	-
13	3	-	2	HeLa S3	15	0.02	No	-
14	3	-	2	HeLa S3	15	0.01	No	-
15	3	short	2	HeLa S3	15	-0.06	No	-
16	3	short	20	HeLa S3	15	0.07	No	-
17	3	-	2	LoVo	0.5	0.01	No	-
18	3	long	2	LoVo	0.5	0	No	-
19	3	-	2	LoVo	15	0.01	No	-
20	3	short	2	LoVo	15	0.01	No	-
21	21	-	2	HeLa S3	0.5	0.05	No	-
22	21	long	2	HeLa S3	0.5	-0.13	Yes	short
23	21	long	20	HeLa S3	0.5	0.01	No	-
24	21	short	2	HeLa S3	1	0.07	No	-
25	21	-	2	HeLa S3	15	0.12	Yes	short
26	21	-	2	HeLa S3	15	-0.02	No	-
27	21	-	20	HeLa S3	15	0.3	Yes	short
28	21	-	2	LoVo	0.5	0.03	No	-
29	21	long	2	LoVo	0.5	-0.11	Yes	short
30	21	long	2	LoVo	15	-0.01	No	-
31	24	-	20	HeLa S3	0.5	0.09	No	-
32	24	short	20	HeLa S3	0.5	0.04	No	-
33	24	-	2	HeLa S3	15	-0.03	No	-
34	24	-	20	HeLa S3	15	0.29	Yes	short
35	24	short	2	HeLa S3	15	0.19	Yes	short
36	24	-	2	LoVo	15	0.05	No	-
37	24	-	2	LoVo	15	0.11	Yes	short
38	24	long	2	LoVo	15	0.01	No	-

$\Delta R = R_{\text{negative control}} - R_{\text{sample exposed to N.E.}}$, where R= short or nicked strand peak height/complementary strand peak height. ΔR values ≠ 0 were considered indicative of repair if $R_{\text{sample exposed to N.E.}} - R_{\text{unprocessed heteroduplex sample}}$ was > 0.1 as $R_{\text{negative control}} - R_{\text{unprocessed heteroduplex sample}}$ was ≤ 0.1. The R-values of all processed sample are listed in Table C.1 (Appendix).

Table 3.14: Summary of repair assessment results for 3, 21 and 24-nucleotide IDLs

IDL size (nt)	Type of repair	Minimum amount of N.E. resulting in detectable repair
3	Nick-dependent & directed MMR-dependent	20 µg HeLa
21	Nick & MMR-independent	2 µg LoVo/HeLa
24	Nick & MMR-independent	2 µg HeLa/LoVo

3.2.3 Summary of results for the *in vitro* IDL and mismatch repair functional assay

- Formation of heteroduplex DNA constructs with 3, 21 and 24-nucleotide IDLs or G.T/A.C mismatches was confirmed using SSCP/heteroduplex analysis.

- Exposure of heteroduplex constructs to nuclear extracts resulted in the formation of homoduplex molecules, which was indicative of repair. Extended overnight exposures to nuclear extracts did not degrade any of the substrates.

- Repair of the G.T mismatch was detected using SSCP/heteroduplex analysis. Repair of IDLs was detected on SSCP gels and on the ABI genetic analyzer.

- Changes in band migration levels on SSCP gels showed that repair could target either the long or short strand of the heteroduplex molecule in the absence of a nick.
    - 2 and 20µg of nuclear extracts gave similar results for exposures longer than 15 minutes; however, 60-minute exposures gave better results than shorter exposures to extracts.
    - Repair of 3-nt IDLs was nick directed when a 5' nick was introduced. Sealing of the nick, which may be due to ligation without repair, was observed in all nicked heteroduplex samples with a close to 100% efficiency; however, repair was not as efficient.

- Analysis on the genetic analyzer allowed semi-quantitative assessment of repair efficiency.
    - Preliminary results obtained using the ΔR value system showed that repair was independent of loop size (3, 21 and 24 nucleotides) and changes in the R-value could be detected for exposures to 2µg (0.08µg/µl) of nuclear extract for as little as 30 seconds.
    - Repair of 3-nt loops was MMR-dependent and nick-directed, whereas repair of larger loops was MMR and nick-independent.

3.3 Development of PGD protocols for *BRCA1* and *MSH2*

PGD workups for three couples that opted for PGD treatment for different *BRCA1* and *MSH2* mutations are described in this section. These couples had undergone genetic counselling and were referred to UCL Centre for PGD. The workup towards a PGD protocol for each couple started after they had undergone the initial PGD consultation and fertility checks at the ACU (UCLH).

The first step of each workup involved direct detection and confirmation of the mutation in the proband and his/her affected relatives and the selection of STR markers that are informative for the couple. The PCRs were carried out as described in Materials and Methods (section 2.4.2.3) using genomic DNA extracted from whole blood.

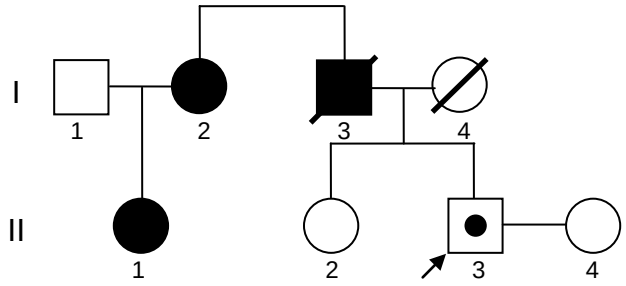
The second step involved the optimisation of a multiplex PCR that amplified the mutation site (when possible) and selected STR markers with genomic DNA. The multiplex PCR protocol was finally optimised using the couples' single lymphocytes in order to reach a protocol with the highest amplification efficiency and lowest allele dropout (ADO) rate achievable. Single-cell PCRs are described in Materials and Methods (section 2.4.2.3).

3.3.1 PGD workup for *MSH2* (c.1277-?_1386+?del)

3.3.1.1 Patient description

The male partner (proband) carried a deletion of exon 8 of the *MSH2* gene. The deletion breakpoints were unknown as the mutation was diagnosed by multiplex ligation-dependent probe amplification (MLPA). At age 40 the proband was asymptomatic but undergoing biannual colonoscopy. His wife was homozygous normal. His father carried the deletion and died from colorectal cancer. The proband's aunt and cousin also carry the germline deletion. Figure 3.41 shows the pedigree. Bloods for DNA extraction were available from the proband, his wife and his affected paternal aunt.

Figure 3.41: Pedigree showing the asymptomatic proband and his affected relatives



II.3 is the proband, carrier of the exon 8 deletion in the *MSH2* gene. I.2, I.3 and II.1 represent the proband's affected paternal aunt, father and cousin, respectively.

3.3.1.2 Informativity tests

As the deletion breakpoints were unknown, the PGD protocol had to rely on linkage analysis. Informativity tests were carried out on genomic DNA from the proband, his wife and paternal aunt. Eleven linked and thirteen unlinked STR markers were investigated using F-PCR followed by fragment length analysis on the ABI Prism™ (section 2.6.4.2). The allele sizes are listed in Table 3.15. Some of the markers were tested by Yasmin Omar and William J. Young as part of their MSc and BSc projects, respectively.

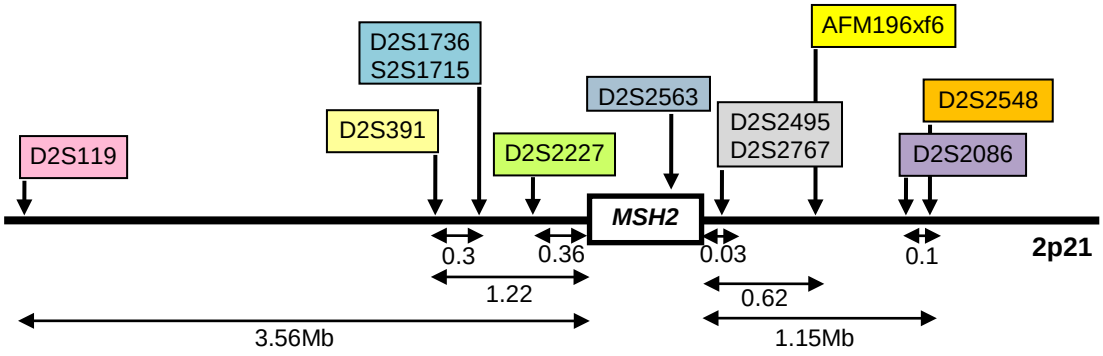
Table 3.15: Allele sizes for different STR markers identified for the couple undergoing PGD for the *MSH2* exon 8 deletion

Primer name	STR marker location	Female partner (bp)	Male partner (proband) (bp)	Affected relative (bp)	Informativity
Linked STR markers					
D2S119*	2p21	211/215	*211/218*	*211/221*	Semi-informative
D2S391	2p21	139	*141*	*141/147*	Not informative
D2S1736	2p21	203	*203*	*203*	Not informative
D2S1715E	2p21	206	*206*	*206*	Not informative
D2S2227*	2p21	151/153	*151/166*	*151/166*	Semi-informative
D2S2563	2p21	187/239	*239*	*239*	Not informative
D2S2495	2p21	205	*205*	*205*	Not informative
D2S2767	2p21	187	*187*	*187*	Not informative
AFM196xf6	2p21	133	*133*	*133*	Not informative
D2S2548	2p21	158/162	*158*	*158/160*	Not informative
D2S2086	2p21	148	*142*	*148*	Not informative
Unlinked STR markers					
D17S855	17q21.31	158	*156*	*151*	Not informative
D17S1185	17q21.31	206	*208*	*211*	Not informative
D17S1343	17q21.31	154/158	*154/158*	*154/158*	Not informative
D17S1338	17q21.31	180/181	*180/181*	*181/187*	Not informative
APOC2*	**19q13.32**	**149/155**	***151/157***	***128/136***	**Informative**
DM	19q13.32	122/170	*122/140*	*123/140*	Semi-informative
D19S112	19q13.32	134	*136*	*129/140*	Not informative
D13S168	13q14.2	141	*141*	*141/146*	Not informative
D13S262	13q14.2	289	*294*	*298*	Not informative
D17S1294	17q11.1	248/256	*244/248*	*248*	Semi-informative
D17S841	17q11.1	**260/268**	***264***	***264/266***	**Informative**
D17S1800	17q11.1	273	*266/274*	*274*	Not informative
IVS26-2.3	17q11.1	211/215	*211*	*211*	Not informative

The affected individuals are represented in italics. Bold text signifies that the alleles are informative and underlined text indicates the phase allele. The markers in blue are linked with the *MSH2* gene. The asterisk denotes the markers that were selected for use in the multiplex PCR reaction.

None of the linked markers tested were found to be fully informative for this couple. Two markers were semi-informative: D2S2227 and D2S119. The allele that was in phase with the mutation was identified for D2S119 as the affected male and his affected aunt shared the 211bp allele only. D2S2227 was approximately 0.36Mb away from *MSH2*, which made it a good candidate for linkage analysis, but D2S119 was >3.5Mb away from the gene. Figure 3.42 shows the locations of the linked markers on chromosome 2 relative to the *MSH2* gene.

Figure 3.42: *MSH2* gene and linked STR markers on chromosome 2

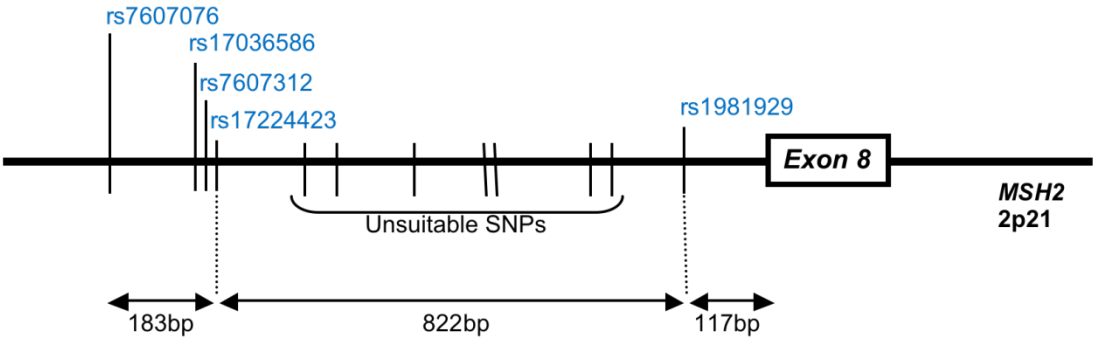


(Location of MSH2 on chromosome 2: 47,525,573–47,526,053)
The displayed distances between the *MSH2* gene and each STR marker are given in Mb. Only one of the markers was Intragenic: D2S2563. Seven out of the eleven linked markers investigated were located within 1Mb away from the gene.

Both the APOC2 and D17S841 unlinked markers were found to be informative as contamination markers as the female partner was heterozygous. However, as the male partner was homozygous for D17S841, the fully informative marker APOC2 was selected as the contamination marker to facilitate multiplex PCR optimisations. The affected male was heterozygous for the two semi-informative linked markers (D2S2227 and D2S119) and APOC2.

In addition to STR markers, SNPs were investigated in an attempt to find an informative SNP for the couple that was located within or in close proximity to exon 8. An informative SNP could be included in the multiplex PCR for PGD for linkage analysis using SSCP or mini-sequencing. If the SNP were within the deleted region, the affected male partner would have a single allele for that SNP. For the SNP to be fully informative, the female partner needed to be homozygous for the other allele. Two PCR primer sets were designed to target five SNPs up to1.5Kb upstream of exon 8. The targeted SNPs were selected based on their location in the *MSH2* gene with respect to exon 8 and their allele frequencies (ideally ~0.5 homozygous frequencies for either allele). Figure 3.43 illustrates the location of the SNPs investigated and Table 3.16 gives the allele frequencies.

Figure 3.43: Schematic diagram showing the locations of the investigated SNPs within the *MSH2* gene



The SNPs investigated (shown in blue) were upstream of exon 8 in the *MSH2* gene and located on human chromosome 2 between 47.52 and 47.54Mb. The SNPs between rs1981929 and rs17224423 were considered unsuitable due to their low frequencies of allele variation (i.e the probability of a given allele is ~0.9-1.0).

Table 3.16: Allele frequencies in European population for the SNPs investigated for the couple undergoing PGD for the *MSH2* exon 8 deletion

SNP	Homozygous frequency for ancestor allele	Homozygous frequency for other allele	Heterozygous frequency (both alleleS)
rs1981929	0.333 (A/A)	0.125 (G/G)	0.542 (A/G)
rs7607076	0.500 (A/A)	0.083 (G/G)	0.417 (A/G)
rs17036586	1.000 (G/G)	0.000 (C/C)	0.000 (G/C)
rs7607312	0.288 (A/A)	0.237 (C/C)	0.475 (A/C)
rs17224423	0.977 (G/G)	0.000 (A/A)	0.023 (G/A)

The four SNPs located in proximity to each other (within 200bp) were targeted by one set of primers (set A: MSH2_4SNPsF/R) shown in Figure 3.44. Another set of primers (B) was designed to amplify the region including the rs1981929 SNP and exon 8 (Figure 3.45). This was the same primer set used to form heteroduplex constructs with a single A.C or G.T mismatch.

Figure 3.44: Genomic DNA sequence of the *MSH2* gene on 2p21 showing the designed primer set A targeting four SNPs

```
47524867 GTTTATATTCTGTCTAAATAGTGCTTCACATGTATGTACTATTTTCTAAATATGTATAGA 47524926
47524927 TGCTTTTGTGATTAATAATAAAACATGAATTCTTAAAACAATTTTGCTGACTTCATAGTA 47524986
                                                  MSH2_4SNPsF
47524987 GCTTTTCACCGTTTTTTCAGTAGCTGCTAAAATTTCTGGAGAAGTTTGGGAACTATTGTT 47525046
47525047 TTGGAGTGAAATGCAGTGTGTTAGATATCACTTGCAGAATTCTTCTAAGGGTATTTATTG 47525106
47525107 GCGATTAGAAAAAAAAATCCTTGTGTTATACCAGTAGTAATACAAAGTAATTGTTCAGCTT 47525166
47525167 CTGTTAAGTGTAAAGGACTATACAAGTATTGTGTATAGTTATCTCATTTATTATTTTCTG 47525226
47525227 GGTAGCTATTGTTATTATTACTTCGTACAAAAAGGGAAAAGGAGGCTCAAAGTATCATGC 47525286
47525287 TCCAGATAACAGAGCCAGTAGGTAGCAGAGCTGGGATTGCTACCCAGGTCTCTAGTCCTG 47525346
                                                  MSH2_4SNPsR
47525347 CTTTTTCACACTATATACTCATTGCTTCACTTACTCCTTCATACATGATTCCCCAGCATG 47525406
47525407 TACTCTTTTTTTTTTTTTTTTTTTTTGTTTGAGATAGAATCTCGCTCTCTGTTGCCCAG 47525466
```

The forward and reverse primers are shown in red and green, respectively. The SNPs are highlighted in green. The 309bp long amplified region is located 863bp upstream of exon 8.

Figure 3.45: Genomic DNA sequence of the *MSH2* gene on 2p21 showing the designed primer set B targeting the rs1981929 SNP

```
47525887 TGAAATTCATTTAGTCATAATTAATGTCATGTTTCTGCATCTATATTACTTGTTGGGTTT 47525946
                                                  rs1981929 F
47525947 ACAGAYGAGGTAGTGTATTATTAGTGGGAAGCWTTGAGTGCTACATCATCTCCCTTTCTA 47526006
47526007 TAAAATAAATTGAGTACGAAACAATTTGAATTAAAACACCTGAGTAAATAGTAACTTTGG 47526066
47526067 AGACCTRCTGTACTATTTGTACCTTTTGGATCAAATGATGCTTGTTTATCTCAGTCAAAA 47526126
47526127 TTTTATGATTTGTATTCTGTAAAATGAGATCTTTTTATTTGTTTGTTTTACTACTTTCTT 47526186
47526187 TTAGGAAAACACCAGAAATTATTGTTGGCAGTTTTTGTGACTCCTCTTACTGATCTTCGT 47526246
47526247 TCTGACTTCTCCAAGTTTCAGGAAATGATAGAAACAACTTTAGATATGGATCAGGTATGC 47526306
                                                  rs1981929 R
47526307 AATATACTTTTTAATTTAAGCAGTAGTTATTTTTAAAAAGCAAAGGCCACTTTAAGAAAG 47526366
```

The forward and reverse primers are shown in red and green, respectively. The amplified region is 328bp long. The rs1981929 SNP is highlighted in green. Exon 8 is shown in bold text.

Genomic DNA from the affected individual, his partner and his aunt was amplified using the two primer sets in separate reactions. Amplification of four SNP sites using primer set A was carried out using standard PCR conditions (0.2µM primer and 1.5mM $MgCl_2$); the annealing temperature was 58ºC. The 309bp products were first checked on a 2% agarose gel (section 2.5.1) then sequenced to identify the alleles at the SNP sites (section 2.6.6). The sequences matched the sequence in Figure 3.44 even at the SNP sites. All three individuals showed the same ancestral allele for the four SNPs. Figure 3.46 displays the sequences read for each individual and the full sequences obtained on the ABI Prism™ 3100 genetic analyzer are shown in the Appendix (Figure C.2).

Figure 3.46: Sequencing results for PCR products amplified with primer set A (MSH2_4SNPs)

The reference sequence shown in black is located on Chromosome 2: 47525063-47525269. The sequence read for each PCR product is shown in blue. The SNPs are highlighted in green on the reference sequences and in yellow on the ABI Prism™ readings.

Primer set B (rs1981929-F/R) was used to amplify the SNP that was closest to exon 8. The optimised PCR conditions described in Table 3.12 were used: 56ºC annealing temperature, 0.2µM primer and 1.5mM $MgCl_2$. Sequencing analysis revealed that the proband (male partner) and his affected aunt were homozygous for this SNP, which was expected due to the deletion. Each of them, however, carried a different allele: the proband's was G/- and his aunt's was A/-. The proband's wife was heterozygous (A/G). The full sequences obtained on the ABI Prism™ 3100 genetic analyzer are shown in Appendix (Figure C.3). Figure 3.47 displays the sequence readings of the targeted fragment and Table 3.17 summarises the results obtained for all five SNPs.

Figure 3.47: Sequencing results for PCR products amplified with primer set B (rs1981929-F/R)

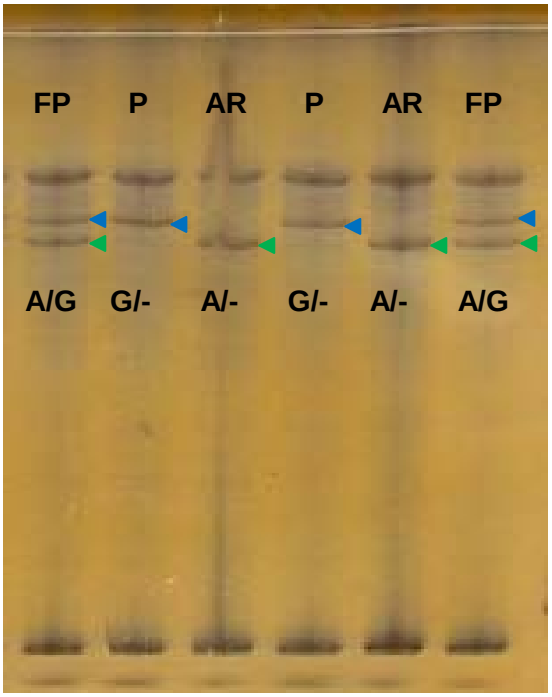
The reference sequence shown in black is located on Chromosome 2: 47526049-47526256. The sequence read for each PCR product is shown in blue. Part of exon 8 was included in the sequence and is shown in bold text. The SNP is highlighted in green on the reference sequences and in yellow on the ABI Prism™ readings. The unaffected partner was heterozygous for the SNP and the affected individuals had only one allele due to the exon 8 deletion.

Table 3.17: Identified alleles at the five SNPs investigated for the proband, his wife and his affected paternal aunt

SNP	Female partner	Male partner (proband)	Affected relative	Informativity
rs1981929	A/G	G/-	A/-	Semi-informative
rs7607076	A/A	A/A or A/-	A/A or A/-	Not informative
rs17036586	G/G	G/G or G/-	G/G or G/-	Not informative
rs7607312	A/A	A/A or A/-	A/A or A/-	Not informative
rs17224423	G/G	G/G or G/-	G/G or G/-	Not informative

SSCP analysis (section 2.6.5) was used for the detection of the different genotypes at the rs1981929 SNP (done in collaboration with Will J. Young during his BSc project at the UCL Centre for PGD). Figure 3.48 shows SSCP analysis of the amplified products for the proband, his wife and his affected aunt.

Figure 3.48: SSCP analysis on PhastGel® Homogeneous 20 run at 10ºC (long pre-run) showing genomic DNA from the proband, his wife and his affected aunt amplified at the rs1981929 locus

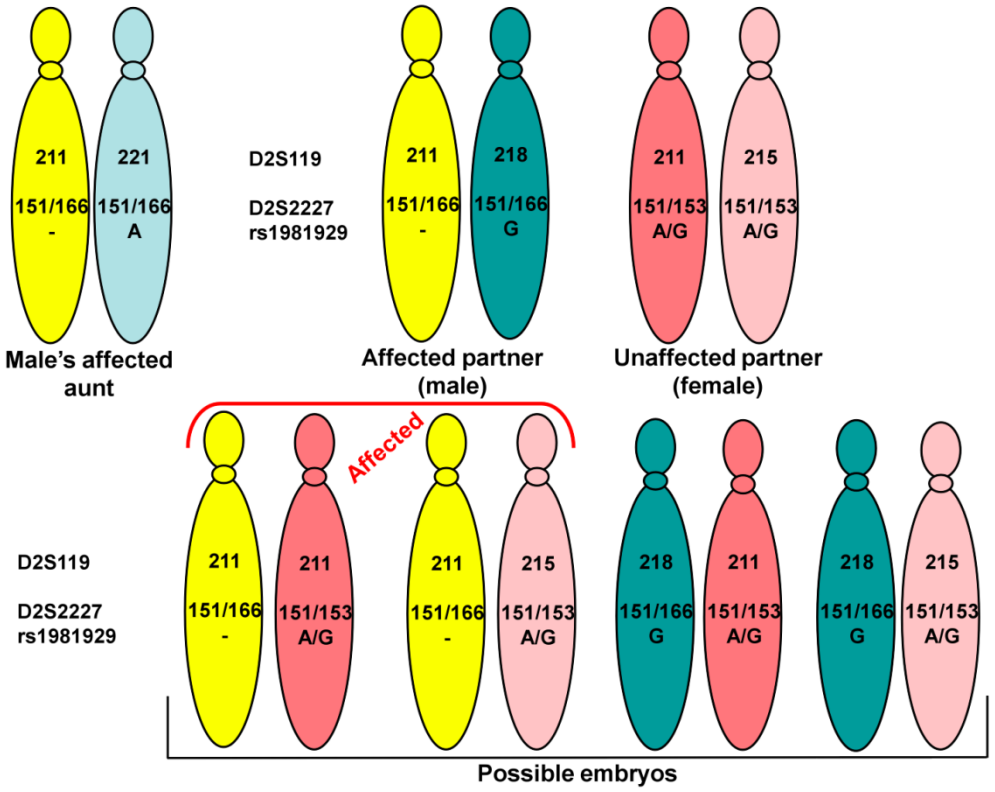


P: proband (male)
FP: female partner
AR: affected relative (proband's paternal aunt)

The bottom most band (indicated by green arrows) corresponds to the A allele, which was missing in the proband who carried the exon 8 deletion. The blue arrows point to the bands corresponding to the G allele. The presence of both bands identifies a normal heterozygous A/G sample. Normal homozygous samples cannot be distinguished from those carrying the deletion using SSCP analysis alone.

The identification of an A/G genotype would indicate a 'normal' or unaffected cell/embryo. However, G/- or A/- could not be distinguished form homozygous G/G or A/A genotypes unless the couple's haplotypes were known. The marker alleles that were in phase with the *MSH2* exon 8 deletion could only be identified for D2S119. Figure 3.49 shows the couple's haplotypes and those of their possible embryos.

Figure 3.49: Possible embryo haplotypes for the couple undergoing PGD for the *MSH2* exon 8 deletion



The chromosome carrying the mutation (shown in yellow) was identifiable by the 211bp allele size for the D2S119 STR marker. Embryos inheriting that paternal chromosome would carry the mutation. Linkage would be confirmed by the D2S2227 linked marker which is 0.36Mb away from the gene. If the 151bp shared allele were in phase with the mutation, detection of 151/166 and 153/166 would indicate a 'normal' or unaffected embryo. If the 166bp allele were in phase with the mutation, detection of 151 homozygous or 151/153 would indicate a normal/unaffected embryo.

In order to determine the couple's haplotypes, other DNA samples from immediate family members were needed; however, this was not available. Alternatively, the phase allele for the D2S2227 marker could have been identified from haploid cells isolated from the proband's semen sample. This could not be obtained during the initial stages of the workup process. However, the phase would have been worked out later on or during the treatment cycle had the couple gone ahead with PGD.

3.3.1.3 Optimisation of PCR protocol

3.3.1.3.1 Multiplex PCR optimisation using genomic DNA

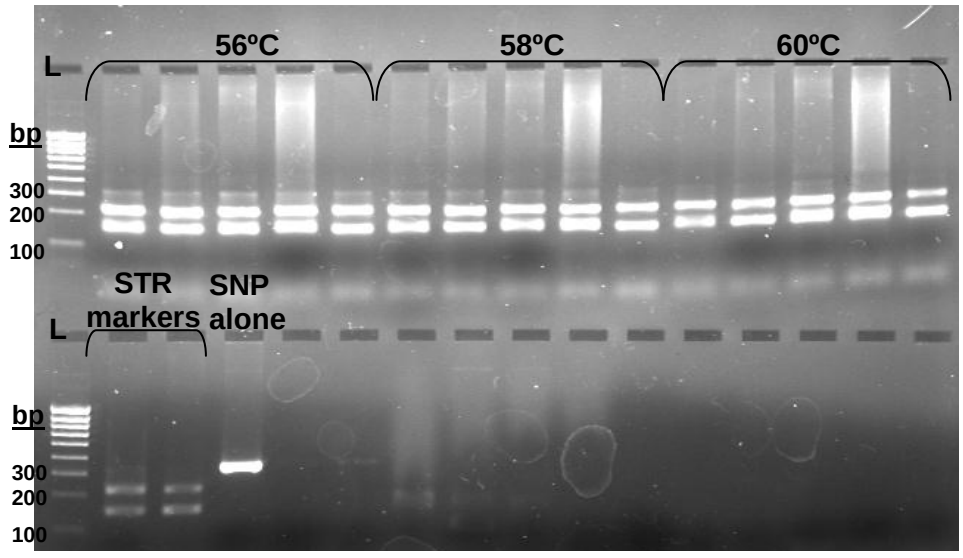
A quadruplex PCR was used to amplify the couple's genomic DNA at the following loci: the two semi-informative STR markers (D2S119 and D2S2227), the semi-informative SNP (rs1981929) and the informative contamination marker (APOC2). As the primers for both linked markers and the SNP worked separately at 60ºC and 56ºC, respectively, a temperature gradient PCR was set up to compare three annealing temperatures: 56, 58 and 60ºC. Table 3.18 summaries the initial conditions tested.

Table 3.18: Conditions of quadruplex PCR during annealing temperature optimisation

Reagent	D2S119 primers	D2S2227 primers	APOC2 primers	rs1981929 primers	$MgCl_2$	Ta
	0.2µM	0.2µM	0.2µM	0.2µM	1.5mM	56ºC
	0.2µM	0.2µM	0.2µM	0.2µM	1.5mM	58ºC
	0.2µM	0.2µM	0.2µM	0.2µM	1.5mM	60ºC

Analysis of the products on a 2% agarose gel revealed that better amplification was achieved at 56 and 58ºC; however, amplification of the SNP region (328bp product) was weak at all temperatures (Figure 3.50).

Figure 3.50: 2% agarose gel (run at 70V) showing quadruplex PCR products at different annealing temperatures: 56, 58 and 60ºC



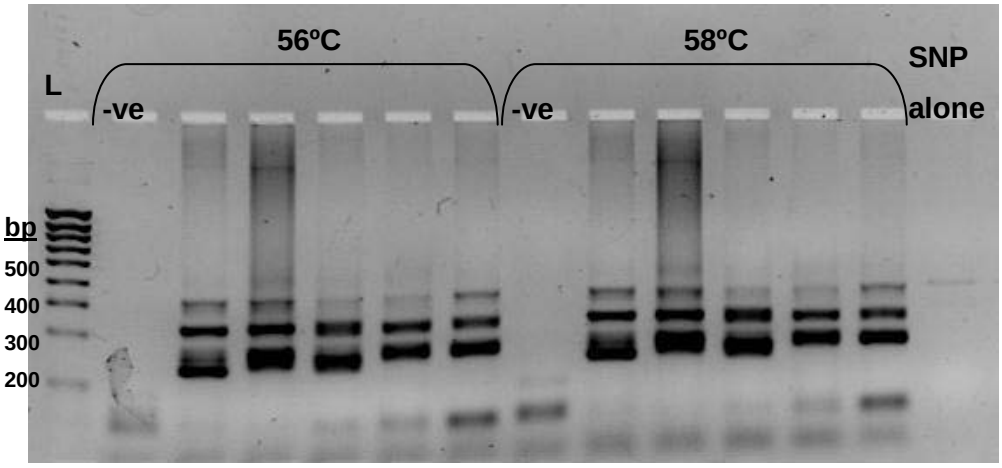
L: 1Kb ladder; Expected product sizes: 328bp for rs1981929 SNP, ~150bp for D2S2227 and APOC2 and 210-220bp for D2S119
Looking at the band intensities it was determined that the best annealing temperature was 56ºC. The STR markers loci were amplified efficiently at all temperatures and best at 56ºC. The SNP locus was weakly amplified at all temperatures, with lowest amplification at 60ºC.

The concentration of the SNP primers was increased to 0.4µM, keeping the rest of the reagents the same, and the PCR was run at 56 and 58ºC (Table 3.19). These conditions gave good results for all loci including the SNP (Figure 3.51).

Table 3.19: Optimisation of primer concentrations for quadruplex PCR using genomic DNA

Reagent	D2S119 primers	D2S2227 primers	APOC2 primers	rs1981929 primers	$MgCl_2$	Ta
	0.2µM	0.2µM	0.2µM	0.4µM	1.5mM	56ºC
	0.2µM	0.2µM	0.2µM	0.4µM	1.5mM	58ºC

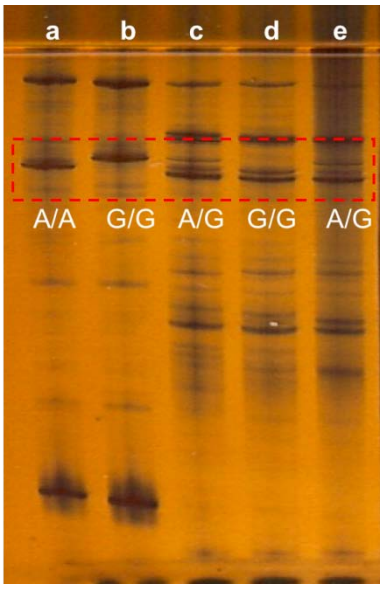
Figure 3.51: 2% agarose gel (run at 70V) showing quadruplex PCR products at 56 and 58ºC annealing temperatures with 0.4µM rs1981929 SNP primers



L: 1Kb ladder; -ve: negative control; Expected product sizes: 328bp for rs1981929 SNP, ~150bp for D2S2227 and APOC2 and 210-220bp for D2S119
The quadruplex PCR had 0.4µM rs1981929 SNP primers and 0.2µM of all other primers (D2S2227, D2S119 and APOC2). All loci were amplified efficiently at both annealing temperatures.

SSCP analysis was carried out on the quadruplex PCR products (Ta=58ºC) to identify the SNP alleles (Figure 3.52). The alleles could be identified; however, due to the presence of extra bands from other amplified loci, some of which were near the rs1981929 product bands, diagnosis would be difficult using this technique. Mini-sequencing primers were thus designed to target the rs1981929 SNP (Figure 3.53).

Figure 3.52: SSCP analysis on PhastGel® Homogeneous 20 run at 4ºC (long pre-run) showing the couple's genomic DNA amplified at four loci, including the rs1981929 locus



a: Singleplex product for rs1981929 from homozygous control DNA (A/A)
b: Singleplex product for rs1981929 from homozygous control DNA (G/G)
c, d, e: Quadruplex products (Ta=58ºC) from genomic DNA of the unaffected female partner, the proband and a heterozygous control, respectively
The red box shows the migration level of the rs1981929 products

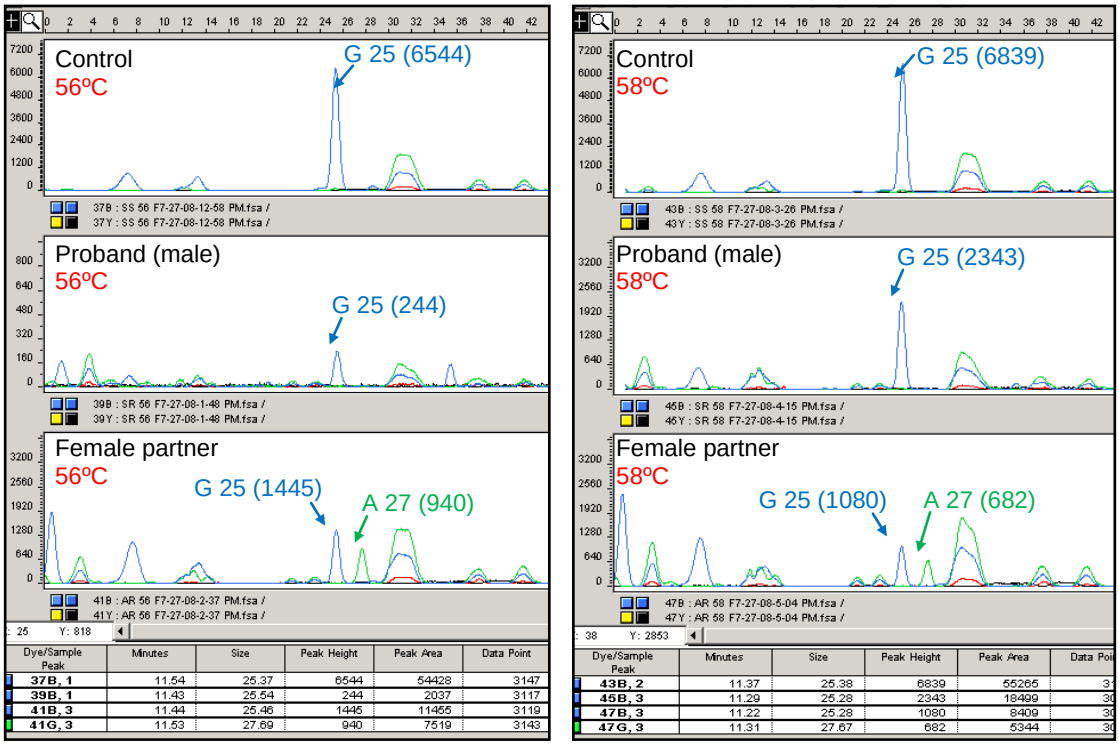
Figure 3.53: Genomic DNA sequence of the *MSH2* gene on 2p21 showing the mini-sequencing primers designed for the rs1981929 SNP

```
47525887 TGAAATTCATTTAGTCATAATTAATGTCATGTTTCTGCATCTATATTACTTGTTGGGTTT 47525946
                                                  rs1981929 F
47525947 ACAGAYGAGGTAGTGTATTATTAGTGGGAAGCWTTGAGTGCTACATCATCTCCCTTTCTA 47526006
                                                          For A/G
47526007 TAAAATAAATTGAGTACGAAACAATTTGAATTAAAACACCTGAGTAAATAGTAACTTTGG 47526066
47526067 AGACCTRCTGTACTATTTGTACCTTTTGGATCAAATGATGCTTGTTTATCTCAGTCAAAA 47526126
                  Rev C/T
47526127 TTTTATGATTTGTATTCTGTAAAATGAGATCTTTTTATTTGTTTGTTTTACTACTTTCTT 47526186
47526187 TTAGGAAAACACCAGAAATTATTGTTGGCAGTTTTTGTGACTCCTCTTACTGATCTTCGT 47526246
47526247 TCTGACTTCTCCAAGTTTCAGGAAATGATAGAAACAACTTTAGATATGGATCAGGTATGC 47526306
                                                  rs1981929 R
47526307 AATATACTTTTTAATTTAAGCAGTAGTTATTTTTAAAAAGCAAAGGCCACTTTAAGAAAG 47526366
```

The initial rs1981929 PCR primers are shown in red and green and the SNaPshot forward (A/G) and reverse (T/C) primers are shown in blue and purple, respectively. The rs1981929 SNP is highlighted in green and exon 8 is shown in bold text.

Mini-sequencing analysis was carried out on the quadruplex PCR products (Table 3.19) using the SNaPshot™ multiplex kit (method described in section 5.6). Initial results showed that mini-sequencing was successful (Figure 3.54) and that amplification of the rs1981929 site in the quadruplex reaction was similar at 56ºC and 58ºC. The correct base was identified for the expected fragment size (primer length + 4, 3 and 2bp, for the A, G and C/T bases, respectively). As some additional random sized peaks were detected, further optimisation on single cells was necessary.

Figure 3.54: Mini-sequencing analysis from quadruplex PCR products (Ta=56 and 58ºC) showing the different rs1981929 SNP alleles for the couple and a control



Dye/Sample Peak	Minutes	Size	Peak Height	Peak Area	Data Point
37B, 1	11.54	25.37	6544	54428	3147
39B, 1	11.43	25.54	244	2037	3117
41B, 3	11.44	25.46	1445	11455	3119
41G, 3	11.53	27.69	940	7519	3143

Dye/Sample Peak	Minutes	Size	Peak Height	Peak Area	Data Poi
43B, 2	11.37	25.38	6839	55265	3
45B, 3	11.29	25.28	2343	18499	3
47B, 3	11.22	25.28	1080	8409	3
47G, 3	11.31	27.67	682	5344	3

Analysis on the ABI Prism™ 310 was expected to show a blue 25bp peak or green 26bp peak for the G and A bases, respectively, when the forward primer was used (ForA/G). The C and T products, obtained with the reverse primer (not shown here), would be visualised as black and red 18bp sized peaks, respectively.
Mini-sequencing analysis allowed the identification of the correct alleles for the proband (G/-), his wife (G/A) and a control homozygous (G/G). Results were similar for both quadruplex reactions performed at 56 and 58ºC (Table 3.19).

The developed quadruplex PCR protocol, involving mini-sequencing for the rs1981929 SNP and fragment length analysis of two linked and one unlinked STR markers, was licensed for PGD treatment by the HFEA. This was the first license issued for a MMR gene in the UK and set the precedence for future cases.

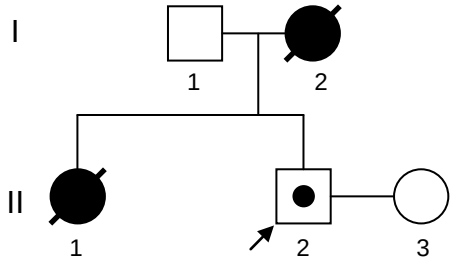
The PCR was never optimised on single cells as the couple had decided not to go through with PGD treatment during the workup process. This was due to the couple's inability to obtain NHS funding for IVF/PGD treatment.

3.3.2 PGD workup for BRCA1 (c. 3339T>G)

3.3.2.1 Patient description

The male partner (proband) carried a non-sense c.3339T>G, p.Tyr1113X mutation (formerly known as c.3458T>G) in exon 11 of the *BRCA1* gene. The proband's wife was homozygous normal. His mother and sister carried the germline mutation and died of breast cancer at ages of 40 and 35, respectively. Figure 3.55 shows the pedigree. Bloods for DNA extraction were available from the proband, his father (homozygous normal) and his wife.

Figure 3.55: Pedigree showing the asymptomatic proband and his affected relatives



II.2 is the proband, carrier of the c.3339T>G mutation in exon 11 of the *BRCA1* gene. I.2 and II.1 represent the proband's affected mother and sister, respectively.

3.3.2.2 Confirmation of mutation using sequencing analysis

A set of non-labelled PCR primers (c.3339T>G) flanking the single base substitution was designed for direct mutation detection using SSCP analysis (Figure 3.56). These primers were initially used to confirm the mutation specified in the patient's genetic report. DNA from the proband, his partner and his father was amplified at the mutation site using these primers. The PCR products for the three individuals were first checked on agarose gels to confirm DNA amplification (and successful DNA extraction after the very first PCR) then sequenced. The sequences were all identical (matched the sequence in Figure 3.56) except for a single base change detected in the male partner's sample at the mutation site. The full sequences obtained on the ABI Prism^{TM} 3100 genetic analyzer are shown in the Appendix (Figure C.4).

Figure 3.56: Genomic DNA sequence of the *BRCA1* gene (reference sequence: L78833, residues 36121-36660) showing the primers designed for the c.3339T>G mutation

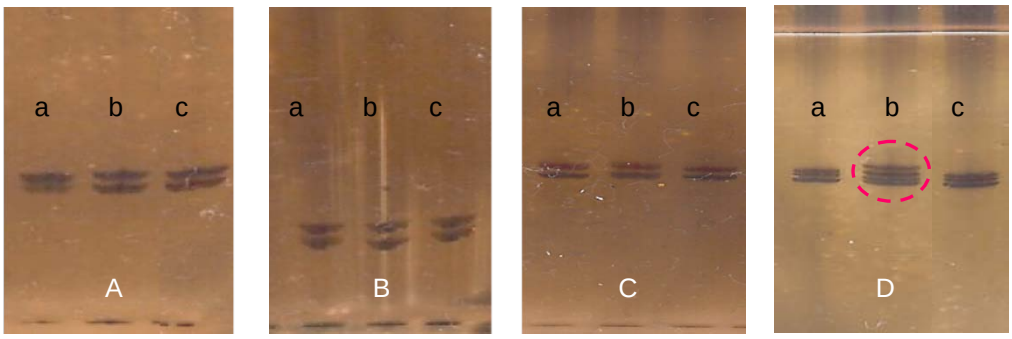
```
36121 ttacaaaacc catatcgtat accaccactt tttcccatca agtcatttgt taaaactaaa
36181 tgtaagaaaa atctgctaga ggaaaacttt gaggaacatt caatgtcacc tgaaagagaa
36241 atgggaaatg agaacattcc aagtacagtg agcacaatta gccgtaataa cattagagaa
36301 aatgttttta aagaagccag ctcaagcaat attaatgaag taggttccag tactaatgaa
36361 gtgggctcca gtattaatga aataggttcc agtgatgaaa acattcaagc agaactaggt
36421 agaaacagag ggccaaaatt gaatgctatg cttagattag gggttttgca acctgaggtc
36481 tataaacaaa gtcttcctgg aagtaattgt aagcatcctg aaataaaaaa gcaagaatat
36541 gaagaagtag ttcagactgt taatacagat ttctctccat atctgatttc agataactta
36601 gaacagccta tgggaagtag tcatgcatct caggtttgtt ctgagacacc tgatgacctg
```

The sequence displayed was selected from of exon 11, which is >3.4kb long. The codon at the mutation site is underlined. The forward and reverse primers are highlighted in yellow and green, respectively. (Product size=136bp)

3.3.2.3 Direct mutation detection using SSCP analysis

Amplified DNA products at the *BRCA1* c.3339T>G mutation site for the proband, his partner and his father were analysed on SSCP gels. Successful separation of the single DNA strands required optimisation of the SSCP conditions (Figure 3.57). The optimal conditions were: 20% homogenous gel, run at 4ºC with long pre-run. The mutation could be detected by the presence of an extra band that was indicative of a heterozygous carrier of the c.3339T>G mutation.

Figure 3.57: SSCP optimisations for *BRCA1* c.3339T>G mutation detection



Gel Conditions	Gel concentration	Temperature	Pre-run	Mutation detection
A	20%	10ºC	short	No
B	20%	15ºC	short	No
C	20%	4ºC	short	No
D	20%	4ºC	long	Yes

Amplified products using the c.3339T>G primers on genomic DNA from the unaffected female partner (a), the proband (b) and his unaffected father (c) were run on a homogeneous 20 PhastGel® at different conditions (A-D). The optimal conditions (D) were: 4ºC with a long pre-run. Product b showed a clearly different pattern (extra band) due to the c.3339T>G mutation carried by the male partner.

3.3.2.4 Informativity tests

Informativity tests were carried out on genomic DNA from the affected male and his partner. Four linked and six unlinked STR markers were investigated using F-PCR followed by fragment length analysis on the ABI Prism™ (Table 3.20). Amongst the linked markers tested, only D17S1338 was found to be fully informative for this couple. Both the DM and D13S168 unlinked markers were found to be informative. However, as the allele sizes would be close together in embryos from the couple for DM, D13S168 was used as the fully informative contamination marker.

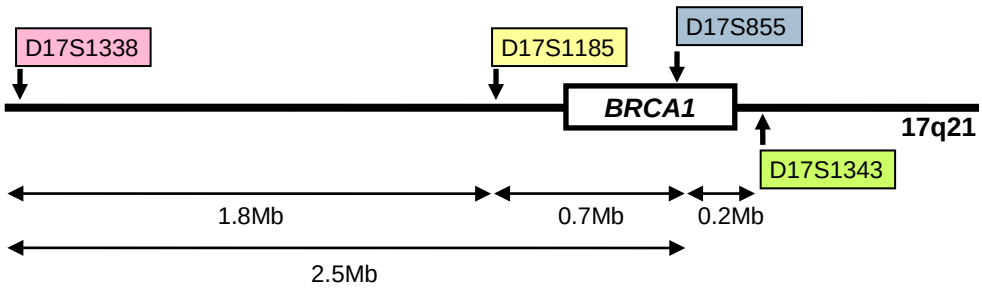
Table 3.20: Allele sizes for different STR markers identified for the couple undergoing PGD for the *BRCA1* c.3339T>G mutation

STR marker	Location	Female partner (bp)	Male partner (proband) (bp)	Informativity
Linked STR markers				
D17S1338 *	**17q21.31**	**183/193**	***179/191***	**Informative**
D17S1185	17q21.31	213/217	213/217	Not informative
D17S855	17q21.31; *BRCA1* intron 11	157	*157/160*	Semi-informative
D17S1343	17q21.31	153/158	*153/158*	Not informative
Unlinked STR markers				
APOC2	19q13.32	150/155	*150/152*	Semi-informative
DM	**19q13.32**	**121/148**	***145/150***	**Informative**
D19S112	19q13.32	130/132	*120/130*	Semi-informative
D13S168 *	**13q14.2**	**126/137**	***141/143***	**Informative**
D13S262	13q14.3	206/325	*206/325*	Not informative
Rb1.20	13q14.2	175/179	*175/187*	Semi-informative

The affected individual is represented in italics. Bold text signifies that the alleles are informative and underlined text indicates the phase allele. The top markers (shown in blue) are linked with the *BRCA1* gene. The asterisk denotes the markers that were selected for use in the multiplex PCR reaction.

The only informative linked marker, D17S1338, was the furthest away from the *BRCA1* gene. The distance between this marker and the gene was approximately 2.5Mb which is greater than the desired limit of 1Mb. Figure 3.58 shows the location of the linked markers on chromosome 17 relative to the *BRCA1* gene. D17S855, which is an intragenic marker, was found to be semi-informative.

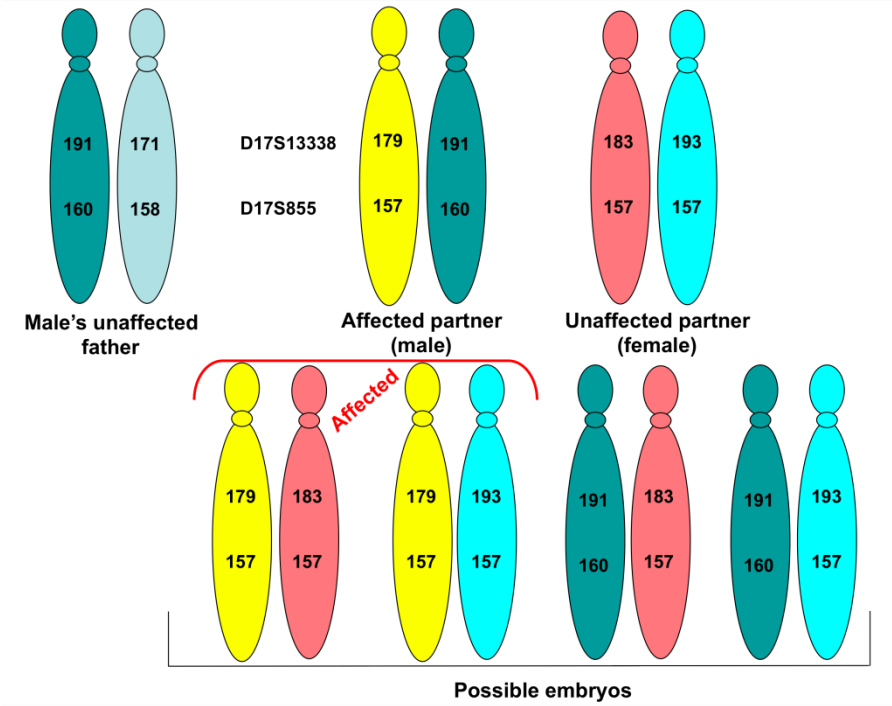
Figure 3.58: *BRCA1* gene and linked STR markers on chromosome 17



The locations of D17S1338, D17S1185, D17S855 and D17S1343 on human chromosome 17 were around 34.5, 36.3, 37.0 and 37.2Mb, respectively. The calculated distance between the *BRCA1* gene and each STR marker is shown in the diagram.

DNA from the unaffected father of the proband was amplified for the two linked markers, D17S1338 and D17S855. This allowed the identification of the marker alleles that were in phase with the *BRCA1* c.3339T>G mutation that the proband inherited from his mother. The 157 and 179bp allele sizes for D17S855 and D17S1338, respectively, were found to be in phase with the mutation. Figure 3.59 illustrates the haplotypes of the couple, the male's unaffected father and the possible embryos.

Figure 3.59: Haplotypes of couple undergoing PGD for the *BRCA1* c.3339T>G mutation and their possible embryos



The chromosome carrying the mutation (shown in yellow) was identifiable by the 157 and 179bp allele sizes for the D17S855 and D17S1338 markers, respectively. Embryos inheriting that paternal chromosome would carry the mutation.

Three STR markers were selected for use in a multiplex PCR along with the mutation primers for the PGD protocol: D17S1338, D17S855 and D13S168.

3.3.2.5 Optimisation of PCR protocol

3.3.2.5.1 Multiplex PCR optimisation using genomic DNA

A quadruplex PCR was initially attempted using basic condition (0.2µM of each primer, 1.5mM $MgCl_2$) to amplify three STR markers (D17S1338, D17S855 and D13S168) and the c.3339T>G mutation. The initial annealing temperature was 60ºC, as all primer sets worked individually at this annealing temperature. All loci were amplified efficiently; the correct allele sizes for the three markers were detected on the genetic analyzer and the mutation was detected by SSCP analysis. A few single cells were then tested using this PCR protocol; amplification was very poor for all single cells at all loci. Thus, further optimisations on genomic DNA from the affected individual and controls were necessary before moving on to single cells.

A. Annealing temperature and $MgCl_2$ concentration:

A magnesium/temperature gradient PCR was set up to check amplification efficiency with 1.5, 2.0 and 2.5mM $MgCl_2$ at annealing temperatures of 58, 60 and 62ºC. All reagents concentrations (other than $MgCl_2$) were kept the same but 10% glycerol was added to the master mixes (Table 3.21). Fragment length analysis on the genetic analyzer showed that the best annealing temperature was 60ºC for all loci. 2.5mM $MgCl_2$ gave the best results for D17S1338 and D17S855; however, 2.0mM $MgCl_2$ gave better results for D13S168 (less unspecific amplification).

Table 3.21: Conditions of quadruplex PCR during $MgCl_2$/annealing temperature optimisation

Ta	Mutation primers	D17S1338 primers	D17S855 primers	D13S168 primers	$MgCl_2$	10% glycerol	Conditions optimal for:
58ºC	0.2µM	0.2µM	0.2µM	0.2µM	1.5mM	Yes	
	0.2µM	0.2µM	0.2µM	0.2µM	2.0mM	Yes	
	0.2µM	0.2µM	0.2µM	0.2µM	2.5mM	Yes	
60ºC	0.2µM	0.2µM	0.2µM	0.2µM	1.5mM	Yes	
	0.2µM	0.2µM	0.2µM	0.2µM	2.0mM	Yes	D13S168
	0.2µM	0.2µM	0.2µM	0.2µM	2.5mM	Yes	D17S138 & D17S855
62ºC	0.2µM	0.2µM	0.2µM	0.2µM	1.5mM	Yes	
	0.2µM	0.2µM	0.2µM	0.2µM	2.0mM	Yes	
	0.2µM	0.2µM	0.2µM	0.2µM	2.5mM	Yes	

B. Split PCRs:

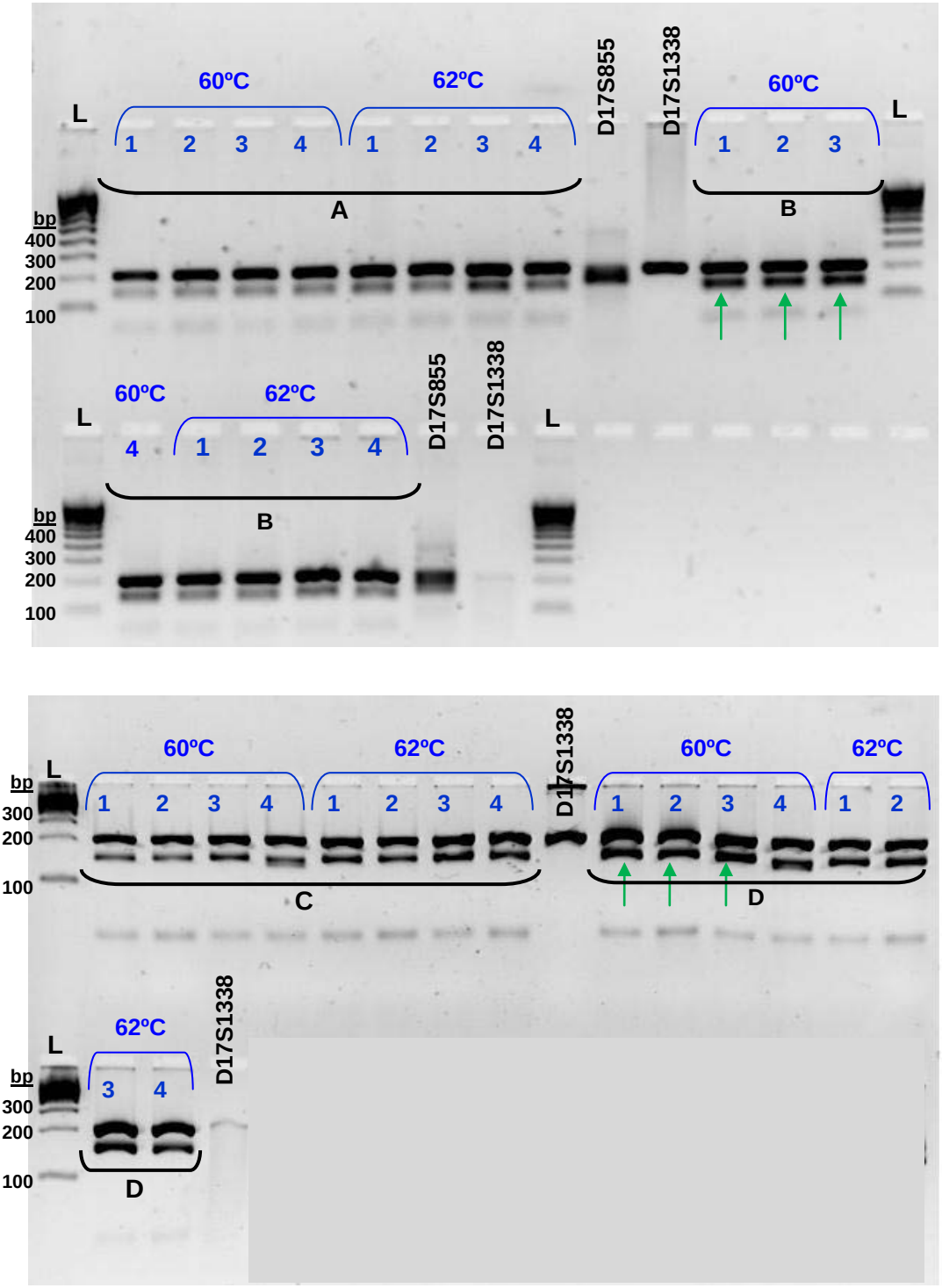
The quadruplex PCR was split into two rounds of amplification. The first round involved the amplification of all loci in 15 cycles. The second round was a full PCR reaction (40 cycles) using 3µl of PCR product obtained from the first round. Round 2 was carried out for D13S168 and c.3339T>G in one PCR and D17S1338 and D17S855 in another PCR.

The addition of glycerol, $MgCl_2$ concentration (2.0 and 2.5mM) and two annealing temperatures (Ta = 60ºC or 62ºC) were tested for both rounds of amplification (Table 3.22). Analysis of all products of round 2 on agarose gels and on the genetic analyzer showed that the best round 1 conditions for all loci were 2.0mM $MgCl_2$ with 10% glycerol and Ta=60ºC. The best conditions for round 2 were 2.5mM $MgCl_2$ with 10% glycerol (Ta=60ºC) for D17S1338 and D17S855 (Figure 3.60), and 2.0mM $MgCl_2$ without glycerol (Ta=62ºC) for D13S168 and c.3339T>G. These conditions were thus tested on single cells.

Table 3.22: Summary of conditions tested for the optimisation of split PCR protocol using genomic DNA

Protocol conditions	Round 1 PCR [$MgCl_2$] (mM)	10% glycerol	Round 2 PCR [$MgCl_2$] (mM)	10% glycerol
A1	2.0	Yes	2.0	No
A2	2.5	Yes	2.0	No
A3	2.0	No	2.0	No
A4	2.5	No	2.0	No
B1	2.0	Yes	2.0	Yes
B2	2.5	Yes	2.0	Yes
B3	2.0	No	2.0	Yes
B4	2.5	No	2.0	Yes
C1	2.0	Yes	2.5	No
C2	2.5	Yes	2.5	No
C3	2.0	No	2.5	No
C4	2.5	No	2.5	No
D1	2.0	Yes	2.5	Yes
D2	2.5	Yes	2.5	Yes
D3	2.0	No	2.5	Yes
D4	2.5	No	2.5	Yes

Figure 3.60: 2% agarose gel (run at 70V) showing optimisation of split PCR for D17S1338 and D17S855 using genomic DNA



L: 1Kb ladder; Samples 1, 2, 3 and 4 were genomic control DNAs.
The optimal PCR conditions were identified by looking at the intensities of two bands: 150-160bp for D17S855 and 180-190bp for D17S1338. Two annealing temperatures (Ta) were tested for the first round: 60 and 62ºC (as shown on top of the gel wells) and Ta was always 60ºC in the second round. The conditions of the first (blue) and second (black) round PCRs are marked on the gel images and listed in Table 3.22. The green arrows designate the best conditions observed on each gel. The optimal conditions for round 1 were: Ta=60ºC, 2.0 or 2.5mM $MgCl_2$ with 10% glycerol. The optimal conditions for round 2 were: 2.5mM $MgCl_2$ with 10% glycerol (Ta=60ºC was constant).

3.3.2.5.2 Multiplex PCR optimisation for single cells

The quadruplex split PCR optimised on genomic DNA was tested on 10 isolated lymphocytes for the couple. Amplification of the mutation site, the unlinked marker and D17S1338 was efficient (≥90%); however, that of the intragenic marker D17S855 was still poor (70%). ADO was observed for one cell for each of the markers (Table 3.23).

Table 3.23: Initial assessment of split PCR protocol on 10 single cells

Loci	*BRCA1* c.3339T>G mutation	D13S168 unlinked marker	D17S1338 linked marker	D17S855 intragenic marker
AE	90% (9/10)	100% (10/10)	90% (9/10)	70% (7/10)
ADO	-	10% (1/10)	11% (1/9)	14% (1/7)

AE: amplification efficiency; ADO: allele dropout

PCR protocol:

- Round 1: Ta=60ºC, 2.0mM $MgCl_2$ with 10% glycerol
- Round 2:
  - Mutation + D13S168: Ta=62ºC, 2.0mM $MgCl_2$ without glycerol
  - D17S1338 + D17S855: Ta=60ºC, 2.5mM $MgCl_2$ with 10% glycerol

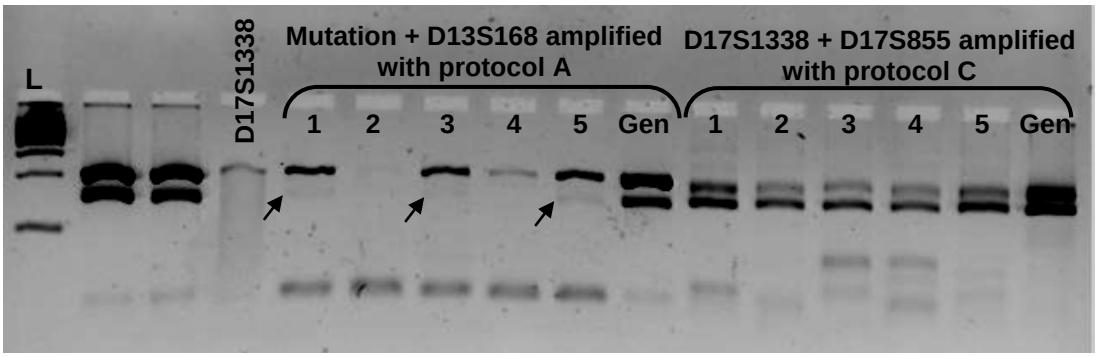
In order to further optimise this split PCR, the first round amplification was performed on five fresh single cells followed by two different conditions for each of the round 2 PCRs. The different protocols along with the Prism analysis results are listed in Table 3.24. Figure 3.61 shows the amplified products obtained from protocols A and C on a 2% agarose gel.

Table 3.24: Summary of PCR protocols tested on 5 single cells and fragment size analysis results

Split reaction	Protocol	Round 1			Round 2			Marker AE & ADO	Comment
		[$MgCl_2$] (mM)	10% glycerol	Ta (ºC)	[$MgCl_2$] (mM)	10% glycerol	Ta (ºC)		
Mutation + D13S168	A	2.0	Yes	60	2.0	Yes	62	5/5 1/5 ADO	Weak amplification
	B	2.0	Yes	60	2.0	No	60	3/5 0% ADO	Amplification better than A
D17S855 + D17S1338	C	2.0	Yes	60	2.5	Yes	60	5/5 for both 0% ADO	
	D	2.0	Yes	60	2.5	No	60	0/5 for both -	

AE: amplification efficiency; ADO: allele dropout
Low peaks were observed on the prism when amplification was weak.

Figure 3.61: 2% agarose gel (run at 70V) showing amplified DNA from the proband's single cells using the split PCR protocols A & C (Table 3.24)



L: 1Kb ladder; Expected product sizes: 126-140bp for D13S168, 136bp for the mutation 150-160bp for D17S855 and 180-190bp for D17S1185
Samples 1-5 designate single lymphocytes isolated from the proband. Gen. designates genomic DNA from the same individual used as positive control. The arrow points at the weak bands obtained from amplified DNA at the mutation and/or the D13S168 locus. Only three out of five cells amplified at these loci using protocol A (Table 3.24). The two linked marker (D17S1338 and D17S855) loci were amplified in 5 out of 5 cells, showing that protocol C was suitable for these markers.

The amplification results of the split quadruplex PCR were still not good enough for a PGD protocol, particularly for the mutation site and the contamination marker (D13S168). As the D17S855 marker was neither fully informative nor necessary for diagnosis, given that the PCR protocol included another fully informative linked marker and a contamination markers, D17S855 was dropped from the PCR protocol.

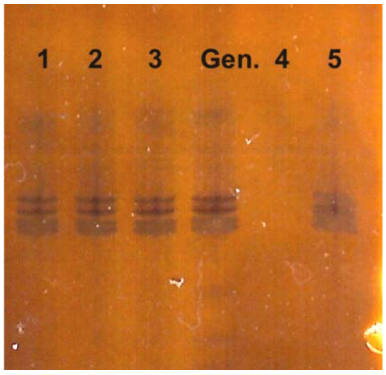
The new protocol included three sets of primers for the following loci: the *BRCA1* c.3339T>G mutation, the D17S1338 and D13S168 linked and unlinked polymorphic Markers, respectively. Based on the amplification efficiencies observed for these loci during the quadruplex optimisation experiments, the triplex PCR was performed in two rounds. The second round was split into a duplex F-PCR targeting the two markers (D17S1338 and D13S168) and a separate PCR for direct mutation detection. The conditions of all three PCRs are summarised in Table 3.25. This protocol gave high amplification efficiencies from single cells and genomic DNA. Initial results from 10 lymphocytes from the proband and his partner showed 90% amplification efficiency (AE) for the mutation and 100% AE for the markers. 0% allele dropout (ADO) was observed for all loci. Figure 3.62 shows amplification results from five single cells at the mutation site on SSCP gels.

Table 3.25: Summary of split PCR protocol tested on 10 single cells

Split PCR	Mutation primers	D17S1338 primers	D13S168 primers	[$MgCl_2$]	10% glycerol	Ta
Round 1	0.2µM	0.2µM	0.2µM	2.5mM	Yes	60**ºC**
Round 2 a	-	0.2µM	0.2µM	2.5mM	Yes	60**ºC**
Round 2 b	0.2µM	-	-	2.0mm	No	60**ºC**
AE	9/10	10/10	10/10			
ADO	0%	0%	0%			

AE: amplification efficiency; ADO: allele dropout

Figure 3.62: SSCP analysis showing *BRCA1* c.3339T>G mutation detection from five lymphocytes after amplification using the split triplex PCR protocol



Homogeneous 20 PhastGel® run at 4ºC (long pre-run)
Samples 1,2,3,4 and 5 designate the amplified DNA at the mutation locus from single lymphosytes isolated from the affected male partner. Gen. designates the genomic DNA samples used as a positive control.
Lymphocyte number 4 failed to amplify at the mutation site. AE was 80% (4 out of 5 cells) and ADO was 0%.

3.3.2.6 Final assessment of PGD protocol

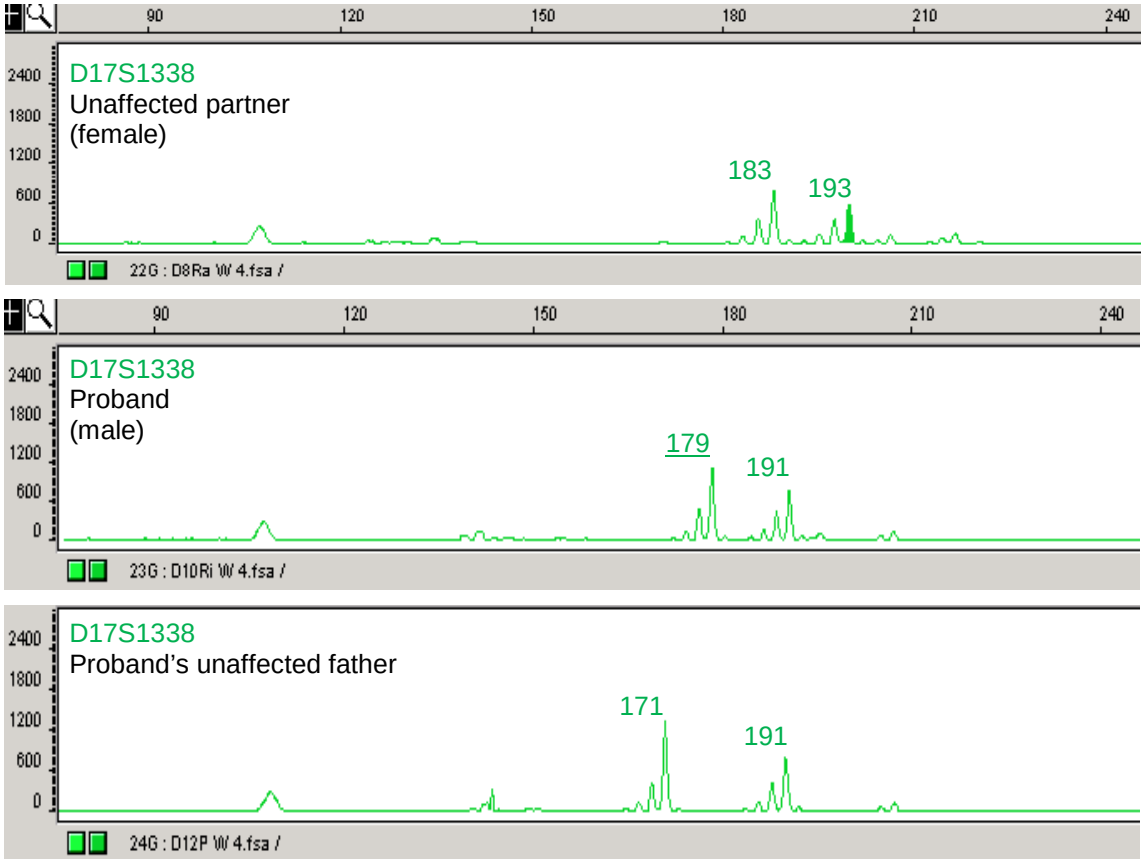
The optimised protocol, shown in Table 3.26, was assessed on 50 lymphocytes that were heterozygous for all loci including the mutation (Figure 3.63). The lymphocytes were isolated from the proband's blood. The outcome is summarised in Table 3.27.

Table 3.26: Summary of optimised split PCR protocol for the *BRCA1* c.3339T>G mutation

Split PCR	Mutation primers	D17S1338 primers	D13S168 primers	[$MgCl_2$]	10% glycerol	Ta	Method of analysis
Round 1	0.2µM	0.2µM	0.2µM	2.5mM	Yes	60**ºC**	
Round 2 a	-	0.2µM	0.2µM	2.5mM	Yes	60**ºC**	ABI genetic analyzer
Round 2 b	0.2µM	-	-	2.0mm	No	60**ºC**	SSCP

Figure 3.63: Analysis of PCR products obtained with the final PGD protocol for the *BRCA1* c.3339T>G mutation

a) Genetic analyzer panels showing the linked informative marker



The underlined allele is in phase with the mutation.

b) Genetic analyzer panels showing the unlinked informative marker

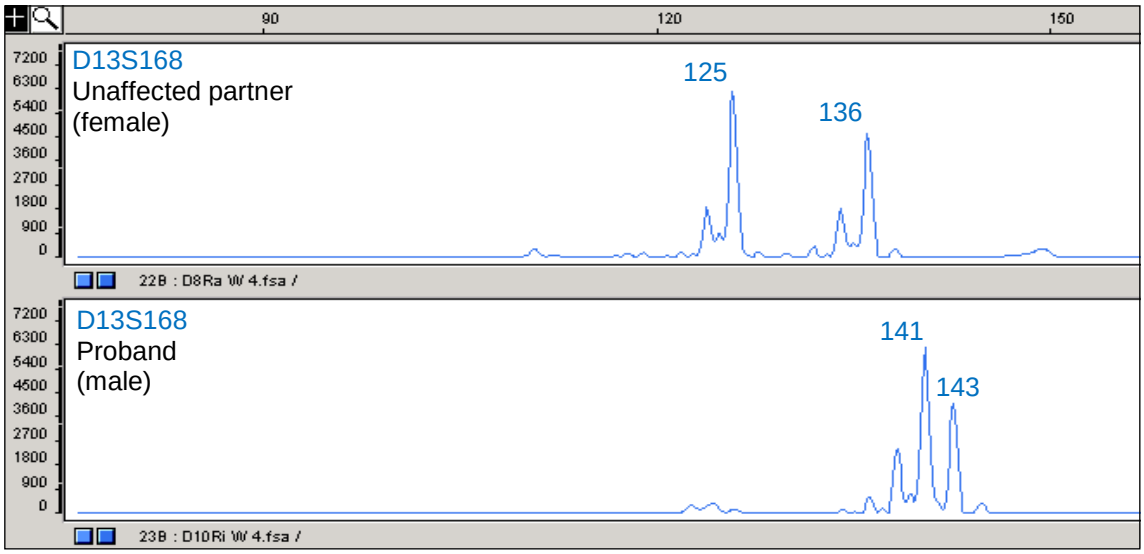
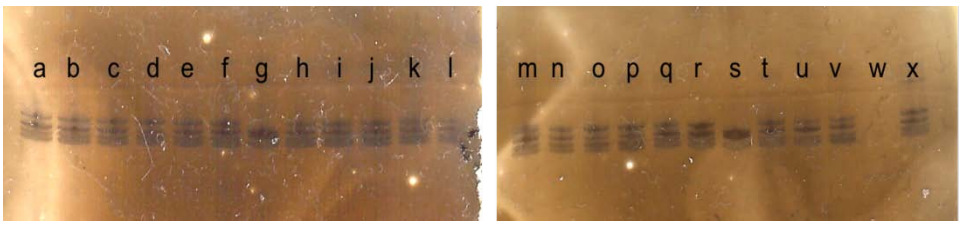


Figure 3.63 (continued): Analysis of PCR products obtained with the final PGD protocol for the *BRCA1* c.3339T>G mutation

c) SSCP analysis showing the *BRCA1* c.3339T>G mutation detection for single lymphocytes from the proband



ADO was observed for samples g and s as the top band representing the mutation allele is missing. Amplification failure was seen for sample w.

Table 3.27: Assessment of triplex PCR developed for PGD of the *BRCA1* c.3339T>G mutation from 50 heterozygous and 10 homozygous lymphocytes

Loci	*BRCA1* c.3339T>G mutation	D17S1338 linked marker	D13S168 unlinked marker
AE	98% (49/50)	92% (46/50)	96% (48/50)
ADO	6% (3/49)	7% (3/46)	2% (1/48)
F.N.	0% (0/46)		
F.P.	0% (0/10)		

AE: amplification efficiency; ADO: allele dropout; F.N.: false negative; F.P.: false positive

The expected F.N. and F.P. rates for diagnosis based on two cells are equal to $(F.N.)^2$ and $(F.P.)^2$; both were 0% (0×0) for this protocol.

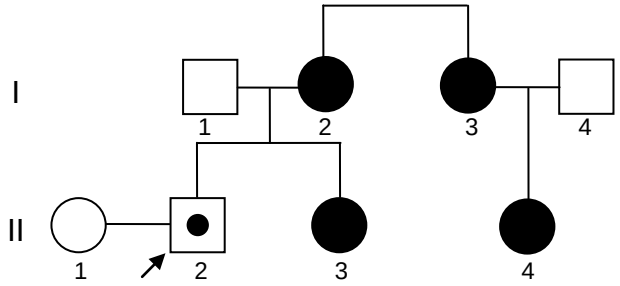
A license for treatment was obtained by the HFEA for this PGD protocol. However, it was never clinically applied as the female partner, aged 34 with no known family history of breast or ovarian cancer, was diagnosed with breast cancer and had to undergo treatment.

3.3.3 PGD workup for *BRCA1* (c.68-69delAG)

3.3.3.1 Patient description

The male partner (proband), aged 29, carried the 2 base pair (bp) deletion c.68_69delAG (formerly known as c.187_188delAG) in exon 2 of the *BRCA1* gene. The proband's wife was homozygous normal. His mother and sister developed breast cancer at the ages of 48 and 26, respectively. His maternal aunt developed ovarian cancer in her 40s and her daughter was diagnosed with breast cancer at the age of 30. Figure 3.64 shows the pedigree. Bloods for DNA extraction were available from the proband and his wife only.

Figure 3.64: Pedigree showing the asymptomatic proband and his affected relatives



II.2 is the proband, carrier of the 2bp deletion in exon 2 of the *BRCA1* gene. I.2, I.3 and II.4 represent the proband's affected mother, maternal aunt and cousin, respectively.

3.3.3.2 Direct mutation detection using F-PCR

A set of fluorescently labelled PCR primers (c.68_69delAG) flanking the 2bp deletion was designed for direct detection of the mutation using F-PCR (Figure 3.65). These primers were initially used to confirm the mutation specified in the patient's genetic report.

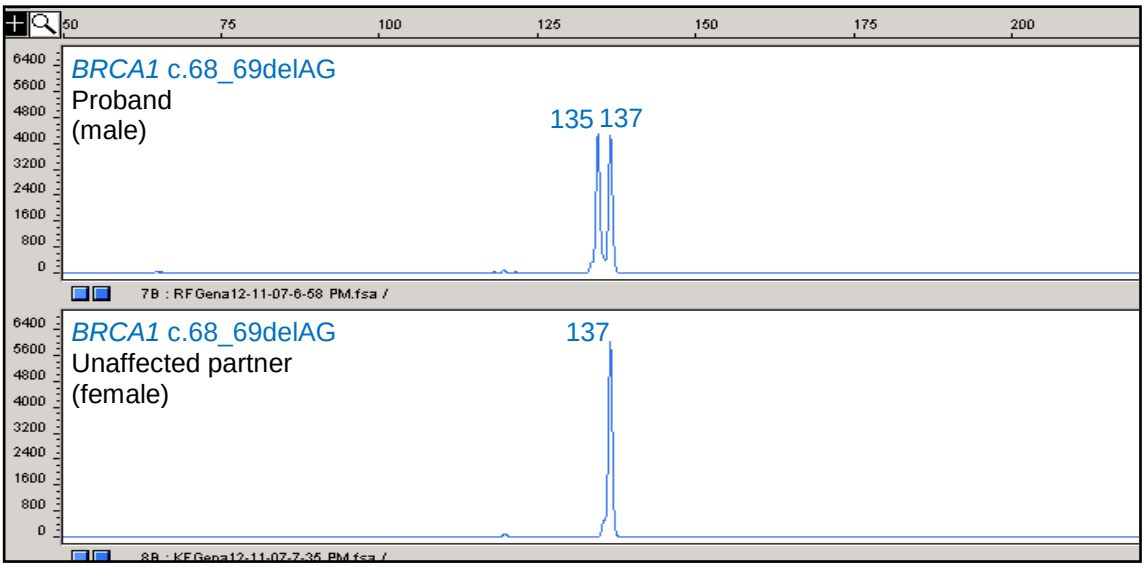
Figure 3.65: Genomic DNA sequence of the *BRCA1* gene (reference sequence L78833, residues 4441-4981) showing the primers designed for the c.68_69delAG mutation

```
4441 acaaaaagca acttctagaa tctttaaaaa taaaggacgt tgtcattagt tctttggttt
4501 gtattattct aaaaccttcc aaatcttaaa tttactttat tttaaaatga taaaatgaag
4561 ttgtcatttt ataaaccttt taaaaagata tatatatatg tttttctaat gtgttaaagt
4621 tcattggaac agaaagaaat ggatttatct gctcttcgcg ttgaagaagt acaaaatgtc   Exon 2
4681 attaatgcta tgcagaaaat cttagagtgt cccatctggt aagtcagcac aagagtgtat   4620-4718
4741 taatttggga ttcctatgat tatctcctat gcaaatgaac agaattgacc ttacatacta
4801 gggaagaaaa gacatgtcta gtaagattag gctattgtaa ttgctgattt ccttaactga
4861 agaactttaa aaatatagaa aatgattcct tgttctccat ccactctgcc tctcccactc
4921 ctctcctttt caacacaaat cctgtggtcc gggaaagaca gggactctgt cttgattggt
```

The sequence of exon 2 is shown in blue; the two bases at the deletion site are underlined. The forward and reverse primers are highlighted in yellow and green, respectively.

After initial PCRs confirmed successful DNA extraction, the asymptomatic affected male and his partner's DNAs were amplified at the mutation site and the products were assessed on the genetic analyzer (Figure 3.66). Fragment length analysis allowed the detection of the mutation when an extra peak that was 2bp smaller than the normal allele size was present; this was indicative of a heterozygous carrier of the mutation. The results confirmed that the proband carried the c.68_69delAG mutation and that his wife was homozygous normal for that locus.

Figure 3.66: ABI Prism™ panels showing detection of the *BRCA1* c.68_69delAG mutation



Fragment length analysis allowed the detection of the c.68_69delAG mutation when an extra peak (135bp), which was 2bp smaller than the normal allele size (137bp), was observed. The proband was 135/137 heterozygous, indicative of a heterozygous carrier of the 2bp deletion, while his wife was homozygous normal (137bp allele only).

3.3.3.3 Informativity tests

Informativity tests were carried out on genomic DNA from the affected male and his partner. Four linked and twelve unlinked STR markers were investigated using F-PCR followed by fragment length analysis on the ABI Prism™ (Table 3.28).

Table 3.28: Allele sizes for different STR markers identified for the couple undergoing PGD for the *BRCA1* c.68_69delAG mutation

Primer name	STR marker location	Female partner (bp)	Male partner (proband) (bp)	Informativity
Linked STR markers				
D17S1338	17q21.31	179/189	*177/179*	Semi-informative
D17S1185*	**17q21.31**	**213**	***202/217***	**Informative**
D17S855	17q21.31; *BRCA1* intron 11	153/160	*153*	Not informative
D17S1343	17q21.31	155/158	*155*	Semi-informative
Unlinked STR markers				
APOC2	19q13.32	152/155	*150/152*	Semi-informative
DM	19q13.32	121	*143/150*	Not informative
D19S112	19q13.32	124/132	*124/127*	Semi-informative
D13S168	13q14.2	126/147	*126/147*	Not informative
Rbivs20	13q14.2	172/184	*164/172*	Semi-informative
IVS26-2.3	17q11.1	212	*207/212*	Not informative
D17S1294*	**17q11.2**	**244/248**	***252/256***	**Informative**
D17S841	17q11.1	262/264	*264*	Not informative
D17S1800	**17q11.1**	**271/276**	***274***	**Informative**
D2S119	2p21	219	*219/221*	Not informative
D2S391	2p21	141	*141*	Not informative
D2S2227	2p21	151/153	*153/164*	Semi-informative

The affected individual is represented in italics. Bold text signifies that the alleles are informative and underlined text indicates the phase allele. The markers in blue are linked with the *BRCA1* gene. The asterisk denotes the markers that were selected for use in the multiplex PCR reaction.

Amongst the linked markers tested, only D17S1185 was found to be informative for this couple and the female partner was homozygous. Both the D17S1294 and D17S1800 unlinked markers could be used as contamination markers. However, as the male partner was homozygous for D17S1800, the fully informative D17S1294 was selected as the contamination marker. This facilitated multiplex PCR optimisations as the affected male was heterozygous for the mutation and both selected markers: D17S1185 and D17S1294.

3.3.3.4 Optimisation of PCR protocol

3.3.3.4.1 Multiplex PCR optimisation using genomic DNA

A triplex PCR was initially tested at the conditions listed in Table 3.29 to amplify two STR markers (D17S1185 and D17S1294) and the c.68_69delAG mutation loci. These initial conditions were selected based on a previously optimised PCR protocol involving the same mutation primers and linked marker (D17S1185).

Table 3.29: Summary of initial triplex PCR tested for the *BRCA1* c.68_69delAG PGD protocol

Reagent	Mutation primers	D17S1185 primers	D17S1294 primers	$MgCl_2$	10% glycerol	Ta
Concentration	0.2µM	0.2µM	0.2µM	1.5mM	Yes	60ºC

Fragment length analysis showed that the mutation and D17S1294 loci were amplified efficiently; however, the peaks for the linked marker D17S1185 were very low. A split PCR was thus used to amplify the D17S1185 loci separately in the second round. The same conditions were used in the fisrt round triplex reaction and the second round PCRs, except no glycerol was added to the singleplex (D17S1185) PCR. This PCR protocol (Table 3.30) was tested on the couple's genomic DNA and five single cells from the proband.

Table 3.30: Summary of split PCR protocol tested on genomic DNA and 5 single cells

Split PCR	Mutation primers	D17S1294 primers	D17S1185 primers	$[MgCl_2]$	10% glycerol	Ta
Round 1	0.2µM	0.2µM	0.2µM	1.5mM	Yes	60ºC
Round 2 a	0.2µM	0.2µM	-	1.5mM	Yes	60ºC
Round 2 b	-	-	0.2µM	1.5mm	No	60ºC
AE	5/5	2/5	0/5			
ADO	1/5	0/2	-			

AE: amplification efficiency; ADO: allele dropout

Amplification was efficient at all loci from genomic DNAs; however, only the mutation locus had amplified efficiently from single cells and ADO of the normal allele (137bp) was seen for one cell.

3.3.3.4.2 Multiplex PCR optimisation for single cells

In order to improve the amplification efficiencies (AE) for the two STR markers, the concentration of these primers was increased to 0.3µM in the first round PCR, while keeping the mutation primers at 0.2µM. Additionally, the concentration of D17S1294 primers was increased to 0.4µM in the second round duplex PCR (mutation + D17S1294). The modified PCR protocol (Table 3.31) was tested on the couple's genomic DNA and five single cells from the proband. This improved AE for D17S1294, however, the D17S1185 loci still failed to amplify from single cells.

Table 3.31: Summary of split PCR protocol tested on genomic DNA and 5 single cells

Split PCR	Mutation primers	D17S1294 primers	D17S1185 primers	[$MgCl_2$]	10% glycerol	Ta
Round 1	0.2µM	0.3µM	0.3µM	1.5mM	Yes	60ºC
Round 2 a	0.2µM	0.4µM	-	1.5mM	Yes	60ºC
Round 2 b	-	-	0.2µM	1.5mm	No	60ºC
AE	5/5	5/5	0/5			
ADO	0%	0%	-			

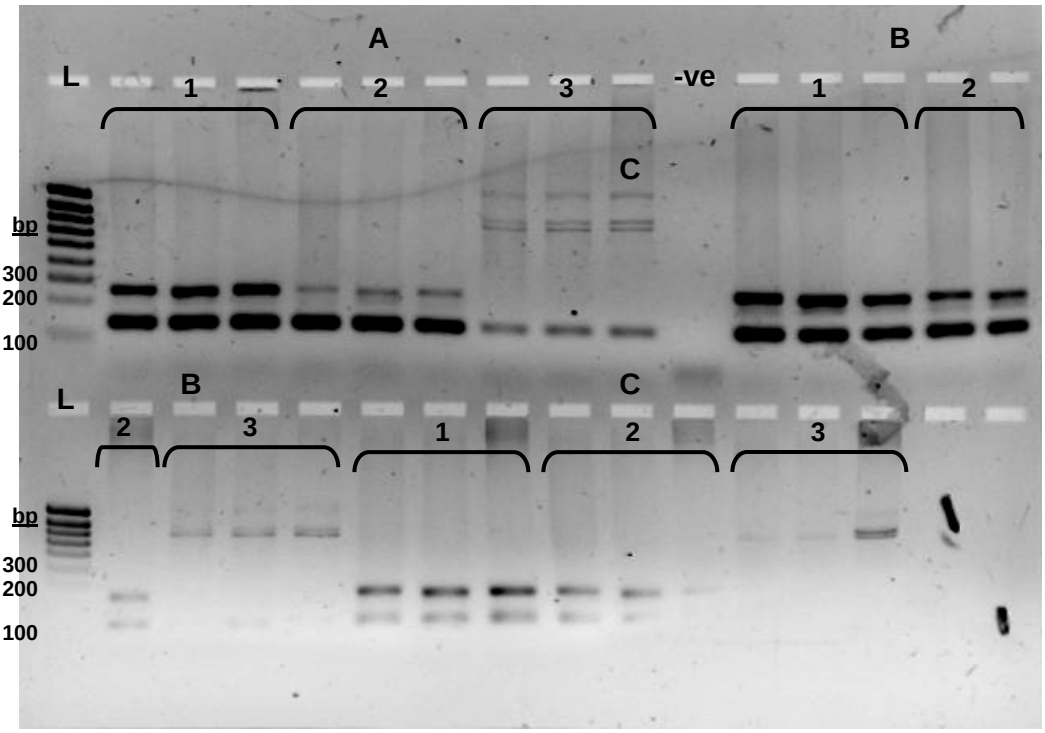
AE: amplification efficiency; ADO: allele dropout

A magnesium/temperature gradient PCR was set up to assess the amplification efficiencies with 1.5, 2.0 and 2.5mM $MgCl_2$ at annealing temperatures of 58, 60 and 62ºC for the first round triplex reaction using genomic DNA. All reagents concentrations (other than $MgCl_2$) were kept the same (Table 3.32). Analysis on the ABI Prism and on a 2% agarose gel (Figure 3.67) showed that the best conditions for all loci were 1.5mM $MgCl_2$ and Ta=60ºC (initial conditions) and 2.5mM $MgCl_2$ with Ta=58ºC, particularly for D17S1185.

Table 3.32: Optimisation of first round triplex PCR using genomic DNA

PCR conditions	Mutation primers	D17S1294 primers	D17S1185 primers	[$MgCl_2$]	10% glycerol	Ta
A1	0.2µM	0.3µM	0.3µM	1.5mM	Yes	58ºC
A2	0.2µM	0.3µM	0.3µM	1.5mM	Yes	60ºC
A3	0.2µM	0.3µM	0.3µM	1.5mM	Yes	62ºC
B1	0.2µM	0.3µM	0.3µM	2.0mM	Yes	58ºC
B2	0.2µM	0.3µM	0.3µM	2.0mM	Yes	60ºC
B3	0.2µM	0.3µM	0.3µM	2.0mM	Yes	62ºC
C1	0.2µM	0.3µM	0.3µM	2.5mM	Yes	58ºC
C2	0.2µM	0.3µM	0.3µM	2.5mM	Yes	60ºC
C3	0.2µM	0.3µM	0.3µM	2.5mm	Yes	62ºC

Figure 3.67: 2% agarose gel (run at 70V) showing products of first round triplex PCR amplified at the conditions listed in table 3.32



L: 1Kb ladder; -ve: negative control; Expected product sizes: 135/137bp for the mutation, ~ 200-220bp for D17S1185 and ~ 250-260bp for D17S1294
Three replicates representing amplified genomic DNA from the proband, his wife and a control (in this order) were run for each protocol tested. Based on the band intensities observed on the gel, procols A1, B1 and C1, which had Ta=58ºC, gave the best results. A1 and B1 gave similar amplifications at all loci; however, protocol C1 seemed to favour amplification of the D17S1294 loci, which gives the larger product.

The first and second round split PCRs were set up with 2.5mM $MgCl_2$ and Ta=58ºC (Table 3.33). Results from 10 single lymphocytes (five from each partner) showed 100% AE and 0% ADO for all three loci. Nonetheless, amplification levels were still much lower for the D17S1185 marker.

Table 3.33: Summary of split PCR protocol tested on 10 single cells

Split PCR	Mutation primers	D17S1294 primers	D17S1185 primers	[$MgCl_2$]	10% glycerol	Ta
Round 1	0.2µM	0.3µM	0.3µM	2.5mM	Yes	58ºC
Round 2 a	0.2µM	0.4µM	-	2.5mM	Yes	58ºC
Round 2 b	-	-	0.2µM	2.5mm	No	58ºC
AE	10/10	10/10	10/10			
ADO	0%	0%	0%			
Comment			Weak amplification			

AE: amplification efficiency; ADO: allele dropout
Low peaks were observed on the ABI Prism™ when amplification was weak.

The second round singleplex PCR was optimised by increasing the primer concentration to 0.3µM and comparing AE at two annealing temperatures: 58 and 60ºC (Table 3.34). Greater amplification levels were obtained for 60ºC annealing temperature.

Table 3.34: Optimisation of the second round of split PCR protocol tested on 10 single cells

Split PCR	Mutation primers	D17S1294 primers	D17S1185 primers	[$MgCl_2$]	10% glycerol	Ta
Round 1	0.2µM	0.3µM	0.3µM	2.5mM	Yes	58ºC
Round 2 b (i)	-	-	0.3µM	2.5mm	No	58ºC
Round 2 b (ii)	-	-	0.3µM	2.5mm	No	60ºC

D17S1294	AE	ADO	Comment
Round 2 Ta=58ºC	10/10	0%	Weaker amplification (lower peaks on ABI Prism™)
Round 2 Ta=60ºC	10/10	0%	Greater amplification (higher peaks on ABI Prism™)

AE: amplification efficiency; ADO: allele dropout

These conditions were used in the final PGD protocol, which is summarised in Table 3.35.

Table 3.35: Summary of optimised split PCR protocol for the *BRCA1* c.68_69delAG mutation

Split PCR	Mutation primers	D17S1294 primers	D17S1185 primers	[$MgCl_2$]	10% glycerol	Ta
Round 1	0.2µM	0.3µM	0.3µM	2.5mM	Yes	58ºC
Round 2 a	0.2µM	0.4µM	-	2.5mM	Yes	58ºC
Round 2 b	-	-	0.3µM	2.5mm	No	60ºC

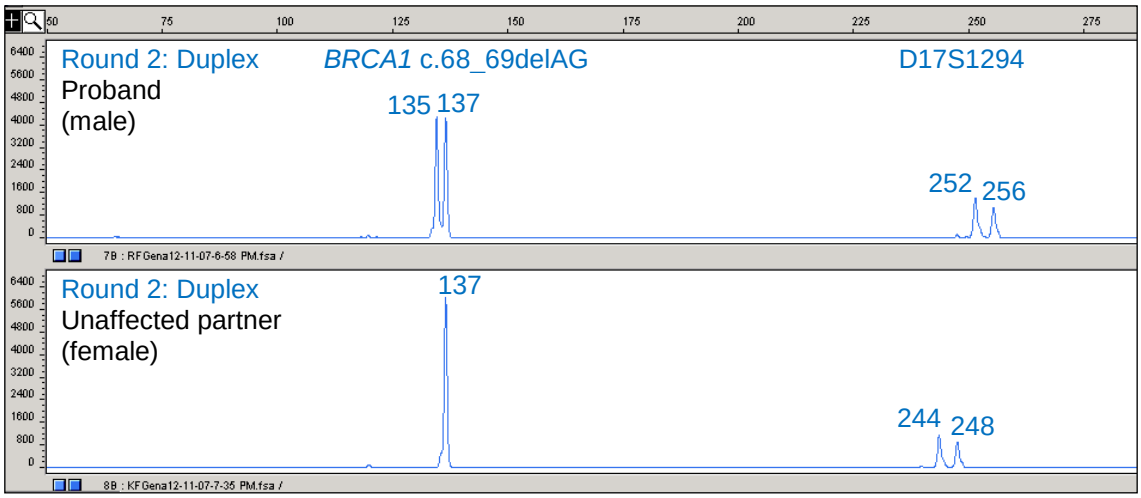
Fragment length analysis of all amplified products was carried out on the ABI genetic analyzer.

3.3.3.5 Final assessment of PGD protocol

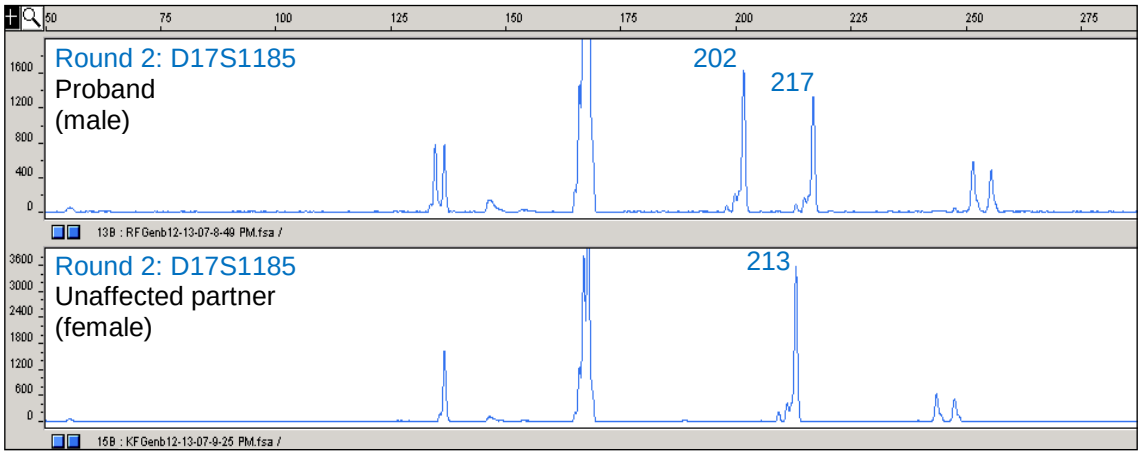
The optimised protocol, shown in Table 3.35, was assessed on 50 lymphocytes that were heterozygous for all loci including the mutation (Figure 3.68). The lymphocytes were isolated from the affected male partner's blood. The outcome is summarised in Table 3.36.

Figure 3.68: Analysis of F-PCR products from single cells obtained with the final PGD protocol: Triplex fluorescent PCR in two rounds of amplification

a) Genetic analyzer panels showing the products of round 2 duplex PCR: *BRCA1* mutation + D17S1294 contamination marker



b) Genetic analyzer panels showing the products of round 2 PCR: D17S1185 linked marker



N.B.: The 168bp peak was constantly observed in all samples amplified with the D17S1185 primers and it was ignored during sizing of alleles.

Table 3.36: Assessment of triplex PCR developed for PGD of the *BRCA1* c.68_69delAG mutation from 50 heterozygous and 10 homozygous lymphocytes

Loci	***BRCA1*** c.68_69delAG **mutation**	**D17S1294 unlinked marker**	**D17S1185 linked marker**
AE	98% (49/50)	96% (48/50)	90% (45/50)
ADO	2% (1/49)	0% (0/48)	9% (4/45)
F.N.	0% (0/45)		
F.P.	0% (0/10)		

AE: amplification efficiency; ADO: allele dropout; F.N.: false negative; F.P.: false positive

The expected F.N. and F.P. rates for diagnosis based on two cells are equal to $(F.N.)^2$ and $(F.P.)^2$; both were 0% (0×0) for this protocol.

This protocol was licensed for treatment by the HFEA and was clinically applied.

3.3.3.6 PGD treatment cycle

Thirteen oocytes were collected and 11 fertilised after ICSI. Two blastomeres were biopsied from each of the 11 embryos. Three cells were collected from embryo number 11 as two cells came out together during biopsy of the second cell. The individual blastomeres were tubed in alkaline lysis buffer (ALB) and transported back to the UCL Centre for PGD laboratory (on a cold rack) for analysis using the PCR protocol described in Table 3.35. The results are presented in Table 3.37.

Five embryos were found to be 'normal' (i.e. free from the *BRCA1* c.68_69delAG mutation) based on the alleles observed for all three markers in two/three cells. Five embryos were found to be 'affected' with increased cancer risk (i.e. carrying the *BRCA1* c.68_69delAG mutation); diagnosis was based on a single cell for three of these embryos (3, 5 and 9).

Table 3.37: Results of the clinical PGD cycle for the *BRCA1* c.68_69delAG mutation

Embryo number	c.68_69delAG mutation	D17S1185 linked	D17S1294 unlinked	Diagnosis	Basis of diagnosis
1	135/- (ADO) 135/137	213/217 213/217	248/252 248/252	Non transferable	2 cells affected
2	**137/137** **137/137**	**202/213** **202/213**	**248/252** **248/252**	**Transferable**	**2 cells normal**
3	(ADO) -/137 -	213/217 -	248/252 -	Non transferable	one cell affected no result from other
4	**137/137** **137/137**	**202/213** **202/213**	**244/256** **244/256**	**Transferable**	**2 cells normal**
5	135/137 -	213/217 -	248/252 -	Non transferable	one cell affected no result from other
6	**137/137** **137/137**	**202/213** **202/213**	**248/256** **248/256**	**Transferable**	**2 cells normal**
7	137/- 137/-	-/213 202/-	248/- -/256	Non transferable	2 cells inconclusive
8	135/137 135/137	213/217 213/217	244/252 244/252	Non transferable	2 cells affected
9	135/137 -	213/217 -	244/252 -	Non transferable	one cell affected no result from other
10	**137/137** **137/137**	**202/213** **202/213**	**248/256** **248/256**	**Transferable**	**2 cells normal**
11	**137/137** **137/137** **137/137**	**202/213** **202/213** **202/213**	**244/256** **244/256** **244/256**	**Transferable**	**3 cells normal**

Maternal alleles are shown in red and paternal alleles in blue. Detection of the 135bp mutation allele indicated that the blastomere/embryo carries the 2bp deletion and is thus 'affected' with increased cancer risk. The mutation was in phase with the 252bp allele for the D17S1294 marker. Transferable embryos are displayed in bold text.

The diagnosed cell from embryo 3 showed paternal phase alleles and was noted as a binucleate at time of biopsy. The cell from embryo 9 that did not give any results was very fragmented and micronuclei were seen. The results were inconclusive for embryo number 7 as both cells appeared to be haploid. One cell showed maternal alleles only and the other cell showed paternal alleles only. At the time of biopsy, it was noted that the second cell from embryo 7 was a binucleate.

A maximum of two embryos can be transferred to the patient's uterus. Selection of the embryos to transfer among the five diagnosed as 'normal' (embryos 2, 4, 6, 10 and 11) was based on morphology and growth rate. Table 3.38 summarises the status of the 11 embryos on days 3 and 5 post fertilisation.

Table 3.38: Description of the 11 embryos from the clinical PGD cycle for the *BRCA1* c.68_69delAG mutation on days 3 and 5 post fertilisation

Embryo number	Day3 stage & grade		Day 5 stage & grade		Diagnosis	Embryo fate
1	11 cells	1-	morula		Affected	research
2	**7 cells**	**2+**	**cavitating morula**		**Unaffected**	**cryopreserved**
3	6 cells	2+/2	vacuolated embryo with disintegrating cells		Affected	research
4	**8 cells**	**1-**	**hatching blastocyst**		**Unaffected**	**cryopreserved**
5	8 cells	1-/2+	6 cells		Affected	research
6	**8 cells**	**1-**	**hatching blastocyst**		**Unaffected**	**transferred**
7	7/8 cells	2+	5 cells	2	Affected	research
8	8 cells	1-/2+	6 cells	3	Affected	research
9	8 cells	2	4 cells	2+/2	Affected	research
10	**8 cells**	**1-**	**hatching blastocyst**		**Unaffected**	**transferred**
11	**8 cells**	**1-**	**fragmented morula**		**Unaffected**	**research**

The 'normal'/transferable embryos are displayed in bold text.

Among the five transferable embryos, embryos 4, 6 and 10 were hatching blastocysts by day 5. Embryos 6 and 10 were selected for transfer. Embryos 2 (cavitating morula) and 4 were vitrified for possible future transfers for this couple. Embryo 11, which was a fragmented morula, was donated for research. All 'affected' embryos appeared to have slower or arrested growth between day 3 and day 5 compared to 'unaffected' embryos. The spare embryos were collected in RNasin solution for research after blastomeres were tubed for confirmation of diagnosis.

This PGD cycle resulted in the birth of a healthy singleton. This was the first baby to be born in the UK after PGD treatment for a *BRCA1* mutation.

3.3.4 Summary of PGD workups

PGD protocols were developed and optimized for two *BRCA1* mutations (*c.3339T>G* and *c.68-69delAG*) and one *MSH2* mutation (*c.1277-?_1386+?del*).

***MSH2 c.1277-?_1386+?del* mutation**

- The developed PGD protocol involved F-PCR fragment length analysis for two linked semi-informative STR markers (D2S2227 and D2S119) and a fully informative contamination marker (APOC2) as well as mini-sequencing for a semi-informative SNP (rs1981929).
- This protocol was licensed for treatment by the HFEA but did reach clinical application due to funding issues.
- This protocol may be used (or adjusted) for future couples seeking treatment for the same condition.

***BRCA1 c.3339T>G* mutation:**

- The developed PGD protocol involved F-PCR fragment length analysis for one linked (D17S1338) and one unlinked (D13S168) informative STR markers as well as direct mutation detection on SSCP.
- The protocol was authorized for treatment by the HFEA but was not clinically applied as the female partner, who had no known family history of breast or ovarian cancer, was diagnosed with breast cancer and had to undergo treatment.

***BRCA1 c.68-69delAG* mutation:**

- The developed PGD protocol involved F-PCR fragment length analysis for the mutation and two informative STR markers (one linked, D17S1185, and one unlinked, D17S1294).
- The protocol was authorized for treatment by the HFEA and was clinically applied, resulting in the birth of a healthy singleton.
- The ‘affected’ embryos obtained from this PGD cycle showed slower growth between day 3 and day 5 post fertilisation compared to unaffected embryos.

4 DISCUSSION

4.1 Expression analysis of DNA repair genes in human oocytes and embryos using microarrays

The first aim of this project was to obtain an expression profile of DNA repair genes in human MII oocytes and *in vitro* derived blastocysts using microarrays in order to investigate two hypotheses: 1) the human oocyte expresses most DNA repair genes to support the early preimplantation embryo and limit DNA damage prior to embryonic genome activation; 2) the expression profile of DNA repair genes in the blastocyst may be different to oocytes due to the high rate of replication and the onset of differentiation.

4.1.1 General expression analysis

The RNA samples extracted from three sets of blastocysts and three sets of oocytes were successfully amplified, hybridised on the microarrays and scanned for analysis. The distinct expression profiles observed using hierarchical clustering (Figure 3.3) and by examining the number of genes expressed at high, medium and low levels (section 2.6.1.2) for each individual chromosome (Figure 3.4) showed that a greater number of genes were expressed at high levels in the blastocyst group compared to the oocyte group. The proportions of transcripts that were expressed at high levels were 25.5% (2,989/11,734) and 34.1% (4,471/13,118) in the MII oocytes and blastocysts, respectively. The investigation of the numbers of genes expressed on each chromosome allowed further assessment of the hybridisation of our samples on the arrays. Global gene expression analysis across chromosomes may allow the determination of relative expression patterns that are characteristic of tissue type and possibly stage of preimplantation development.

A large number of genes was expressed at higher levels in blastocysts compared to oocytes. Eleven out of the 12 commonly selected housekeeping genes investigated were detected in both oocytes and blastocysts and five out of the eight differentially expressed genes (63%) showed a greater than three fold increase in blastocysts compared to oocytes. This observation remained consistent when a larger number of genes were investigated. Analysis of 560 housekeeping genes conducted by Dr Georgia Kakourou using data obtained from the same samples revealed that 211 genes were differentially expressed ($p<0.05$) and 155/211 (73%) had greater expression levels in blastocysts compared to oocytes.

The *in vitro* culture of oocytes and embryos can affect mRNA transcript levels. The effect is largely dependent on the culture medium and conditions of incubation (Rizos *et al.*, 2002; Watson *et al.*, 2000; Wrenzycki *et al.*, 2001). Several studies have shown that *in vitro* culture influences the expression of developmentally important genes in mammalian MII oocytes (Katz-Jaffe *et al.*, 2005b; Lonergan *et al.*, 2003a) and blastocysts (Fernandez-Gonzalez *et al.*, 2009; Lonergan *et al.*, 2003b; Natale *et al.*, 2001) leading to increased apoptosis and a reduced number of cells compared to *in vivo* derived embryos (Chandrakanthan *et al.*, 2006).

Microarray analysis of *in vivo* and *in vitro* produced bovine blastocysts demonstrated that the extent of the change in expression levels for most of the genes was small and most differentially expressed genes were up-regulated under *in vitro* culture conditions (Fernandez-Gonzalez *et al.*, 2009). The up-regulation of the genome expression *in vitro*, which suggests loss of the epigenetic mark (Fernandez-Gonzalez *et al.*, 2009), agrees with the 'quiet embryo' hypothesis (Leese, 2002) that suggests that embryos exhibit increased cellular metabolic activity under stress.

Katz-Jaffe *et al.* (2009) compared the expression patterns of *in vivo* and *in vitro* matured (in special media for over 20 hours) bovine oocytes. Out of approximately 12,000 genes investigated, 10 genes affecting mainly imprinting were found to be differentially expressed (Katz-Jaffe *et al.*, 2009). The MII oocytes used in our study were kept in culture for 4 hours. Therefore, the influence of culture on the mRNA transcript levels is expected to be considerably smaller than that of *in vitro* maturation. The availability of human oocytes that are optimal for clinical use is severely limited and the use of oocytes with delayed maturation, like the ones used in this study, may overcome some of the effects reported in studies that have used *in vitro* matured oocytes.

Freezing and thawing of the blastocysts used in this study may have affected mRNA transcript levels. There is little evidence showing the effect of freezing on expression in mammalian blastocysts. One study showed that thawed vitrified bovine blastocysts exhibit increased fragmentation rates and higher expression levels of apoptosis related genes (*survivin*, *Fas*, *Caspase*-3, and *Hsp70*) compared to non-frozen embryos (Park *et al.*, 2006). Additionally, an increased incidence of chromosomally chaotic embryos was observed among frozen-thawed human blastocysts assessed after they had undergone cell divisions compared to immediately after thawing (Salumets *et al.*, 2004).

4.1.2 Expression of DNA repair genes in MII oocytes and blastocysts

All DNA repair pathways were represented in human MII oocytes and blastocysts and distinct expression patterns were observed for the two groups. While most differentially expressed genes had higher expression levels in the blastocyst compared to the oocyte, most differentially expressed DNA repair genes (73%) were detected in lower levels in the blastocyst compared to the oocyte. This may be because the oocyte must have sufficient mRNA templates to support genome integrity until embryonic genome activation (EGA).

4.1.2.1 Base excision repair (BER)

UNG, *APEX1* and *POLB* mRNAs were detected at high levels in both human oocytes and blastocysts. This agrees with microarray data from another group reporting that *UNG*, *APEX1* and *POLB* are highly expressed in human oocytes at the GV stage (Menezo *et al.*, 2007) given that the mRNA content in the oocyte does not change between the GV and the MII stage as there is no transcription at the final stages of oocyte maturation (El Mouatassim *et al.*, 1999). Degradation of some mRNA transcripts may occur, which may explain the lower number of transcripts picked up in MII oocytes versus GVs (Wells and Patrizio, 2008). El-Mouatassim *et al.* (2007) detected *APEX1* expression in human GV oocytes, 2-cell embryos, morulas and blastocysts and in lower levels in the MII oocytes using real-time PCR analysis. Our data showed that *OGG1* mRNA was detected at medium levels in the MII oocytes and low levels in the blastocysts; Menezo *et al.* (2007) also detected medium expression levels of *OGG1* in human GV oocytes. The low expression levels of *OGG1* mRNA in the blastocyst may indicate that despite the high levels of *APEX1* mRNA, and *POLB* (which plays other key roles in the cell), BER of 8-oxoguanine residues is probably not as important in the blastocyst as it is in the oocyte. In mouse embryos the expression of most BER genes were detected, however *OGG1* and *UNG* mRNAs were not detected at the morula and blastocyst stages (Zheng *et al.*, 2005). Maynard *et al.* (2008) suggested that the lower levels of 8-oxoguanine in human embryonic stem cells (hESCs) compared to differentiated cells indicated a greater OGG1 repair efficiency in hESCs; however, *OGG1* mRNA levels detected by microarrays were not significantly higher and OGG1 incision activity using a functional assay showed similar levels of activity in extracts from hESCs and differentiated cells (Maynard *et al.*, 2008).

One possible explanation proposed by Maynard *et al.* was that hESCs possess higher antioxidant activity, which leads to slower buildup of 8-oxoguanine lesions. Lower levels of ROS were shown to be maintained in mouse and human ESC due to higher expression of major antioxidant genes (such as thioredoxin and glutathione reductase and Mn-superoxide dismutase) (Saretzki *et al.*, 2004; Saretzki *et al.*, 2008). The mRNA expression of genes coding for four enzymes involved in the protection against free radicals (Cu-Zn-superoxide dismutase (*Cu-Zn-SOD* or *SOD1*), Mn-superoxide dismutase (*Mn-SOD* or *SOD2*), glutathione peroxidase (*GPX*) and γ-glutamylcysteine synthetase (*GCS*)) was detected in mouse GV and MII oocytes.

Using reverse-transcriptase PCR, all four enzymes were found to be expressed in human oocytes at the MII stage with particularly high levels for *SOD1* but not all were detected at the GV stage due to maturation-specific polyadenylation of transcripts (El Mouatassim *et al.*, 1999). Our results showed that *SOD1*, *SOD2* and *GPX* were highly expressed in human MII oocytes and blastocysts. *GCS*, however, was not detected in MII oocytes and had low expression levels in blastocysts.

4.1.2.2 Double Strand Break Repair (DSBR)

DSBR via the HR pathway seems to be predominantly active in MII oocytes and blastocysts, with NHEJ acting as a backup pathway in the blastocysts. This is expected for the oocytes, since HR is active during M phase, and for the blastocysts, since HR has greater fidelity of DNA repair than NHEJ and thus may be the preferred mechanism for DSBR. The short G1 and G2 phases in rapidly dividing blastomeres support the assumption that HR is the dominant DSBR mechanism in the blastocyst. Both HR and NHEJ were shown to be active in the mouse zygotes; HR being more active in the male pronucleus (Derijck *et al.*, 2008). NHEJ acts as a backup pathway whenever HR is difficult. This could be important in MII oocytes when DSBs occur at highly repetitive sequences which make homology searches difficult and HR can result in translocations (Agarwal *et al.*, 2006).

HR repair in mammalian cells is highly dependent on the MRE11A-RAD50-NBS1 (MRN) complex, which binds DNA, and on the exonuclease EXO1. The expression levels of *MRE11A*, *RAD50*, *NBS1* and *EXO1* mRNAs indicated that DSBR via HR may be more important in human MII oocytes than in blastocysts.

Initiation of HR involves DNA end binding by RAD52, which competes with KU to direct DSBR toward HR rather than NHEJ (Bassing and Alt, 2004; Van Dyck *et al.*, 1999). *RAD52* was highly expressed in MII oocytes and blastocysts. The interaction of BRCA2 with RAD51 also activates repair using HR (via the cyclin-dependent kinase (CDK) regulatory pathway). *RAD51* was expressed at high levels in both groups, however, *BRCA2* was only detected in the MII oocytes (medium level). *BRCA2* expression was previously detected in human GV oocytes at medium levels (Menezo *et al.*, 2007) and in human MII oocytes with lower levels in preimplantation embryos (increasing from 2-cell to the blastocyst stage) (Wells *et al.*, 2005b).

The interaction of FANCD2 or FANCI with BRCA1 also triggers HR via ATR (Ataxia telangiectasia and RAD3 related) regulation. *FANCD2* mRNA expression levels were significantly higher in the blastocyst versus the oocyte group (Table 3.5). *BRCA1* mRNA expression was detected at medium levels in the blastocysts but was not detected in MII oocytes. This is contrary to data by Menezo *et al.* (2007), where *BRCA1* was shown to be expressed at high levels in human GV oocytes, and previous studies revealed that *BRCA1* was expressed in human MII oocytes at higher levels than in blastocysts (Wells *et al.*, 2005b). *BRCA1* mRNA expression was actually detected with medium signal in only one out of the three replicates; the other two had S.N. ratios <3. Further investigation with real-time PCR will be necessary to determine the expression of *BRCA1* in human MII oocytes. Activation of the HR pathway via ATR involves the MRN complex and BRCA1 as well as 53BP1, MDC1 and later on CHEK1 or CHEK2 (Lobrich and Jeggo, 2007). mRNA templates for *ATR*, *53BP1*, *MDC1* and *CHEK1* were detected in human MII oocytes and blastocysts; this suggests that this pathway may be active in both groups.

During S phase, NHEJ and HR proteins can compete over binding of DSBs. It has been suggested that RAD18 and PARP1 may suppress the DNA binding of KU proteins in favour of HR (Saberi *et al.*, 2007). *PARP1* showed higher mRNA expression levels in blastocyts versus oocytes (although not statistically significant) and *RAD18* mRNA was only detected in the blastocysts. RBBP8 (CtIP) is necessary for repair of DSBs by HR in S/G2 cells and requires the recruitment of BRCA1 (Yun and Hiom, 2009). *RBBP8* had significantly higher mRNA expression levels in MII oocytes versus blastocysts. The main activation processes of DSBR via HR may be different in the blastocyst and oocyte, which would explain the differences in expression levels of various genes.

4.1.2.3 Mismatch Repair (MMR)

MMR is the main DNA repair pathway active post DNA replication (G2/M checkpoint). Our data showed that *MSH2* and *MSH6* were both highly expressed in human MII oocytes and blastocysts. Elevated *MSH2* and *MSH6* expression levels have been reported in GV oocytes and it was hypothesised that 8-oxo nucleotides are removed via the MMR pathway rather than the more common *OOG1* pathway (Menezo *et al.*, 2007). High levels of *MSH2* can result in one of two vital functions, leading the cell to repair its DNA or to undergo apoptosis (DeWeese *et al.*, 1998). Significantly higher levels of *MSH2* were detected in mouse embryonic stem cells (mESCs) compared to differentiated cells (Roos *et al.*, 2007). Additionally, *MSH2* seems to be particularly important in embryonic stem cells, reducing mutation levels and preventing the accumulation of errors, through its possibly more critical role in promoting apoptosis in response to stress (Tichy and Stambrook, 2008).

Previous studies reported that all key MMR genes are expressed in rhesus monkey oocytes and embryos (Zheng *et al.*, 2005). However, MMR may be compromised in the rhesus monkey preimplantation embryos after activation of its genome and thus in the blastocyst due to the over-expression of *MSH3* in comparison to *MSH6* and *PMS2,* as well as the lower expression of *MLH1* (Zheng *et al.*, 2005). High expression of *MSH3,* relative to *MSH2* and *MSH6,* can reduce the formation of the MutSα complex by sequestering MSH2 and degrading MSH6 (Marra *et al.*, 1998).

The concentration of MutSα needs to exceed that of MutSβ to ensure efficient single base mismatch repair in human cells, however, relatively low physiological concentration of MutSα are sufficient for efficient repair of insertion/deletion loops (IDLs) (Drummond *et al.*, 1997; Zhang *et al.*, 2005). The lower expression levels of *MSH3* relative to *MSH2* and *MSH6* in human MII oocytes and blastocysts demonstrated that MMR is potentially efficient for single base mismatches and IDLs.

MSH2, *MSH3*, and *PMS1* were expressed at significantly higher levels in the human MII oocytes versus the blastocysts (Table 3.4). This may indicate that while MMR activity may be important in the human blastocyst, it is of greater significance for the oocyte and early preimplantation embryo that avoid apoptosis as cell death at these stages reduces the chances of survival of the embryo.

Recognition of mismatched DNA may lead to activation of apoptosis in the blastocyst. In fact, apoptosis is commonly observed in some cells at the late cleavage or blastocyst stages while there is no indication of apoptosis in the normally developing human embryo prior to compaction (Antczak and van Blerkom, 1999; Haouzi *et al.*, 2008; Hardy, 1999). Embryos seem to require a 'moderate level' of apoptosis for normal development possibly to regulate the cell number and prevent overgrowth of particularly the inner cell mass (Hardy *et al.*, 2003).

Our data showed that the anti-apoptotic genes *BCL-2* and *BCL-W* were both expressed at low levels in MII oocytes and blastocysts and *BCL-XL* was not detected in either group. The pro-apoptotic genes *BAX* and *BAK* were expressed at low and medium levels in MII oocytes and blastocysts, respectively. *BAD* (another pro-apoptotic gene from the BCL-2 family) was only detected in the blastocysts. Previous studies focusing on genes in the BCL-2 family demonstrated that the anti-apoptotic genes *Bcl-2*, *Bcl-w* and *Bcl-x* were expressed at all stages of development in the mouse from GVs to blastocysts while the pro-apoptotic genes *Bax* and *Bad* were found to be expressed at low levels in oocytes (with higher mRNA levels in GVs versus MII oocytes) and expression of both genes increased in the embryo steadily up to the blastocyst stage (Jurisicova *et al.*, 1998).

The expression of most genes belonging to the BCL-2 family has been confirmed (Metcalfe *et al.*, 2004; Warner *et al.*, 1998). Only *BAX* seems to be constitutively expressed while the others have different expression levels throughout development with contradictory patterns reported in the literature (Haouzi *et al.*, 2008; Metcalfe *et al.*, 2004; Warner *et al.*, 1998). A recent study investigated the expression of *Bax* and *Bcl-2* in immature bovine oocytes using real time RT-PCR and revealed that early apoptosis is indicative of the oocyte's developmental competence (Li *et al.*, 2009).

4.1.2.4 Nucleotide Excision Repair (NER)

The oocyte seems to have the necessary mRNA transcripts for NER. Genes involved in the transcription-coupled repair pathway (TCR), *ERCC6* (*CSB*), *GTF2H1,2* & *5* and *MMS19L*, had medium or high mRNA expression levels in the human MII oocyte. These TCR genes were found to be highly expressed in the human GV oocytes (Menezo *et al.*, 2007). Although there is no transcription before the 4-cell stage (Braude *et al.*, 1988; Telford *et al.*, 1990), the mRNA transcripts laid down for the TCR proteins may be translated and used in the preimplantaion embryo when it begins its own transcription processes.

As NER is active during the G1 phase of the cell cycle, the mRNA templates found in the oocytes are most likely solely stored for translation and used by the early preimplantation embryo. The G1 phase is also short in the rapidly dividing cells of the blastocyst; this may mean that NER is rather limited in these cells as well. In fact, the levels of expression of most NER genes were lower in the blastocyst compared to the oocyte.

4.1.2.5 Other DNA repair pathways

The post-replication repair genes *UBE2A* had significantly higher expression levels in human MII oocytes compared to blastocysts, respectively. *UBE2B* had lower expression levels in oocytes versus blastocysts but the differences in expression levels were not significant. Data from other studies on Rhesus monkey oocytes and embryos also showed that *UBE2A* was mostly expressed as a maternal mRNA and was progressively down-regulated during blastocyst formation, whereas, *UBE2B* mRNA levels increased gradually after the 8-cell stage (Zheng *et al.*, 2005).

4.1.3 Conclusion of DNA repair gene expression from microarray analysis

The expression patterns detected in the oocytes used in this study could be influenced by a variety of factors, including stimulation protocols and culture conditions. For the blastocysts, the effects of cryopreservation also need to be considered. Despite any changes in gene expression that may have occurred, it seems that all DNA repair pathways are potentially functional in human MII oocytes and blastocysts, as a large number of DNA repair genes involved in different repair pathways are expressed at these stages.

Expression levels of DNA repair genes in human MII oocytes and blastocysts suggest differences in DNA repair mechanisms pre and post EGA. Investigation of the BER pathway indicated greater protection against free radicals pre EGA versus post EGA. The main pathway for DSBR appears to be HR but the activation of the pathway may be different between the blastocyst and the oocyte. Expression of the MMR genes in MII oocytes indicates that the pathway is directed towards repair rather than apoptosis. In blastocysts, however, MMR genes were expressed at significantly lower levels than in MII oocytes while the pro-apoptotic genes were expressed at higher levels. The oocyte has all the mRNA transcripts required for NER; however, post EGA, the role of this pathway may be limited as the levels of expression of most NER genes were lower in the blastocyst compared to the oocyte. This is expected, as the G1 phase is short in the cells of the blastocyst.

4.2 Development of a cell free *in vitro* functional assay for IDL and mismatch repair

The second aim of this project was to develop an *in vitro* functional assay for insertion/deletion loop (IDL) and mismatch repair that was more sensitive than available assays and could be applied to cell-free extracts from a limited number of human oocytes or embryos.

4.2.1 Formation of heteroduplex DNA constructs with IDLs and G.T or A.C mismatches

4.2.1.1 Design of DNA templates

4.2.1.1.1 DNA constructs containing insertion/deletion loops (IDLs)

DNA insertion/deletion loops (IDLs) can range in size from one to thousands of nucleotides. Loops that are smaller than 20 nucleotides are mostly formed during DNA replication at repeat sequences and larger loops are created during recombination events (Henderson and Petes, 1992; McCulloch *et al.*, 2003b; Strand *et al.*, 1993). Repair of IDLs can be mismatch repair dependent or independent, depending on the size of the loop. DNA loops that are up to 16 nucleotides in size are processed by MMR dependent and independent pathways; however, IDLs that are greater than 16 nucleotides in size are repaired by MMR independent pathways only (McCulloch *et al.*, 2003a).

The PCR based strategy used for the formation of heteroduplex DNA constructs containing insertion/deletion loops (IDLs) relied on the variable allele sizes at the *DMPK* triplet repeat, which made it possible to generate DNA constructs with different IDL sizes. The *DMPK* (CTG)n repeat is particularly interesting because it is characterised by intergenerational instability that can take place in the oocyte and/or early embryo, however, the mechanism of expansion is still unknown (Savouret *et al.*, 2003). The efficiencies of loop recognition by the MMR complexes (MutSα and MutSβ) vary according to the size of the loop and the extrahelical nucleotides and their flanking sequences (Marti *et al.*, 2002; Palombo *et al.*, 1996). For example, 2 to 4-nucleotide IDLs are preferentially detected by MutSβ; however the GT IDL can be preferentially bound by MutSα in a specific sequence context (Palombo *et al.*, 1996).

Three homozygous samples, with allele sizes of 122, 143 and 146, were used to form heteroduplex DNA contructs with 3, 21 and 24-nucleotide (nt) IDLs. The sizes of the IDLs were selected deliberately to include loops that are smaller and greater than 16 nucleotides in size, in order to assess IDL repair that is dependent or independent of MMR proteins. Additionally, the single CTG insertion (3-nt IDL) was of fundamental interest as understanding its recognition and repair efficiencies by the MMR machinery could shed light on the instability of the *DMPK* repeat during human preimplantation development. It was possible to create heteroduplexes with IDLs of various sizes by simply changing the homozygous samples used.

4.2.1.1.2 DNA constructs containing a G.T or A.C mismatch

The G.T mismatch was selected as the main base-base mismatch substrate because it is the most efficiently repaired by MMR efficient cells, particularly HeLa S3 extracts, followed by the C.A mismatch (Hays *et al.*, 2005; Holmes *et al.*, 1990; Thomas *et al.*, 1991). These mismatches are also known to be bound with greater affinity by MutSα than MutSβ (Marti *et al.*, 2002; Palombo *et al.*, 1996).

Homozygous samples for the rs1981929 SNP site were used for the formation of constructs with a G.T or A.C mismatch. The same approach can be applied with different SNP sites to form heteroduplexes containing G.A or C.T mismatches or several mismatches by targeting a DNA sequence that contains several SNPs within ~ 300bp region in the genome; but it was not possible to form A.A, T.T, G.G or C.C mismatches using this PCR based strategy. Alternatively, DNA strands with the desired sequences and lengths can be purchased (e.g. oligonucleotides listed in Table 2.9). However, synthesis of long oligonucleotide sequences can sometimes be difficult and obtaining the required amounts can be costly.

4.2.1.2 Optimisation of PCR conditions

Amplification of the DMPK $(CTG)_n$ repeat region using the MMR primer sets was difficult due to the large product size (558-639bp) and the high GC (guanine and cytosine) content, which causes the DNA to coil around itself increasing secondary structures that make it less accessible to primers. In order to overcome this problem, 5% DMSO was added. DMSO reduces the formation of stable secondary and tertiary structures in the DNA templates, thus enhancing the PCR amplification by increasing the ratio of full-length products to shortened products (Kang *et al.*, 2005).

4.2.1.3 Formation and detection of heteroduplex DNA constructs

Five heteroduplex DNA constructs were generated in this project: constructs with IDL sizes of 3, 21 and 24 residues and constructs with a G.T or A.C mismatch. The constructs could be easily generated in large amounts and modified to vary the size of the IDL or the base-base mismatch.

Heteroduplex analysis on SSCP gels showed that an overnight incubation at 37ºC was necessary for the mismatched strands to hybridise. Some loss of product was detected at the single strand separation and purification stages. However, recovery rate was not a big issue at that stage since the amount of ssDNA obtained was quite high (as seen from the ABI Prism^{TM} peak heights in Figure 3.22) and the PCR products could be easily generated in large amounts making their loss affordable.

The labelled MMR products (558-639bp long) and the derived ssDNA could not be visualised on SSCP gels. This was probably due to the presence of the fluorescent label that affected the fragments' run properties in the gels. When non-labelled primers were used, the detection of the DNA fragments was significantly improved, especially for duplex DNA structures that could not be initially seen.

Another reason the MMR products were difficult to visualise on SSCP gels was probably the large size of the DNA fragments. SSCP sensitivity is optimal for products up to 200bp in size and the sensitivity is decreased as the fragment size increases (Sheffield *et al.*, 1993). More than 90% of single-base changes can be detected for fragments of about 200bp, whereas the detection rate is lowered to 80% for fragments of about 400bp (Hayashi and Yandell, 1993; Nataraj *et al.*, 1999). Mutations can be detected in fragments as large as 800bp; however, the sensitivity of the assay would probably be lowered (Nataraj *et al.*, 1999). One way to deal with large fragments is to cut the PCR product into smaller fragments prior to SSCP/heteroduplex analysis. In this project, the MMR2S primer was used to create shorter products/heteroduplexes (277-358bp) that included the same target CTG repeat. The formation of tightly coiled tertiary structures in the DNA, caused by the high GC content in the DNA fragments, should however improve the SSCP/heteroduplex analysis sensitivity by making fragments of different sequences more discernible (Nataraj *et al.*, 1999).

The heteroduplexes containing a 3-nucleotide IDL and G.T or A.C mismatch could not be easily distinguished from their corresponding homoduplexes (PCR products) on the SSCP system used (Figures 3.24 and 3.25), however, the nicked heteroduplex molecules were clearly different. The change to the overall structure of the DNA molecules containing a single base mismatch or a 3-nucleotide IDL was relatively subtle compared to that caused by 21 and 24-nucleotide IDLs, making them less discernible from homoduplexes. Generally, insertions and deletions create structural changes in DNA fragments characterised as ‘bulges’ that can cause kinks or bending of the DNA, while single and multiple base substitutions create smaller changes characterised as ‘bubbles’ that are less likely to cause bending (Bhattacharyya and Lilley, 1989). It is expected that larger distortions to the DNA molecule result in greater differential migration in the gel compared to homoduplex DNA than mismatches causing less structural distortions (Highsmith *et al.*, 1999). Kinks in DNA fragments formed by 3-nucleotide ‘bulges’ (80 ± 10^{o}) or ‘gaps’ (different angles) have been visualised by electron microscopy at increased frequencies compared to homoduplex fragments (Wang *et al.*, 1992).

Investigation of different physical parameters influencing sensitivity of mutation/mismatch detection by heteroduplex analysis showed that GC content, fragment length (between 100 and 600bp) and the position of the mismatch (centrally located or 50bp from either end) had little or no effect on sensitivity (Highsmith *et al.*, 1999). The nature of the mismatch was the main factor affecting the differential migration of homoduplexes and heteroduplexes, with G.G/C.C having the greatest separation followed by A.C/T.G and A.G/T.C (equally), and finally A.A/T.T (Highsmith *et al.*, 1999).

Another difficulty faced during the production of heteroduplex constructs from non-labelled PCR products was matching the concentrations of the separated ssDNA fragments in order to avoid having a large excess of a particular strand and promote heteroduplex formation. For this reason, commercially synthesised oligonucleotides containing the same sequences were obtained and prepared at fixed concentrations.

4.2.2 Repair assessment after exposure of heteroduplex constructs to nuclear extracts from HeLa S3 and LoVo cells

Assessment of DNA loop repair using heteroduplex analysis on the PhastSystem™ was difficult for the 3-nucleotide (nt) IDLs, as the difference between homoduplex and heteroduplex DNA molecules was not always obvious. However, minor changes in electrophoretic migration were consistently observed, especially for the heteroduplex samples exposed to HeLa S3 N.E. compared to the original heteroduplex sample and the reaction's negative control. A conformational change from heteroduplex to homoduplex gave a clear indication of DNA repair.

Heteroduplex samples, containing a G.T mismatch or a 3-nucleotide IDL, that were exposed to 2µg HeLa S3 nuclear extracts (N.E.) showed shifts in electrophoretic migrations on SSCP gels that matched those of homoduplex samples. This was indicative of IDL repair. Repair was nick directed in presence of a nick, as expected (Thomas *et al.*, 1991; Hays *et al.*, 2005), but seemed to target either strand in non-nicked heteroduplex constructs.

Sealing of the nick was approximately 100% efficient when nicked heteroduplexes were exposed to nuclear extracts for 15 minutes or longer; however, repair was not as efficient. This situation is referred to as 'escaped' repair and was detected in other studies for G.T substrates (Holmes *et al.*, 1990; Thomas *et al.*, 1991) and CTG insertions with a 3' nick (Panigrahi *et al.*, 2005).

Assessment of IDL repair using semi-quantitative fluorescence analysis (on the ABI Prism^{TM} 310) and heteroduplex analysis showed that repair was independent of loop size (3, 21 and 24 nucleotides) and was detected for exposures to as little as 2μg (0.08μg/μl) of nuclear extracts. Repair was MMR and nick-dependent and nick-directed for 3-nucleotide IDLs. Repair of larger loops was MMR and nick-independent. Generally, greater ΔR values were obtained with 20μg of HeLa S3 nuclear extracts compared to 2μg of HeLa S3 or LoVo. The greatest ΔR value was detected with 3-nucleotide IDL after a 15 minute exposure to 20μg of HeLa S3 nuclear extracts (Table 3.13). Exposures of heteroduplex constructs to LoVo nuclear extracts resulted in less repair compared to HeLa S3 for all IDL sizes. These observations are consistent with previous studies which showed that IDLs can be repaired in a nick-directed manner by the MMR pathway (up to16-nt IDLs) or MMR-independent pathways for larger loops in human extracts (Littman *et al.*, 1999; McCulloch *et al.*, 2003a; McCulloch *et al.*, 2003b).

Interestingly, our results showed that MMR-independent repair of larger loops was not nick-directed when the nicked strand contained the loop (i.e. long strand was nicked). Another study that used DM1 clone substrates containing pure $(CTG)_n$ repeats, with n=30 or 50, reported that CAG/CTG hairpin repair occurs in an error-free manner, when the hairpin is in the continuous strand, and error-prone manner, when the hairpin is in the nicked strand, particularly for CTG versus CAG repeats due to their hairpin versus random coils conformations (Panigrahi *et al.*, 2005). In our study, the insertions in the nicked strand were all CAG repeats. Panigrahi *et al.* suggested that incorrect repair is the result of a faulty excision step that does not result in complete removal of the 'slipped-out' repeats and is followed by gap filling and ligation. The expansion bias that is observed in human trinucleotide repeat diseases may be explained by the fact that repair of insertions or expansion intermediates in the CTG/CAG repeat is error-prone but repair of deletions is not.

A recent study by Hou *et al.* (2009), which investigated CTG/CAG hairpin repair in HeLa nuclear extracts, showed that repair of CTG/CAG hairpins was error-free in a strand-specific and nick-dependent manner. Southern blot analysis of repair intermediates and products showed that repair was initiated by endonucleotlytic incisions targeting the repeat region on the nicked strand, which required PCNA (Hou *et al.*, 2009). The discrepancies between this study and the one by Panigrahi *et al.* may be due to the assay systems used (Southern blotting with ^{12}P-labelled probes versus incorporation of [α-^{32}P]-dNTPs) or differences in sequence contexts. Therefore, it may not be appropriate to suggest a single model to explain trinucleotide repeat expansions, as subtly different repair enzymes may be involved in trinucleotide repeat hairpin repair in different biological contexts (Hou *et al.*, 2009).

PCNA, which is an important factor in DNA replication, has been shown to be necessary for the initiation of MMR (Guo *et al.*, 2004) and hairpin repair (Hou *et al.*, 2009). Our microarray results showed that *PCNA* was highly expressed in human MII oocytes and blastocysts, with higher levels in the oocytes ($p<0.05$).

Exposure time to nuclear extracts affected repair efficiency; however, beyond 60 minutes, repair was unchanged. No substantial exonucleolytic degradation due to free DNA ends was detected, even after long exposures to N.E., in either short (~300bp) or long (~600bp) heteroduplex molecules. Although circular DNA molecules are normally used as substrates for *in vitro* DNA repair reactions, it has been demonstrated that linear molecules are also competent to assess MMR using human nuclear extracts (Iams *et al.*, 2002). Additionally, Iams *et al.*'s study showed that the location of both the nick and the mismatch could affect MMR capacity. If the nick or the mismatch was located too closely to the DNA ends (<50 or 150bp, respectively) no MMR was detected. None of the heteroduplex constructs used in this project had a nick or mismatch location known to jeopardise MMR, however, the specific structure of each construct could affect repair efficiency as the sequence context can influence recognition and repair of DNA lesions.

In our study, brief exposures of heteroduplexes with 3-nucleotide loops to 2µg of nuclear extracts for 30 or 60 seconds resulted in the excision of one strand (which confirmed recognition of the mismatched region and initiation of repair) without resynthesis. This agrees with previous studies that state that excision starts as early as 40 seconds after initiation (Wang and Hays, 2002b) and reaches a plateau within 6 to 7 minutes (Hays *et al.*, 2005). Incomplete repair was also detected for the same substrates on SSCP gels after exposures to 2µg of N.E. for 60 minutes. This could have resulted from insufficient amounts of dNTPs or the necessary proteins for resynthesis of the excised DNA fragments. Interpretation of the genetic analyzer results showed that it was probably insufficient amounts of MMR efficient nuclear extracts that resulted in incomplete or lack of repair. 3-nt IDL repair was MMR and nick-dependent and was detected for exposures to 20µg of HeLa N.E. but never for 2µg exposures using the ΔR-value system.

SSCP/heteroduplex analysis proved to be somewhat problematic for robust assessment of repair; however, it did show conformational changes in constructs exposed to nuclear extracts and allowed the detection of excision as well as complete repair. These changes were consistent and seemed to confirm repair, as all observations were concordant with the obtained R-value results.

Detection of IDL repair was more sensitive on the genetic analyzer, which allowed the detection of relative repair efficiencies using R-values. This type of analysis also helped overcome the issue of incomplete nicking of the heteroduplex constructs. When two-colour labelled nicked heteroduplexes were used, repair was clearly detected by the appearance of a new peak representing the repaired strand (Figure 3.40). However, when both strands of the heteroduplex molecule were FAM-labelled, the relative proportions of short/long strand or nicked/non-nicked strand can change in the absence of a complete DNA repair reaction. Initiation of repair through excision of a fragment on the nicked (or targeted) strand would alter the R-value. While this is fairly indicative of DNA repair activity, it was not interpreted as complete repair.

In order to improve base-base mismatch repair assessment, mini-sequencing analysis could be applied on the G.T and C.A heteroduplex constructs. The sizes of mini-sequencing products are different for each base and visualised in different colours on the ABI Prism™ 310 genetic analyzer. This would allow accurate detection of the base on each strand of the DNA molecule at the SNP/mismatch site. Assessment of repair efficiencies would also be possible based on the peak heights and calculation of R-values.

While the genetic analyzer gave more sensitive detection of repair, it could only give information on single stranded DNA and could not always distinguish excision of the targeted DNA strand from complete repair (i.e. fragment resynthesis). The most informative results were obtained using the initial design, which involved FAM/HEX double-labelled constructs. SSCP/heteroduplex analysis was necessary to detect the formation of duplex DNA constructs with different secondary structures, which act as substrates for DNA repair proteins, and confirm complete repair via the detection of newly formed homoduplex molecules.

This assay was up to 50-fold more sensitive than available methods, however, in order to make it suitable for application on human oocytes and embryos it is crucial to reach much greater sensitivity. The average total protein content in mammalian oocytes is approximately 0.1μg and ranges in preimplantation embryos from 0.16μg (at the 2-cell stage) up to 50μg (at the blastocyst stage) (Grealy *et al.*, 1996; Morgan and Kane, 1993; Thompson *et al.*, 1998). Several MutSα or MutSβ complexes are required to efficiently repair one molecule of mismatched DNA, with a minimum of 4-fold molar excess of MutSα compared to DNA substrates (Zhang *et al.*, 2005). This means that fewer heteroduplex DNA molecules, that would require fewer repair proteins and thus less nuclear extract, need to be used in the assay.

It has been demonstrated that 50µg of HeLa nuclear extracts, which contains ~ 0.1µg MutSα, is necessary to repair more than 50% of 24fmol of G.T mismatched heteroduplex constructs in 15 minutes (Dzantiev *et al.*, 2004; Zhang *et al.*, 2005). This is equivalent to 2µg of nuclear extract per fmol of substrate and is similar to the amounts (~1.33µg per fmol of substrate) used by other studies that assessed the repair of single base mismatches and IDLs in nuclear extracts (Littman *et al.*, 1999; Wang and Hays, 2002a; Wang and Hays, 2002b).

In this project, approximately 1 to 2fmol of substrate was used; bringing this down by a further 10-fold may allow the future assessment of repair in nuclear/whole cell extracts from human oocytes and preimplantation embryos. Initially, pooling of oocytes or embryos will probably be necessary for the extraction of proteins. The developed assay could probably be used with the same system of analysis to begin with; nevertheless, a new system of analysis with greater sensitivity of DNA detection would be valuable to improve assessment of DNA repair in limited amounts of nuclear extracts.

4.2.3 Conclusion of functional assessment of IDL repair and MMR in human nuclear extracts

The main limitation of functional assays at present is that they require large amounts of proteins. In this project a novel approach was used in the development of an *in vitro* functional assay for MMR that may be optimised for analysis of nuclear extracts from pooled oocytes or embryos. The developed assay does not rely on cloning for the generation of substrates and requires neither bacterial cultures nor radioactive labeling (like Southern blotting) for analysis.

Two types of heteroduplex constructs were created using a PCR based method for the assessment of MMR and IDL repair efficiencies: heteroduplexes containing a G.T or A.C mismatch and heteroduplexes containing a 3, 21 or 24-nucleotide insertion/deletion loop. These constructs could be easily generated and modified to vary the size of the insertion/deletion loop or the single base mismatch to be repaired.

The heteroduplex constructs were exposed to nuclear extracts from human cells to repair the mismatch. The efficiency of repair was determined semi-quantitatively by assessing the ratio of heteroduplexes to repaired homoduplexes using the genetic analyzer and qualitatively using SSCP/heteroduplex analysis. The preliminary results obtained from semi-quantitative assessment of IDL repair were in agreement with other studies and demonstrated that the developed assay works. Further optimisations will be necessary to improve the sensitivity of this assay and allow its future application on pooled human oocytes or embryos.

4.3 PGD for *MSH2* and *BRCA1* mutations

The third aim of this project was to investigate the effect of germline mutations in DNA repair genes on early embryonic development. Clinical PGD for *MSH2* and *BRCA1* was initiated in order to achieve this aim.

4.3.1 Development of PGD protocols

Single cell PCR strategies for preimplantation genetic diagnosis generally rely on direct mutation detection with at least one linked polymorphic marker to detect allele dropout (ADO), contamination and recombination (Spits *et al.*, 2007). The use of linked markers alone is becoming increasingly common as it makes it possible to apply the same gene-specific PGD test to many couples with different mutations. However, this would require the couple seeking PGD treatment to be informative for at least two markers flanking the gene of interest and the availability of DNA from affected family members to establish haplotypes (Goossens *et al.*, 2003).

The linked polymorphic markers can be STRs or SNPs. The main advantage of SNPs is that they are densely dispersed within all genes; however, their main disadvantage is that they are less likely to be informative than STRs, which are more polymorphic. This means that potentially a large number of SNPs need to be investigated before finding an informative one. Additionally, analysis of SNPs is more labour intensive than STRs as it requires sequencing/mini-sequencing or restriction enzyme digestion reactions post PCR (Spits *et al.*, 2007).

Three protocols for preimplantation genetic diagnosis were developed for couples with inherited mutations in DNA repair genes, one for *MSH2* and two for *BRCA1* mutations.

4.3.1.1 Workup for MSH2 (c.1277-?_1386+?del)

The *MSH2* mutation carried by the male partner was an exon 8 deletion diagnosed by multiplex ligation-dependent probe amplification (MLPA). MLPA is a relatively recent quantitative technique designed to detect insertions and deletions within exonic sequences using exon-specific oligonucleotides probes (Schouten *et al.*, 2002). MLPA has become the preferred method for rapid genetic screening for insertions and deletions as it is less time consuming and cheaper than other techniques. The main drawback of this technique is that it does not provide specific breakpoints for the insertions or deletions. This can be a problem for designing PGD protocols.

Identification of the deletion breakpoints was attempted by Yasmin Omar and William Young during the course of their MSc and BSc research projects, respectively, at the UCL Centre for PGD laboratory. Their work involved long-range PCRs with primers designed at different locations in introns 7 and 8. While exon 8 is only 109bp long, the two introns surrounding it are very large (~15.5 and 17Kb) and include several Alu repeat sequences. This made the PCR amplifications very difficult in this region. A 14.9Kb deletion between two Alu repeats in introns 7 and 8 reported by Thiffault *et al.* (2004) was suspected; however, it could not be confirmed using the designed primers for long-range PCR.

Southern blot analysis, which would have allowed accurate sizing of the deletion, is currently rarely used as it involves radioactive labeling. Fluorescent *in situ* hybridisation (FISH) and metaphase array comparative genomic hybridisation (CGH) are often used for the detection of large deletions. The resolution of these techniques can be up to 1Mb (Oostlander *et al.*, 2004; Volpi and Bridger, 2008). However, since the maximum size of the deletion including exon8 can be ~32.6Kb, FISH and array CGH would not be suitable in this case.

As the deletion breakpoints were not defined, the PGD protocol could not involve direct mutation detection and the diagnosis had to be based on linkage. Eleven linked STR markers, including one intragenic marker, were investigated for the couple. None were found to be fully informative; two were semi-informative: D2S227 and S2S119. The allele that was in phase with the deletion could be identified for D2S119 only from the affected paternal aunt's DNA. D2S119 was located approximately 3.5Mb upstream of the *MSH2*. As this is outside the targeted 1Mb range, which limits the chances of recombination to 1% (Collins, 2009), PGD could not rely on this STR marker alone. Identifying the phase allele for D2S2227, which is located 0.36Mb upstream of *MSH2*, could help increase the diagnosis rate. It is generally preferable to base the diagnosis of a normal embryo on heterozygous genotype in order to avoid the event of ADO leading to a misdiagnosis. Thus, it would be better if the non-shared (166bp) D2S2227 allele turned out to be in phase with the mutation. However, as two linked markers were used it would have been possible to detect ADO and identify the parental origin of the observed allele.

DNA was not available from any other relatives; however, the haplotye could have been identified from the affected male's sperm if the couple was going ahead with the PGD treatment. The investigation of any additional linked STR markers would have required the use markers that are further away from the gene. Instead, SNP sites were investigated. Although the chances of finding informative loci for SNPs are much lower than for STR markers, a large number of intragenic SNPs are available.

There were no known SNPs within exon 8 of the *MSH2* gene, which is only 109bp long. Five SNPs that were up to ~1kb upstream of exon 8 were investigated (Figure 3.42 and Table 3.17). The rs1981929 SNP was found to be semi-informative, which would confirm 50% of the normal/unaffected embryos. The other SNPs investigated were probably in linkage disequilibrium as all three individuals were homozygous for the four SNPs. This illustrates the main disadvantage of using SNPs as polymorphic markers since this type of SNP characterisation is not yet reported in public databases and can only be discovered empirically (Spits *et al.*, 2007). Adding this problem to the fact that the deletion breakpoints were unknown, the selection process of a fully informative SNP for this case could potentially be extremely time consuming and labour intensive.

The difficulties of finding informative SNPs could have been overcome by the use of SNP microarrays designed to target specific genes or chromosomes on the couples genomic DNA. This would have also helped identify the breakpoints of the exon 8 deletion. However, SNP microarray services can be very expensive and in this case the PGD workup budget could not cover the costs.

The combination of the three linked loci (D2S119, D2S2227 and rs1981929) with the fully informative contamination marker (APOC2) resulted in an acceptable PGD protocol. The PCR based protocol, involving fragment size analysis and mini-sequencing, reached the preliminary stages of optimisation on single cells. The workup was not continued as the couple decided not to go ahead with PGD treatment due to financial issues.

Clearly, this PGD protocol was compromised due to the undetermined breakpoints of the exon 8 deletion. Detection of copy number variation of several SNPs located within the target region using real-time PCR should be investigated as an alternative strategy for marker selection and to help identify the breakpoints of the mutation for future patients diagnosed with large duplications/deletions by MLPA. This will be important for future cases as it is becoming increasingly difficult to find Southern blot analysis services.

This protocol was licensed by the HFEA and initiated the PGD service for Lynch syndrome at the UCL Centre for PGD. Since then, several couples have been referred for PGD treatment for deletions in the *MSH2* and *MLH1* mismatch repair genes. As a license for PGD treatment for *MSH2* has been obtained, the workup time, which is one of the main problems of PGD particularly for patients with genetic predispositions to cancer, will be much shorter for future *MSH2* cases.

4.3.1.2 Workup for BRCA1 (c.3339T>G)

A PCR-based protocol was developed for direct mutation detection by single stranded conformational polymorphism (SSCP) and the detection of ADO at the mutation locus and contamination using a linked and unlinked STR markers. The triplex PCR was optimised on single cells for the amplification of the following loci: *BRCA1* c.3339T>G mutation, D17S1338 (linked marker) and D13S168 (unlinked marker).

The intragenic D17S855 marker was dropped out of the initial quadruplex PCR protocol since it was difficult to optimise and was not strictly necessary for the PGD strategy. D17S13338, the only fully-informative marker for this couple, was located at ~2.5Mb upstream of the *BRCA1* gene. This was not ideal, as it exceeded the guideline 1Mb limit for a 1% recombination risk (Collins, 2009). However, this protocol involved direct mutation detection and was not reliant on the linked marker alone for diagnosis.

The final triplex PCR protocol allowed direct mutation detection with a high amplification efficiency (AE) of 98% and an acceptable allele dropout (ADO) rate of 6%. The linked marker had a 92% AE and 7% ADO rate, while the contamination marker had a 96% AE and 2% ADO rate. In the event of ADO at the mutation site, the linked marker could help identify the allele passed on from the affected male partner. The presence of the intragenic marker or another linked marker on the other side of the mutation to detect recombination would have been extremely helpful in such a situation.

A license for treatment was obtained by the HFEA for this PGD protocol. However, it was never clinically applied as the 34 year old female partner, who had no family history of breast or ovarian cancers, was diagnosed with breast cancer prior to initiation of IVF treatment. It is important to point out that this patient was of Ashkenazi Jewish origin. Her condition may have resulted from a *de novo* mutation in the *BRCA1/2* genes or she may have a genotype that includes specific polymorphisms associated with an increased risk of breast cancer.

It is circumstances like these that compel us to understand why there were so many ethical debates concerning PGD for *BRCA1/2*. PGD for cancer predispositions with 'later' age of onset and incomplete penetrance was not allowed in the UK until 2006. (The word 'later' is used here to imply during adulthood as opposed to early childhood). Several recent papers have discussed the ethical issues involved (Clancy, 2009; Krahn, 2009; Menon *et al.*, 2007; Noble *et al.*, 2008; Williams *et al.*, 2007); all agree on the importance of proper counseling for these patients. It is crucial that patients seeking PGD treatment are well aware of the workup time and limitations of the technique. In the case of breast cancer, PGD is only testing for a specific mutation in the *BRCA1/2* genes. This means that any other genetic factors that could lead to an increased risk of developing breast cancer or other associated cancers are not examined. Such factors include polymorphisms that may be specific to certain populations (Easton *et al.*, 2007;

Song *et al.*, 2009) or the combination of low penetrance genes (Tempfer *et al.*, 2006). At best, PGD can lower the cancer risk to the general population level. The timing of the PGD treatment is critical when the female partner carries the mutation, as the window of opportunity is much smaller for patients who are likely to require prophylactic treatment or cancer therapy.

It is possible that PGD treatment, which involves hormone stimulation for IVF, can exacerbate the condition of patients with a genetic predisposition to breast cancer. A matched case-control study, conducted on 1,380 pairs of women with a *BRCA1* or *BRCA2* mutation, showed that the use of fertility medications does not increase the risk of breast cancer among these patients (Kotsopoulos *et al.*, 2008).

4.3.1.3 Workup for BRCA1 (c.68-69delAG or 185delAG)

A PCR-based protocol was developed for direct mutation detection by fragment length analysis, including a linked and an unlinked STR marker. The triplex PCR was optimised on single cells for the amplification of the following loci: *BRCA1* c.68-69delAG mutation, D17S1185 (linked marker located at ~0.7Mb upstream of the *BRCA1* gene) and D17S1294 (unlinked marker). The fully-informative unlinked marker was necessary for the detection of maternal contamination as the unaffected female partner was homozygous for the linked marker. This protocol was licensed for treatment by the HFEA and was the first clinically applied *BRCA1* PGD protocol in the UK that resulted in the birth of a healthy singleton.

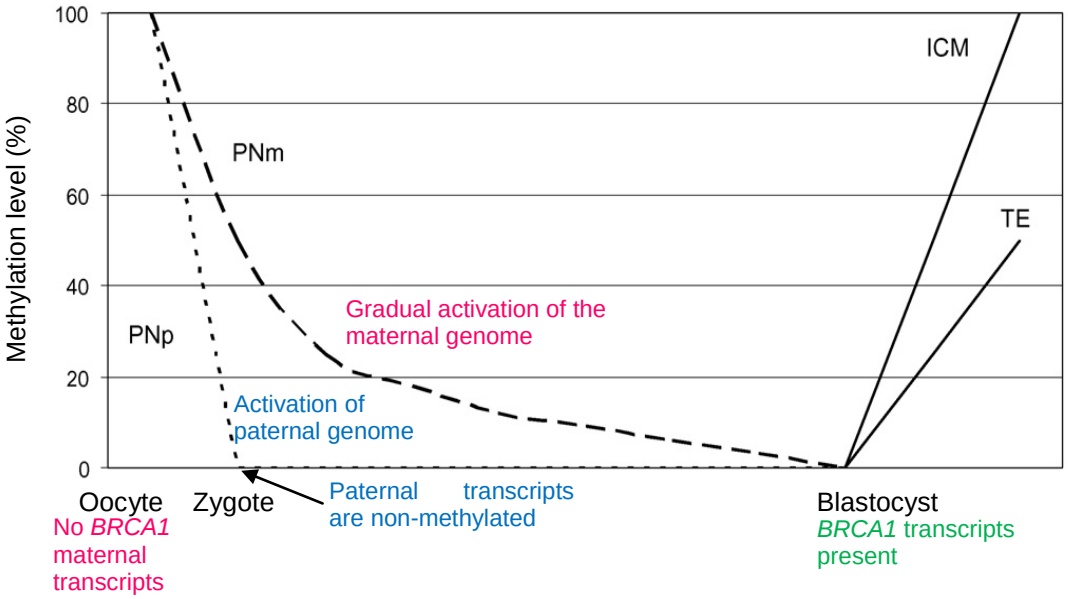
4.3.2 Outcome of a clinical PGD cycle for *BRCA1*

Five out of eleven embryos were diagnosed as transferable/unaffected and five were found to be affected with the *BRCA1* c.68_69delAG mutation (91% diagnosis rate). All embryos carrying the paternal mutation appeared to have slower or arrested growth between day 3 and day 5 (Table 3.20) but some of them had developed into blastocysts by day 6. It is possible that heterozygosity for *BRCA1* mutations may affect the embryo's development. In fact, it has been shown that the fidelity of DSBR was impaired in cell lines with heterozygous *BRCA1* missense mutations (Coupier *et al.*, 2004).

However, in two PGD cycles carried out for another couple where the female partner carried the the same mutation, no differences in embryo morphology were observed between embryos with and without the *BRCA1* mutation (personal communication with Thalia Mamas and Dr Sioban SenGupta). Therefore, it is interesting to hypothesise that the sex of the affected parent may have an impact on early embryonic development. Future cases and the investigation of *BRCA1* expression in the affected embryos collected for research will reveal if the hypothesis stays accurate.

It is possible that once the embryo has depleted the maternal mRNA transcripts and before full activation of its genome, it is dependent on the expression of the paternal genome. It has been shown that EGA occurs in two stages in mammalian embryos (Jeanblanc *et al.*, 2008). The first is a minor activation mainly targeting the paternal pronucleus; the second is a major activation step, which results in full activation of the embryonic genome (of maternal and paternal origin). Mayer *et al.* (2000) showed that the male pronucleus undergoes demethylation very soon after fertilisation. This demethylation is fast (complete eight hours post-fertilisation) and active, taking place independently of replication. However, the maternal genome undergoes passive demethylation (with levels decreasing gradually with each cell cycle) until the morula stage (Yamazaki *et al.*, 2007). Figure 4.1 illustrates the demethylation process in preimplantation embryos from the one-cell to the blastocyst stages.

Figure 4.1: DNA demethylation in preimplantation embryos



Adapted from Gynécologie Obstétrique & Fertilité, 34, Jeanblanc et al., Embryonic Genome Activation, pages 1126-1132, Copyright (2008), with permission from Elsevier

PNp: paternal pronuclei; PNm: maternal pronuclei; ICM: inner cell mass; TE: trophectoderm

Our microarray results showed that *BRCA1* mRNA was not detected in the human MII oocytes but was detected at medium levels in the blastocyst (post EGA). Even if the oocyte did express *BRCA1* but this was not picked up by our arrays, at a certain stage the maternal mRNA templates would be depleted and the embryo would have to rely on the transcription of the paternal *BRCA1* gene until full EGA occurs. Therefore, if the paternal copy of the gene carries the missense mutation, the embryo is likely to be more vulnerable to stress and/or may undergo abnormal development. Future cases and the investigation of *BRCA1* expression in the affected embryos collected for research will reveal if the paternally derived *BRCA1* transcripts are more abundant than maternally derived transcripts between day 3 and day 5 post fertilisation.

4.3.3 Conclusion of PGD workups for mutations in DNA repair gene

PGD protocols were developed and optimized for one *MSH2* mutation (*c.1277-?_1386+?del*) and two *BRCA1* mutations (*c.3339T>G* and *c.68-69delAG*). One of the *BRCA1* PGD protocols was clinically applied and resulted in the birth of a healthy singleton. The other two protocols, which were licensed for treatment by the HFEA, may be used or adjusted for other couples seeking treatment for the same condition in the future.

With the advancement of microarray platforms and whole genome amplification (WGA) methods, PGD for single gene disorders may be carried out using microarray analysis in the near future, especially for cases where the mutation is not well characterised. The application of SNP microarrays on individual blastomeres has provided DNA fingerprints for the embryos and allowed the detection of aneuploidy (Ling *et al.*, 2009; Wells *et al.*, 2008). This technique has been shown to allow the detection of polymorphisms associated with a monogenic disorder in single cells (Handyside *et al.*, 2008). As the genetics of cancer predisposition is further understood, the use of microarrays may increase the diagnostic capability of single cell analysis to include many polymorphisms that affect cancer risk.

Only a small proportion of the referred couples decide to go through with PGD treatment. Among all couples referred to the UCL Centre for PGD for genetic predispositions to cancers, the largest group was for breast and ovarian cancer (31/80 referals) due to *BRCA1* and *BRCA2* mutations. However, only 31% (11/31) of refered couples opted for PGD for *BRCA1/2* mutations compared to 87%, 71% and 65% for *NF1*, *RB1* and *APC* mutations, respectively (personal communication with Dr Sioban SenGupta). This may be due to the difference in penetrance of the disease or the lengthy process of funding, protocol development and HFEA licensing, which can interfere with their planned prophylactic treatment.

Of the three cycles that have been carried out for *BRCA1* to date the observation of morphology of embryos suggested that the parental origin of the germline mutation may affect embryonic development between days 3 and 5. The initiation of a PGD service for couples with inherited mutations for DNA repair genes will allow future investigation of DNA repair in affected embryos donated for research.

5 CONCLUDING REMARKS AND FUTURE WORK

The overall DNA repair capacity in the human oocyte or blastocyst is complex as there is a level of redundancy between different repair pathways and many DNA repair proteins are also involved in other cellular pathways. Moreover, variations in gene expression levels do not necessarily correlate with DNA repair ability as not all mRNA templates detected are fully translated to functional proteins; however, gene expression is one of the key mechanisms that can influence DNA repair activity.

The data from the microarray study showed that human MII oocytes and embryos can potentially carry out all types of repair as a large number of DNA repair genes involved in the different repair pathways were expressed at these stages. As hypothesised in the aims, the expression profiles in the blastocysts were different to the oocytes. The different expression patterns of DNA repair genes in MII oocytes and blastocysts indicate different DNA repair mechanisms pre and post embryonic genome activation (EGA). Possibly, as the maternal transcripts decline before EGA, there may be a stage during which the embryo is more vulnerable to DNA damage. Further expression analysis of embryos at different stages of development up to the blastocyst stage may identify key indicators of normal *in vitro* development and implantation potential.

This project provided preliminary evidence on the DNA repair pathways that may be active in the MII oocyte and the blastocyst. However, in order to fully understand DNA repair activity in the human preimplantation embryo, we intend to investigate each repair pathway separately and in greater detail. The highly expressed and differentially expressed DNA repair genes that have been identified in this project will serve as primary targets for further investigation by real-time PCR analysis on single preimplantation embryos at different stages of development up to the blastocyst stage.

The preimplantation embryos will be classified into three main categories: normal embryos, embryos affected with a single gene disorder (from PGD cases) and chromosomally abnormal embryos (from PGS cases). Embryo classification follows the criteria of Delhanty *et al.,* (1997) which sorts the embryos into two main groups: normal and abnormal. The chromosomally abnormal embryos can be further classified into abnormal non-mosaic, mosaic and chaotic embryos (Delhanty *et al.,* 1997).

Collected embryos will be of different morphological quality; thus they will be further categorised according to grade. The DNA repair genes that will first be investigated are listed in Table D.1 (Appendix) and were selected based on the microarray data as well as data available in the literature.

Using fluorescent immunostaining, we could demonstrate that DNA repair genes found to be expressed in the human preimplantation embryo are in fact translated into protein; however, the only way to prove DNA repair activity in the human embryo is by performing functional assays. The complexity of the DNA repair pathways justifies the need for a functional assay that measures DNA repair activity, which is the total outcome of several factors including gene expression, mRNA stability, effects of inhibitors, stimulators and environmental factors. The *in vitro* insertion/deletion loop (IDL) and mismatch repair functional assay designed and optimised in the course of this project was up to 50-fold more sensitive than available methods. However, in order to make it suitable for application on human oocytes and embryos, it requires further optimisation to achieve greater sensitivity.

Fewer molecules of heteroduplex DNA constructs need to be used so that less nuclear extract is necessary to repair the templates. Initially, this should be attempted using double labelled DNA constructs with 3, 21 and 24-nucleotide IDLs and analysis on the ABI Prism^{TM}. This would allow the detection of created homoduplex molecules even if the repair efficiency was low. The use of mini-sequencing analysis for the detection of base-base mismatch repair may also improve sensitivity. Furthermore, it may be necessary to circularise the DNA constructs or to protect the 5' ends of the DNA molecules with biotin in order to prevent exonuclease degradation. Although, this does not seem to be a major concern based on our initial results. Finally, the ultimate goal will be to assess the repair of a single molecule of DNA. This may be possible in the near future due to the development of lab on a chip technologies such as "digital microfluidics" (Advanced Liquid Logic, Inc.), which allows any reaction to take place in a liquid droplet that is electrically manipulated on a chip, minimising reaction volumes.

The initiation of a PGD service for couples with inherited mutations for DNA repair genes will allow future investigation of DNA repair in affected embryos donated for research. DNA repair capacity may have an effect on the preimplantation embryo's morphology and development between day 3 and day 5. Preimplantation embryos carrying the *BRCA1* c.68-69delAG mutation (better known as 185delAG) were collected for gene expression analysis on day 6 post fertilisation. Preliminary observations of the affected embryos suggest that development may be impaired when the mutation is paternally inherited. This may be because of the different demethylation rates of the parental genomes, which may result in a greater transcription potential for the paternal genome around day 3. Future cases and the investigation of *BRCA1* mRNA transcripts in the affected embryos using real-time PCR targeting SNPs will reveal if the hypothesis that the sex of the affected parent may have an impact on embryonic development is true. Time-lapse photography of PGD embryos from fertilisation to embryo transfer may help pinpoint the time at which the embryo may be more vulnerable to exogenous stress.

Although the *MSH2* PGD case did not reach clinical application, the premise has been set for future cases. It will be interesting to observe the effect of transitional *MSH2* deficiency on embryonic development in situations where the father carries the mutation. Additionally, it would be important to investigate microsatellite instability (DNA content) in conjunction with gene expression analysis (mRNA content) in these embryos with the possibility of assessing functional MMR (protein content) in the future.

Evaluation of the embryo's mismatch repair capacity can be achieved by investigating the extent of microsatellite instability (MSI) in embryos carrying mutations in MMR genes and normal/control embryos. The investigation of chromosome breaks (H2AX loci) together with fluorescent immunostaining of BRCA1 and other important factors involved in double strand break repair (DSBR) (like MRE11, RAD50 and NBS1) in embryos carrying *BRCA1* mutations, aneuploid and normal/control embryos may help understand the importance of DNA DSBR in the early human embryo.

The study of DNA repair gene expression in conjunction with the functional assessment of MMR and/or IDL repair efficiency will shed light on DNA repair activity in early human embryos in the context of IVF. The two parts of the project complement each other and with the initiation of PGD for germline mutations in DNA repair genes provide a model system to study the role of specific genes in early human development.

REFERENCE LIST

Adiga SK, Toyoshima M, Shimura T, Takeda J, Uematsu N and Niwa O (2007) Delayed and stage specific phosphorylation of H2AX during preimplantation development of gamma-irradiated mouse embryos. *Reproduction* 133, 415-422.

Alani E, Sokolsky T, Studamire B, Miret JJ and Lahue RS (1997) Genetic and biochemical analysis of Msh2p-Msh6p: role of ATP hydrolysis and Msh2p-Msh6p subunit interactions in mismatch base pair recognition. *Mol Cell Biol* 17, 2436-2447.

Ames BN and Shigenaga MK (1992) Oxidants are a major contributor to aging. *Ann N Y Acad Sci* 663, 85-96.

An Q, Robins P, Lindahl T and Barnes DE (2005) C --> T mutagenesis and gamma-radiation sensitivity due to deficiency in the Smug1 and Ung DNA glycosylases. *EMBO J* 24, 2205-2213.

Anderson RA and Pickering S (2008) The current status of preimplantation genetic screening: British Fertility Society Policy and Practice Guidelines. *Hum Fertil (Camb)* 11, 71-75.

Antoniou A, Pharoah PD, Narod S, Risch HA, Eyfjord JE, Hopper JL, Loman N, Olsson H, Johannsson O, Borg A *et al* (2003) Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet* 72, 1117-1130.

Ao A, Erickson RP, Winston RM and Handyside AH (1994) Transcription of paternal Y-linked genes in the human zygote as early as the pronucleate stage. *Zygote* 2, 281-287.

Aquilina G, Crescenzi M and Bignami M (1999) Mismatch repair, G(2)/M cell cycle arrest and lethality after DNA damage. *Carcinogenesis* 20, 2317-2326.

Assou S, Anahory T, Pantesco V, Le Carrour T, Pellestor F, Klein B, Reyftmann L, Dechaud H, De Vos J and Hamamah S (2006) The human cumulus--oocyte complex gene-expression profile. *Hum Reprod* 21, 1705-1719.

Aten JA, Stap J, Krawczyk PM, van Oven CH, Hoebe RA, Essers J and Kanaar R (2004) Dynamics of DNA double-strand breaks revealed by clustering of damaged chromosome domains. *Science* 303, 92-95.

Baarends WM, van der LR and Grootegoed JA (2001) DNA repair mechanisms and gametogenesis. *Reproduction* 121, 31-39.

Baker SM, Bronner CE, Zhang L, Plug AW, Robatzek M, Warren G, Elliott EA, Yu J, Ashley T, Arnheim N *et al* (1995) Male mice defective in the DNA mismatch repair gene PMS2 exhibit abnormal chromosome synapsis in meiosis. *Cell* 82, 309-319.

Baker SM, Plug AW, Prolla TA, Bronner CE, Harris AC, Yao X, Christie DM, Monell C, Arnheim N, Bradley A *et al* (1996) Involvement of mouse Mlh1 in DNA mismatch repair and meiotic crossing over. *Nat Genet* 13, 336-342.

Baldeyron C, Jacquemin E, Smith J, Jacquemont C, De O, I, Gad S, Feunteun J, Stoppa-Lyonnet D and Papadopoulo D (2002) A single mutated BRCA1 allele leads to impaired fidelity of double strand break end-joining. *Oncogene* 21, 1401-1410.

Barnes DE, Lindahl T and Sedgwick B (1993) DNA repair. *Curr Opin Cell Biol* 5, 424-433.

Barton TS, Robaire B and Hales BF (2007) DNA damage recognition in the rat zygote following chronic paternal cyclophosphamide exposure. *Toxicol Sci* 100, 495-503.

Basille C, Frydman R, Aly AE, Hesters L, Fanchin R, Tachdjian G, Steffann J, Lelorc'h M and Achour-Frydman N (2009) Preimplantation genetic diagnosis: State of the art. *Eur J Obstet Gynecol Reprod Biol*. 145, 9-13.

Bassing CH and Alt FW (2004) The cellular response to general and programmed DNA double strand breaks. *DNA Repair (Amst)* 3, 781-796.

Baumann CG, Morris DG, Sreenan JM and Leese HJ (2007) The quiet embryo hypothesis: molecular characteristics favoring viability. *Mol Reprod Dev* 74, 1345-1353.

Baumann P and West SC (1997) The human Rad51 protein: polarity of strand transfer and stimulation by hRP-A. *EMBO J* 16, 5198-5206.

Baumann P and West SC (1999) Heteroduplex formation by human Rad51 protein: effects of DNA end-structure, hRP-A and hRad52. *J Mol Biol* 291, 363-374.

Bell CE, Calder MD and Watson AJ (2008) Genomic RNA profiling and the programme controlling preimplantation mammalian development. *Mol Hum Reprod* 14, 691-701.

Bennett SE, Umar A, Oshima J, Monnat RJ, Jr. and Kunkel TA (1997) Mismatch repair in extracts of Werner syndrome cell lines. *Cancer Res* 57, 2956-2960.

Benson FE, Baumann P and West SC (1998) Synergistic actions of Rad51 and Rad52 in recombination and DNA repair. *Nature* 391, 401-404.

Blackwell LJ, Bjornson KP and Modrich P (1998) DNA-dependent activation of the hMutSalpha ATPase. *J Biol Chem* 273, 32049-32054.

Blackwell LJ, Martik D, Bjornson KP, Bjornson ES and Modrich P (1998) Nucleotide-promoted release of hMutSalpha from heteroduplex DNA is consistent with an ATP-dependent translocation mechanism. *J Biol Chem* 273, 32055-32062.

Boelens MC, te Meerman GJ, Gibcus JH, Blokzijl T, Boezen HM, Timens W, Postma DS, Groen HJ and van den BA (2007) Microarray amplification bias: loss of 30% differentially expressed genes due to long probe - poly(A)-tail distances. *BMC Genomics* 8, 277.

Bolton VN, Hawes SM, Taylor CT and Parsons JH (1989) Development of spare human preimplantation embryos in vitro: an analysis of the correlations among gross morphology, cleavage rates, and development to the blastocyst. *J In Vitro Fert Embryo Transf* 6, 30-35.

Branzei D and Foiani M (2008) Regulation of DNA repair throughout the cell cycle. *Nat Rev Mol Cell Biol* 9, 297-308.

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, 459-461.

Brendel V, Brocchieri L, Sandler SJ, Clark AJ and Karlin S (1997) Evolutionary comparisons of RecA-like proteins across all major kingdoms of living organisms. *J Mol Evol* 44, 528-541.

Brison DR (2000) Apoptosis in mammalian preimplantation embryos: regulation by survival factors. *Hum Fertil (Camb)* 3, 36-47.

Brison DR, Hollywood K, Arnesen R and Goodacre R (2007) Predicting human embryo viability: the road to non-invasive analysis of the secretome using metabolic footprinting. *Reprod Biomed Online* 15, 296-302.

Brook JD, McCurrach ME, Harley HG, Buckler AJ, Church D, Aburatani H, Hunter K, Stanton VP, Thirion JP, Hudson T *et al* (1992) Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* 69, 385.

Brose MS, Rebbeck TR, Calzone KA, Stopfer JE, Nathanson KL and Weber BL (2002) Cancer risk estimates for BRCA1 mutation carriers identified in a risk evaluation program. *J Natl Cancer Inst* 94, 1365-1372.

Buccione R, Schroeder AC and Eppig JJ (1990) Interactions between somatic cells and germ cells throughout mammalian oogenesis. *Biol Reprod* 43, 543-547.

Buchholz TA, Wu X, Hussain A, Tucker SL, Mills GB, Haffty B, Bergh S, Story M, Geara FB and Brock WA (2002) Evidence of haplotype insufficiency in human cells containing a germline mutation in BRCA1 or BRCA2. *Int J Cancer* 97, 557-561.

Carethers JM, Hawn MT, Chauhan DP, Luce MC, Marra G, Koi M and Boland CR (1996) Competency in mismatch repair prohibits clonal expansion of cancer cells treated with N-methyl-N'-nitro-N-nitrosoguanidine. *J Clin Invest* 98, 199-206.

Charbonnier F, Olschwang S, Wang Q, Boisson C, Martin C, Buisine MP, Puisieux A and Frebourg T (2002) MSH2 in contrast to MLH1 and MSH6 is frequently inactivated by exonic and promoter rearrangements in hereditary nonpolyposis colorectal cancer. *Cancer Res* 62, 848-853.

Charbonnier F, Raux G, Wang Q, Drouot N, Cordier F, Limacher JM, Saurin JC, Puisieux A, Olschwang S and Frebourg T (2000) Detection of exon deletions and duplications of the mismatch repair genes in hereditary nonpolyposis colorectal cancer families using multiplex polymerase chain reaction of short fluorescent fragments. *Cancer Res* 60, 2760-2763.

Cho EA, Prindle MJ and Dressler GR (2003) BRCT domain-containing protein PTIP is essential for progression through mitosis. *Mol Cell Biol* 23, 1666-1673.

Christmann M, Tomicic MT, Roos WP and Kaina B (2003) Mechanisms of human DNA repair: an update. *Toxicology* 193, 3-34.

Cline SD and Hanawalt PC (2003) Who's on first in the cellular response to DNA damage? *Nat Rev Mol Cell Biol* 4, 361-372.

Collins A (2009) Allelic association: linkage disequilibrium structure and gene mapping. *Mol Biotechnol* 41, 83-89.

Coupier I, Baldeyron C, Rousseau A, Mosseri V, Pages-Berhouet S, Caux-Moncoutier V, Papadopoulo D and Stoppa-Lyonnet D (2004) Fidelity of DNA double-strand break repair in heterozygous cell lines harbouring BRCA1 missense mutations. *Oncogene* 23, 914-919.

Cousineau I and Belmaaza A (2007) BRCA1 haploinsufficiency, but not heterozygosity for a BRCA1-truncating mutation, deregulates homologous recombination. *Cell Cycle* 6, 962-971.

Cromie GA, Connelly JC and Leach DR (2001) Recombination at double-strand breaks and DNA ends: conserved mechanisms from phage to humans. *Mol Cell* 8, 1163-1174.

Daphnis DD, Fragouli E, Economou K, Jerkovic S, Craft IL, Delhanty JD and Harper JC (2008) Analysis of the evolution of chromosome abnormalities in human embryos from Day 3 to 5 using CGH and FISH. *Mol Hum Reprod* 14, 117-125.

de Laat WL, Jaspers NG and Hoeijmakers JH (1999) Molecular mechanism of nucleotide excision repair. *Genes Dev* 13, 768-785.

de Wind N, Dekker M, Berns A, Radman M and te RH (1995) Inactivation of the mouse Msh2 gene results in mismatch repair deficiency, methylation tolerance, hyperrecombination, and predisposition to cancer. *Cell* 82, 321-330.

Debrock S, Melotte C, Spiessens C, Peeraer K, Vanneste E, Meeuwis L, Meuleman C, Frijns JP, Vermeesch JR and D'Hooghe TM (2009) Preimplantation genetic screening for aneuploidy of embryos after in vitro fertilization in women aged at least 35 years: a prospective randomized trial. *Fertil Steril*. In Press.

Delhanty JD and Handyside AH (1995) The origin of genetic defects in the human and their detection in the preimplantation embryo. *Hum Reprod Update* 1, 201-215.

Delhanty JD and Harper JC (2000) Pre-implantation genetic diagnosis. *Baillieres Best Pract Res Clin Obstet Gynaecol* 14, 691-708.

Derijck A, van der HG, Giele M, Philippens M and de Boer P (2008) DNA double-strand break repair in parental chromatin of mouse zygotes, the first cell cycle as an origin of de novo mutation. *Hum Mol Genet* 17, 1922-1937.

Devreker F and Englert Y (2000) In vitro development and metabolism of the human embryo up to the blastocyst stage. *Eur J Obstet Gynecol Reprod Biol* 92, 51-56.

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. *Proc Natl Acad Sci U S A* 102, 737-742.

Dikomey E, Dahm-Daphi J, Brammer I, Martensen R and Kaina B (1998) Correlation between cellular radiosensitivity and non-repaired double-strand breaks studied in nine mammalian cell lines. *Int J Radiat Biol* 73, 269-278.

Dobson AT, Raja R, Abeyta MJ, Taylor T, Shen S, Haqq C and Pera RA (2004) The unique transcriptome through day 3 of human preimplantation development. *Hum Mol Genet* 13, 1461-1470.

Dominguez F, Gadea B, Esteban FJ, Horcajadas JA, Pellicer A and Simon C (2008) Comparative protein-profile analysis of implanted versus non-implanted human blastocysts. *Hum Reprod* 23, 1993-2000.

Dronkert ML, Beverloo HB, Johnson RD, Hoeijmakers JH, Jasin M and Kanaar R (2000) Mouse RAD54 affects DNA double-strand break repair and sister chromatid exchange. *Mol Cell Biol* 20, 3147-3156.

Drost JB and Lee WR (1995) Biological basis of germline mutation: comparisons of spontaneous germline mutation rates among drosophila, mouse, and human. *Environ Mol Mutagen* 25 Suppl 26, 48-64.

Duftner N, Larkins-Ford J, Legendre M and Hofmann HA (2008) Efficacy of RNA amplification is dependent on sequence characteristics: implications for gene expression profiling using a cDNA microarray. *Genomics* 91, 108-117.

Dumitrescu RG and Cotarla I (2005) Understanding breast cancer risk -- where do we stand in 2005? *J Cell Mol Med* 9, 208-221.

Duranthon V, Watson AJ and Lonergan P (2008) Preimplantation embryo programming: transcription, epigenetics, and culture environment. *Reproduction* 135, 141-150.

Dyrkheeva NS, Khodyreva SN and Lavrik OI (2008) Interaction of APE1 and other repair proteins with DNA duplexes imitating intermediates of DNA repair and replication. *Biochemistry (Mosc)* 73, 261-272.

Eckert KA and Hile SE (2009) Every microsatellite is different: Intrinsic DNA features dictate mutagenesis of common microsatellites present in the human genome. *Mol Carcinog* 48, 379-388.

Edelmann W, Cohen PE, Kane M, Lau K, Morrow B, Bennett S, Umar A, Kunkel T, Cattoretti G, Chaganti R *et al* (1996) Meiotic pachytene arrest in MLH1-deficient mice. *Cell* 85, 1125-1134.

Edelmann W, Yang K, Umar A, Heyer J, Lau K, Fan K, Liedtke W, Cohen PE, Kane MF, Lipford JR *et al* (1997) Mutation in the mismatch repair gene Msh6 causes cancer susceptibility. *Cell* 91, 467-477.

Egozcue J, Santalo J, Gimenez C, Perez N and Vidal F (2000) Preimplantation genetic diagnosis. *Mol Cell Endocrinol* 166, 21-25.

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nat Rev Genet* 5, 435-445.

Fang WH and Modrich P (1993) Human strand-specific mismatch repair occurs by a bidirectional mechanism similar to that of the bacterial reaction. *J Biol Chem* 268, 11838-11844.

Feldmann E, Schmiemann V, Goedecke W, Reichenberger S and Pfeiffer P (2000) DNA double-strand break repair in cell-free extracts from Ku80-deficient cells: implications for Ku serving as an alignment factor in non-homologous DNA end joining. *Nucleic Acids Res* 28, 2585-2596.

Fernandez-Gonzalez R, Moreira PN, Perez-Crespo M, Sanchez-Martin M, Ramirez MA, Pericuesta E, Bilbao A, Bermejo-Alvarez P, de Dios HJ, de Fonseca FR *et al* (2008) Long-term effects of mouse intracytoplasmic sperm injection with DNA-fragmented sperm on health and behavior of adult offspring. *Biol Reprod* 78, 761-772.

Fishel R, Lescoe MK, Rao MR, Copeland NG, Jenkins NA, Garber J, Kane M and Kolodner R (1993) The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell* 75, 1027-1038.

Fleming TP, Sheth B and Fesenko I (2001) Cell adhesion in the preimplantation mammalian embryo and its role in trophectoderm differentiation and blastocyst morphogenesis. *Front Biosci* 6, D1000-D1007.

Foray N, Randrianarison V, Marot D, Perricaudet M, Lenoir G and Feunteun J (1999) Gamma-rays-induced death of human cells carrying mutations of BRCA1 or BRCA2. *Oncogene* 18, 7334-7342.

Friedberg EC, Bond JP, Burns DK, Cheo DL, Greenblatt MS, Meira LB, Nahari D and Reis AM (2000) Defective nucleotide excision repair in xpc mutant mice and its association with cancer predisposition. *Mutat Res* 459, 99-108.

Friedberg EC and Meira LB (2000) Database of mouse strains carrying targeted mutations in genes affecting cellular responses to DNA damage. Version 4. *Mutat Res* 459, 243-274.

Friedberg EC and Meira LB (2004) Database of mouse strains carrying targeted mutations in genes affecting biological responses to DNA damage (Version 6). *DNA Repair (Amst)* 3, 1617-1638.

Friedberg EC and Meira LB (2006) Database of mouse strains carrying targeted mutations in genes affecting biological responses to DNA damage Version 7. *DNA Repair (Amst)* 5, 189-209.

Genschel J, Bazemore LR and Modrich P (2002) Human exonuclease I is required for 5' and 3' mismatch repair. *J Biol Chem* 277, 13302-13311.

Gomes-Pereira M, Fortune MT, Ingram L, McAbney JP and Monckton DG (2004) Pms2 is a genetic enhancer of trinucleotide CAG.CTG repeat somatic mosaicism: implications for the mechanism of triplet repeat expansion. *Hum Mol Genet* 13, 1815-1825.

Goossens V, Harton G, Moutou C, Traeger-Synodinos J, Van Rij M and Harper JC (2009) ESHRE PGD Consortium data collection IX: cycles from January to December 2006 with pregnancy follow-up to October 2007. *Hum Reprod*. 24, 1786-810.

Gottlieb TM and Jackson SP (1993) The DNA-dependent protein kinase: requirement for DNA ends and association with Ku antigen. *Cell* 72, 131-142.

Gradia S, Acharya S and Fishel R (2000) The role of mismatched nucleotides in activating the hMSH2-hMSH6 molecular switch. *J Biol Chem* 275, 3922-3930.

Gradia S, Subramanian D, Wilson T, Acharya S, Makhov A, Griffith J and Fishel R (1999) hMSH2-hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. *Mol Cell* 3, 255-261.

Gurtu VE, Verma S, Grossmann AH, Liskay RM, Skarnes WC and Baker SM (2002) Maternal effect for DNA mismatch repair in the mouse. *Genetics* 160, 271-277.

Haber JE (2000) Partners and pathwaysrepairing a double-strand break. *Trends Genet* 16, 259-264.

Hakem R (2008) DNA-damage repair; the good, the bad, and the ugly. *EMBO J* 27, 589-605.

Hales BF (2005) DNA repair disorders causing malformations. *Curr Opin Genet Dev* 15, 234-240.

Hamatani T, Ko MS, Yamada M, Kuji N, Mizusawa Y, Shoji M, Hada T, Asada H, Maruyama T and Yoshimura Y (2006) Global gene expression profiling of preimplantation embryos. *Hum Cell* 19, 98-117.

Hammarsten O, DeFazio LG and Chu G (2000) Activation of DNA-dependent protein kinase by single-stranded DNA ends. *J Biol Chem* 275, 1541-1550.

Handyside AH, Kontogianni EH, Hardy K and Winston RM (1990) Pregnancies from biopsied human preimplantation embryos sexed by Y-specific DNA amplification. *Nature* 344, 768-770.

Handyside AH, Thornhill AR, Affara NA, Harton GL, Mariani BD and Griffin DK (2008) Recombination mapping: a universal molecular karyotyping method for preimplantation genetic diagnosis of inherited disease. *Fertil Steril* 90 Suppl 1, S24.

Hardarson T, Hanson C, Lundin K, Hillensjo T, Nilsson L, Stevic J, Reismer E, Borg K, Wikland M and Bergh C (2008) Preimplantation genetic screening in women of advanced maternal age caused a decrease in clinical pregnancy rate: a randomized controlled trial. *Hum Reprod* 23, 2806-2812.

Hardy K (1999) Apoptosis in the human embryo. *Rev Reprod* 4, 125-134.

Hardy K, Stark J and Winston RM (2003) Maintenance of the inner cell mass in human blastocysts from fragmented embryos. *Biol Reprod* 68, 1165-1169.

Harper J, Sermon K, Geraedts J, Vesela K, Harton G, Thornhill A, Pehlivan T, Fiorentino F, SenGupta S, Die-Smulders C *et al* (2008) What next for preimplantation genetic screening? *Hum Reprod* 23, 478-480.

Harrison RH, Kuo HC, Scriven PN, Handyside AH and Ogilvie CM (2000) Lack of cell cycle checkpoints in human cleavage stage embryos revealed by a clonal pattern of chromosomal mosaicism analysed by sequential multicolour FISH. *Zygote* 8, 217-224.

Hartley KO, Gell D, Smith GC, Zhang H, Divecha N, Connelly MA, Admon A, Lees-Miller SP, Anderson CW and Jackson SP (1995) DNA-dependent protein kinase catalytic subunit: a relative of phosphatidylinositol 3-kinase and the ataxia telangiectasia gene product. *Cell* 82, 849-856.

Hawn MT, Umar A, Carethers JM, Marra G, Kunkel TA, Boland CR and Koi M (1995) Evidence for a connection between the mismatch repair system and the G2 cell cycle checkpoint. *Cancer Res* 55, 3721-3725.

Hernandez-Gonzalez I, Gonzalez-Robayna I, Shimada M, Wayne CM, Ochsner SA, White L and Richards JS (2006) Gene expression profiles of cumulus cell oocyte complexes during ovulation reveal cumulus cells express neuronal and immune-related genes: does this expand their role in the ovulation process? *Mol Endocrinol* 20, 1300-1321.

Hogervorst FB, Cornelis RS, Bout M, van Vliet M, Oosterwijk JC, Olmer R, Bakker B, Klijn JG, Vasen HF, Meijers-Heijboer H *et al* (1995) Rapid detection of BRCA1 mutations by the protein truncation test. *Nat Genet* 10, 208-212.

Hsieh P and Yamane K (2008) DNA mismatch repair: molecular mechanism, cancer, and ageing. *Mech Ageing Dev* 129, 391-407.

Iaccarino I, Marra G, Palombo F and Jiricny J (1998) hMSH2 and hMSH6 play distinct roles in mismatch binding and contribute differently to the ATPase activity of hMutSalpha. *EMBO J* 17, 2677-2686.

Jackson SP (2002) Sensing and repairing DNA double-strand breaks. *Carcinogenesis* 23, 687-696.

Jansen RP, Bowman MC, de Boer KA, Leigh DA, Lieberman DB and McArthur SJ (2008) What next for preimplantation genetic screening (PGS)? Experience with blastocyst biopsy and testing for aneuploidy. *Hum Reprod* 23, 1476-1478.

Jaroudi S and SenGupta S (2007) DNA repair in mammalian embryos. *Mutat Res* 635, 53-77.

Jeanblanc M, Salvaing J, Mason K, Debey P and Beaujean N (2008) [Embryonic genome activation]. *Gynecol Obstet Fertil* 36, 1126-1132.

Jeggo PA, Hafezparast M, Thompson AF, Broughton BC, Kaur GP, Zdzienicka MZ and Athwal RS (1992) Localization of a DNA repair gene (XRCC5) involved in double-strand-break rejoining to human chromosome 2. *Proc Natl Acad Sci USA* 89, 6423-6427.

Jiricny J (1998) Replication errors: cha(lle)nging the genome. *EMBO J* 17, 6427-6436.

Johnson RD and Jasin M (2000) Sister chromatid gene conversion is a prominent double-strand break repair pathway in mammalian cells. *EMBO J* 19, 3398-3407.

Jones S, Emmerson P, Maynard J, Best JM, Jordan S, Williams GT, Sampson JR and Cheadle JP (2002) Biallelic germline mutations in MYH predispose to multiple colorectal adenoma and somatic G:C-->T:A mutations. *Hum Mol Genet* 11, 2961-2967.

Jurisicova A, Latham KE, Casper RF, Casper RF and Varmuza SL (1998) Expression and regulation of genes associated with cell death during murine preimplantation embryo development. *Mol Reprod Dev* 51, 243-253.

Kadyrov FA, Genschel J, Fang Y, Penland E, Edelmann W and Modrich P (2009) A possible mechanism for exonuclease 1-independent eukaryotic mismatch repair. *Proc Natl Acad Sci U S A* 106, 8495-8500.

Kagawa W, Kurumizaka H, Ikawa S, Yokoyama S and Shibata T (2001) Homologous pairing promoted by the human Rad52 protein. *J Biol Chem* 276, 35201-35208.

Kan R, Sun X, Kolas NK, Avdievich E, Kneitz B, Edelmann W and Cohen PE (2008) Comparative analysis of meiotic progression in female mice bearing mutations in genes of the DNA mismatch repair pathway. *Biol Reprod* 78, 462-471.

Karran,P. and Bignami,M. (1999) Mismatch repair and cancer. In Smith,P.J. and Jones C.J. (eds), *DNA Recombination and Repair*. Oxford University Press, New York, pp. 66-98.

Katz-Jaffe MG, Linck DW, Schoolcraft WB and Gardner DK (2005) A proteomic analysis of mammalian preimplantation embryonic development. *Reproduction* 130, 899-905.

Katz-Jaffe MG, McReynolds S, Gardner DK and Schoolcraft WB (2009) The role of proteomics in defining the human embryonic secretome. *Mol Hum Reprod* 15, 271-277.

Katz-Jaffe MG, Schoolcraft WB and Gardner DK (2006) Analysis of protein expression (secretome) by human and mouse preimplantation embryos. *Fertil Steril* 86, 678-685.

Kearns WG, Pen R, Benner A, Kittai A, Widra E and Leach R (2008) SNP microarray genetic analyses to determine 23-chromosome ploidy, structural chromosome abberations and genome-wide scans to identify disease risks from a single embryonic cell. *Fertil Steril* 90 Suppl 1, S23.

Kemp Z, Thirlwell C, Sieber O, Silver A and Tomlinson I (2004) An update on the genetics of colorectal cancer. *Hum Mol Genet* 13 Spec No 2, R177-R185.

King MC, Marks JH and Mandell JB (2003) Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2. *Science* 302, 643-646.

Koonin EV, Altschul SF and Bork P (1996) BRCA1 protein products ... Functional motifs.. *Nat Genet* 13, 266-268.

Labhart P (1999) Ku-dependent nonhomologous DNA end joining in Xenopus egg extracts. *Mol Cell Biol* 19, 2585-2593.

Lahiri DK and Nurnberger JI, Jr. (1991) A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. *Nucleic Acids Res* 19, 5444.

Larson JS, Stringer SL and Stringer JR (2004) Impact of mismatch repair deficiency on genomic stability in the maternal germline and during early embryonic development. *Mutat Res* 556, 45-53.

Latham KE, Garrels JI, Chang C and Solter D (1992) Analysis of embryonic mouse development: construction of a high-resolution, two-dimensional gel protein database. *Appl Theor Electrophor* 2, 163-170.

Leach FS, Nicolaides NC, Papadopoulos N, Liu B, Jen J, Parsons R, Peltomaki P, Sistonen P, Aaltonen LA, Nystrom-Lahti M *et al* (1993) Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell* 75, 1215-1225.

Leadon SA (1999) Transcription-coupled repair of DNA damage: unanticipated players, unexpected complexities. *Am J Hum Genet* 64, 1259-1263.

Leber R, Wise TW, Mizuta R and Meek K (1998) The XRCC4 gene product is a target for and interacts with the DNA-dependent protein kinase. *J Biol Chem* 273, 1794-1801.

Lee JW, Inamdar KV, Hannah MF, Lees-Miller SP and Povirk LF (2003) DNA end sequestration by DNA-dependent protein kinase and end joining of sterically constrained substrates in whole-cell extracts. *Environ Mol Mutagen* 42, 279-287.

Lee ML, Kuo FC, Whitmore GA and Sklar J (2000) Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations. *Proc Natl Acad Sci U S A* 97, 9834-9839.

Leese HJ (2002) Quiet please, do not disturb: a hypothesis of embryo metabolism and viability. *Bioessays* 24, 845-849.

Lehmann AR (2003) DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. *Biochimie* 85, 1101-1111.

Li GM (2008) Mechanisms and functions of DNA mismatch repair. *Cell Res* 18, 85-98.

Li M, DeUgarte CM, Surrey M, Danzer H, DeCherney A and Hill DL (2005) Fluorescence in situ hybridization reanalysis of day-6 human blastocysts diagnosed with aneuploidy on day 3. *Fertil Steril* 84, 1395-1400.

Li YC, Korol AB, Fahima T, Beiles A and Nevo E (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Mol Ecol* 11, 2453-2465.

Libby RT, Monckton DG, Fu YH, Martinez RA, McAbney JP, Lau R, Einum DD, Nichol K, Ware CB, Ptacek LJ *et al* (2003) Genomic context drives SCA7 CAG repeat instability, while expressed SCA7 cDNAs are intergenerationally and somatically stable in transgenic mice. *Hum Mol Genet* 12, 41-50.

Liebens FP, Carly B, Pastijn A and Rozenberg S (2007) Management of BRCA1/2 associated breast cancer: a systematic qualitative review of the state of knowledge in 2006. *Eur J Cancer* 43, 238-257.

Lipkin SM, Moens PB, Wang V, Lenzi M, Shanmugarajah D, Gilgeous A, Thomas J, Cheng J, Touchman JW, Green ED *et al* (2002) Meiotic arrest and aneuploidy in MLH3-deficient mice. *Nat Genet* 31, 385-390.

Lipkin SM, Wang V, Jacoby R, Banerjee-Basu S, Baxevanis AD, Lynch HT, Elliott RM and Collins FS (2000) MLH3: a DNA mismatch repair gene associated with mammalian microsatellite instability. *Nat Genet* 24, 27-35.

Lipton L, Halford SE, Johnson V, Novelli MR, Jones A, Cummings C, Barclay E, Sieber O, Sadat A, Bisgaard ML *et al* (2003) Carcinogenesis in MYH-associated polyposis follows a distinct genetic pathway. *Cancer Res* 63, 7595-7599.

Littman SJ, Fang WH and Modrich P (1999) Repair of large insertion/deletion heterologies in human nuclear extracts is directed by a 5' single-strand break and is independent of the mismatch repair system. *J Biol Chem* 274, 7474-7481.

Liu B, Parsons R, Papadopoulos N, Nicolaides NC, Lynch HT, Watson P, Jass JR, Dunlop M, Wyllie A, Peltomaki P *et al* (1996) Analysis of mismatch repair genes in hereditary non-polyposis colorectal cancer patients. *Nat Med* 2, 169-174.

Lonergan P, Rizos D, Gutierrez-Adan A, Fair T and Boland MP (2003a) Oocyte and embryo quality: effect of origin, culture conditions and gene expression patterns. *Reprod Domest Anim* 38, 259-67.

Lonergan P, Rizos D, Gutiérrez-Adán A, Fair T and Boland MP (2003b) Effect of culture environment on embryo quality and gene expression - experience from animal studies. *Reprod Biomed Online* 7, 657-663.

Longley MJ, Pierce AJ and Modrich P (1997) DNA polymerase delta is required for human mismatch repair in vitro. *J Biol Chem* 272, 10917-10921.

Lukas J, Lukas C and Bartek J (2004) Mammalian cell cycle checkpoints: signalling pathways and their organization in space and time. *DNA Repair (Amst)* 3, 997-1007.

Lynch HT, Shaw TG and Lynch JF (2004) Inherited predisposition to cancer: a historical overview. *Am J Med Genet C Semin Med Genet* 129C, 5-22.

Mamo S, Bodo S, Kobolak J, Polgar Z, Tolgyesi G and Dinnyes A (2006a) Gene expression profiles of vitrified in vivo derived 8-cell stage mouse embryos detected by high density oligonucleotide microarrays. *Mol Reprod Dev* 73, 1380-1392.

Mamo S, Gal AB, Bodo S and Dinnyes A (2007) Quantitative evaluation and selection of reference genes in mouse oocytes and embryos cultured in vivo and in vitro. *BMC Dev Biol* 7, 14.

Mamo S, Sargent CA, Affara NA, Tesfaye D, El Halawany N, Wimmers K, Gilles M, Schellander K and Ponsuksili S (2006b) Transcript profiles of some developmentally important genes detected in bovine oocytes and in vitro-produced blastocysts using RNA amplification and cDNA microarrays. *Reprod Domest Anim* 41, 527-534.

Manley K, Shirley TL, Flaherty L and Messer A (1999) Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice. *Nat Genet* 23, 471-473.

Marquez N, Chappell SC, Sansom OJ, Clarke AR, Court J, Errington RJ and Smith PJ (2003) Single cell tracking reveals that Msh2 is a key component of an early-acting DNA damage-activated G2 checkpoint. *Oncogene* 22, 7642-7648.

Martensson S and Hammarsten O (2002) DNA-dependent protein kinase catalytic subunit. Structural requirements for kinase activation by DNA ends. *J Biol Chem* 277, 3020-3029.

Marti TM, Kunz C and Fleck O (2002) DNA mismatch repair and mutation avoidance pathways. *J Cell Physiol* 191, 28-41.

Martorell L, Johnson K, Boucher CA and Baiget M (1997) Somatic instability of the myotonic dystrophy (CTG)n repeat during human fetal development. *Hum Mol Genet* 6, 877-880.

Maser RS, Monsen KJ, Nelms BE and Petrini JH (1997) hMre11 and hRad50 nuclear foci are induced during the normal cellular response to DNA double-strand breaks. *Mol Cell Biol* 17, 6087-6096.

Mastenbroek S, Twisk M, Echten-Arends J, Sikkema-Raddatz B, Korevaar JC, Verhoeve HR, Vogel NE, Arts EG, de Vries JW, Bossuyt PM *et al* (2007) In vitro fertilization with preimplantation genetic screening. *N Engl J Med* 357, 9-17.

Matheson EC and Hall AG (2003) Assessment of mismatch repair function in leukaemic cell lines and blasts from children with acute lymphoblastic leukaemia. *Carcinogenesis* 24, 31-38.

May A, Kirchner R, Muller H, Hartmann P, El Hajj N, Tresch A, Zechner U, Mann W and Haaf T (2009) Multiplex rt-PCR expression analysis of developmentally important genes in individual mouse preimplantation embryos and blastomeres. *Biol Reprod* 80, 194-202.

Maynard S, Swistowska AM, Lee JW, Liu Y, Liu ST, Da Cruz AB, Rao M, Souza-Pinto NC, Zeng X and Bohr VA (2008) Human embryonic stem cells have enhanced repair of multiple forms of DNA damage. *Stem Cells* 26, 2266-2274.

McCulloch SD, Gu L and Li GM (2003a) Bi-directional processing of DNA loops by mismatch repair-dependent and -independent pathways in human cells. *J Biol Chem* 278, 3891-3896.

McCulloch SD, Gu L and Li GM (2003b) Nick-dependent and -independent processing of large DNA loops in human cells. *J Biol Chem* 278, 50803-50809.

Menezo Y, Jr., Russo G, Tosti E, El Mouatassim S 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. *J Assist Reprod Genet* 24, 513-520.

Menezo YJ (2004) Blastocyst freezing. *Eur J Obstet Gynecol Reprod Biol* 115 Suppl 1, S12-S15.

Menezo YJ (2006) Paternal and maternal factors in preimplantation embryogenesis: interaction with the biochemical environment. *Reprod Biomed Online* 12, 616-621.

Menon U, Harper J, Sharma A, Fraser L, Burnell M, ElMasry K, Rodeck C and Jacobs I (2007) Views of BRCA gene mutation carriers on preimplantation genetic diagnosis as a reproductive option for hereditary breast and ovarian cancer. *Hum Reprod* 22, 1573-1577.

Mersereau JE, Pergament E, Zhang X and Milad MP (2008) Preimplantation genetic screening to improve in vitro fertilization pregnancy rates: a prospective randomized controlled trial. *Fertil Steril* 90, 1287-1289.

Meyer LR, Klipstein S, Hazlett WD, Nasta T, Mangan P and Karande VC (2009) A prospective randomized controlled trial of preimplantation genetic screening in the "good prognosis" patient. *Fertil Steril* 91, 1731-1738.

Meyer-Ficca ML, Lonchar J, Credidio C, Ihara M, Li Y, Wang ZQ and Meyer RG (2009) Disruption of Poly(ADP-Ribose) Homeostasis Affects Spermiogenesis and Sperm Chromatin Integrity in Mice. *Biol Reprod*. 81, 46-55.

Mi H, Lazareva-Ulitsky B, Loo R, Kejariwal A, Vandergriff J, Rabkin S, Guo N, Muruganujan A, Doremieux O, Campbell MJ *et al* (2005) The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic Acids Res* 33, D284-D288.

Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W *et al* (1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* 266, 66-71.

Mirkin SM (2006) DNA structures, repeat expansions and human hereditary disorders. *Curr Opin Struct Biol* 16, 351-358.

Mirkin SM (2007) Expandable DNA repeats and human disease. *Nature* 447, 932-940.

Moayeri SE, Allen RB, Brewster WR, Kim MH, Porto M and Werlin LB (2008) Day-3 embryo morphology predicts euploidy among older subjects. *Fertil Steril* 89, 118-123.

Modrich P and Lahue R (1996) Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu Rev Biochem* 65, 101-133.

Morgan HD, Santos F, Green K, Dean W and Reik W (2005) Epigenetic reprogramming in mammals. *Hum Mol Genet* 14 Spec No 1, R47-R58.

Moshous D, Callebaut I, de Chasseval R, Corneo B, Cavazzana-Calvo M, Le Deist F, Tezcan I, Sanal O, Bertrand Y, Philippe N *et al* (2001) Artemis, a novel DNA double-strand break repair/V(D)J recombination protein, is mutated in human severe combined immune deficiency. *Cell* 105, 177-186.

Mtango NR, Potireddy S and Latham KE (2008) Oocyte quality and maternal control of development. *Int Rev Cell Mol Biol* 268, 223-290.

Munne S, Alikani M, Tomkin G, Grifo J and Cohen J (1995) Embryo morphology, developmental rates, and maternal age are correlated with chromosome abnormalities. *Fertil Steril* 64, 382-391.

Munne S, Marquez C, Reing A, Garrisi J and Alikani M (1998) Chromosome abnormalities in embryos obtained after conventional in vitro fertilization and intracytoplasmic sperm injection. *Fertil Steril* 69, 904-908.

Munne S, Sandalinas M, Escudero T, Fung J, Gianaroli L and Cohen J (2000) Outcome of preimplantation genetic diagnosis of translocations. *Fertil Steril* 73, 1209-1218.

Munne S, Velilla E, Colls P, Garcia BM, Vemuri MC, Steuerwald N, Garrisi J and Cohen J (2005) Self-correction of chromosomally abnormal embryos in culture and implications for stem cell production. *Fertil Steril* 84, 1328-1334.

Muttukrishna S, McGarrigle H, Wakim R, Khadum I, Ranieri DM and Serhal P (2005) Antral follicle count, anti-mullerian hormone and inhibin B: predictors of ovarian response in assisted reproductive technology? *BJOG* 112, 1384-1390.

Narayanan L, Fritzell JA, Baker SM, Liskay RM and Glazer PM (1997) Elevated levels of mutation in multiple tissues of mice deficient in the DNA mismatch repair gene Pms2. *Proc Natl Acad Sci U S A* 94, 3122-3127.

Nelms BE, Maser RS, MacKay JF, Lagally MG and Petrini JH (1998) In situ visualization of DNA double-strand break repair in human fibroblasts. *Science* 280, 590-592.

New JH, Sugiyama T, Zaitseva E and Kowalczykowski SC (1998) Rad52 protein stimulates DNA strand exchange by Rad51 and replication protein A. *Nature* 391, 407-410.

Oda N, Saxena JK, Jenkins TM, Prasad R, Wilson SH and Ackerman EJ (1996) DNA polymerases alpha and beta are required for DNA repair in an efficient nuclear extract from Xenopus oocytes. *J Biol Chem* 271, 13816-13820.

Orita M, Iwahana H, Kanazawa H, Hayashi K and Sekiya T (1989a) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc Natl Acad Sci U S A* 86, 2766-2770.

Palombo F, Gallinari P, Iaccarino I, Lettieri T, Hughes M, D'Arrigo A, Truong O, Hsuan JJ and Jiricny J (1995) GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells. *Science* 268, 1912-1914.

Palombo F, Iaccarino I, Nakajima E, Ikejima M, Shimada T and Jiricny J (1996) hMutSbeta, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. *Curr Biol* 6, 1181-1184.

Pan H, Ma P, Zhu W and Schultz RM (2008) Age-associated increase in aneuploidy and changes in gene expression in mouse eggs. *Dev Biol* 316, 397-407.

Papp J, Kovacs ME and Olah E (2007) Germline MLH1 and MSH2 mutational spectrum including frequent large genomic aberrations in Hungarian hereditary non-polyposis colorectal cancer families: implications for genetic testing. *World J Gastroenterol* 13, 2727-2732.

Park MS, Ludwig DL, Stigger E and Lee SH (1996) Physical interaction between human RAD52 and RPA is required for homologous recombination in mammalian cells. *J Biol Chem* 271, 18996-19000.

Park Y and Gerson SL (2005) DNA REPAIR DEFECTS IN STEM CELL FUNCTION AND AGING *. *Annu Rev Med* 56, 495-508.

Parsons R, Li GM, Longley M, Modrich P, Liu B, Berk T, Hamilton SR, Kinzler KW and Vogelstein B (1995) Mismatch repair deficiency in phenotypically normal human cells. *Science* 268, 738-740.

Patel OV, Suchyta SP, Sipkovsky SS, Yao J, Ireland JJ, Coussens PM and Smith GW (2005) Validation and application of a high fidelity mRNA linear amplification procedure for profiling gene expression. *Vet Immunol Immunopathol* 105, 331-342.

Pearson CE, Nichol EK and Cleary JD (2005) Repeat instability: mechanisms of dynamic mutations. *Nat Rev Genet* 6, 729-742.

Peltomaki P, Gao X and Mecklin JP (2001) Genotype and phenotype in hereditary nonpolyposis colon cancer: a study of families with different vs. shared predisposing mutations. *Fam Cancer* 1, 9-15.

Peltomaki P and Vasen HF (1997) Mutations predisposing to hereditary nonpolyposis colorectal cancer: database and results of a collaborative study. The International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer. *Gastroenterology* 113, 1146-1158.

Peng X, Wood CL, Blalock EM, Chen KC, Landfield PW and Stromberg AJ (2003) Statistical implications of pooling RNA samples for microarray experiments. *BMC Bioinformatics* 4, 26.

Petranovic M, Vlahovic K, Zahradka D, Dzidic S and Radman M (2000) Mismatch repair in xenopus egg extracts is not strand-directed by DNA methylation. *Neoplasma* 47, 375-381.

Pfeiffer P, Goedecke W and Obe G (2000) Mechanisms of DNA double-strand break repair and their potential to induce chromosomal aberrations. *Mutagenesis* 15, 289-302.

Piyamongkol W, Harper JC, Sherlock JK, Doshi A, Serhal PF, Delhanty JD and Wells D (2001) A successful strategy for preimplantation genetic diagnosis of myotonic dystrophy using multiplex fluorescent PCR. *Prenat Diagn* 21, 223-232.

Prolla TA, Baker SM, Harris AC, Tsao JL, Yao X, Bronner CE, Zheng B, Gordon M, Reneker J, Arnheim N *et al* (1998) Tumour susceptibility and spontaneous mutation in mice deficient in Mlh1, Pms1 and Pms2 DNA mismatch repair. *Nat Genet* 18, 276-279.

Reeves WH, Sthoeger ZM and Lahita RG (1989) Role of antigen selectivity in autoimmune responses to the Ku (p70/p80) antigen. *J Clin Invest* 84, 562-567.

Reik W, Dean W and Walter J (2001) Epigenetic reprogramming in mammalian development. *Science* 293, 1089-1093.

Reitmair AH, Schmits R, Ewel A, Bapat B, Redston M, Mitri A, Waterhouse P, Mittrucker HW, Wakeham A, Liu B *et al* (1995) MSH2 deficient mice are viable and susceptible to lymphoid tumours. *Nat Genet* 11, 64-70.

Rich T, Allen RL and Wyllie AH (2000) Defying death after DNA damage. *Nature* 407, 777-783.

Riis B, Risom L, Loft S and Poulsen HE (2002) OGG1 mRNA expression and incision activity in rats are higher in foetal tissue than in adult liver tissue while 8-oxo-2'-deoxyguanosine levels are unchanged. *DNA Repair (Amst)* 1, 709-717.

Robertson AB, Klungland A, Rognes T and Leiros I (2009) DNA repair in mammalian cells: Base excision repair: the long and short of it. *Cell Mol Life Sci* 66, 981-993.

Roig I, Liebe B, Egozcue J, Cabero L, Garcia M and Scherthan H (2004) Female-specific features of recombinational double-stranded DNA repair in relation to synapsis and telomere dynamics in human oocytes. *Chromosoma* 113, 22-33.

Ronen A and Glickman BW (2001) Human DNA repair genes. *Environ Mol Mutagen* 37, 241-283.

Rothfuss A, Schutz P, Bochum S, Volm T, Eberhardt E, Kreienberg R, Vogel W and Speit G (2000) Induced micronucleus frequencies in peripheral lymphocytes as a screening test for carriers of a BRCA1 mutation in breast cancer families. *Cancer Res* 60, 390-394.

Rozen S and Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 132, 365-386.

Rudnicki M, Eder S, Schratzberger G, Mayer B, Meyer TW, Tonko M and Mayer G (2004) Reliability of t7-based mRNA linear amplification validated by gene expression analysis of human kidney cells using cDNA microarrays. *Nephron Exp Nephrol* 97, e86-e95.

Sahu B, Ozturk O, Deo N, Fordham K, Ranierri M and Serhal P (2008) Response to controlled ovarian stimulation and oocyte quality in women with myotonic dystrophy type I. *J Assist Reprod Genet* 25, 1-5.

Sampson JR, Jones S, Dolwani S and Cheadle JP (2005) MutYH (MYH) and colorectal cancer. *Biochem Soc Trans* 33, 679-683.

Sancar A, Lindsey-Boltz LA, Unsal-Kacmaz K and Linn S (2004) Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu Rev Biochem* 73, 39-85.

Savouret C, Brisson E, Essers J, Kanaar R, Pastink A, te RH, Junien C and Gourdon G (2003) CTG repeat instability and size variation timing in DNA repair-deficient mice. *EMBO J* 22, 2264-2273.

Savouret C, Garcia-Cordier C, Megret J, te RH, Junien C and Gourdon G (2004) MSH2-dependent germinal CTG repeat expansions are produced continuously in spermatogonia from DM1 transgenic mice. *Mol Cell Biol* 24, 629-637.

Schultz RM (2002) The molecular foundations of the maternal to zygotic transition in the preimplantation embryo. *Hum Reprod Update* 8, 323-331.

Shi CZ, Collins HW, Garside WT, Buettger CW, Matschinsky FM and Heyner S (1994) Protein databases for compacted eight-cell and blastocyst-stage mouse embryos. *Mol Reprod Dev* 37, 34-47.

Shuck SC, Short EA and Turchi JJ (2008) Eukaryotic nucleotide excision repair: from understanding mechanisms to influencing biology. *Cell Res* 18, 64-72.

Singleton,B.K. and Jeggo,P.A. (1999) Double-strand break repair and V(D)J recombination. In Smith,P.J. and Jones C.J. (eds), *DNA Recombination and Repair*. Oxford University Press Inc., New York, pp. 16-37.

Singleton MR, Wentzell LM, Liu Y, West SC and Wigley DB (2002) Structure of the single-strand annealing domain of human RAD52 protein. *Proc Natl Acad Sci USA* 99, 13492-13497.

Sipley JD, Menninger JC, Hartley KO, Ward DC, Jackson SP and Anderson CW (1995) Gene for the catalytic subunit of the human DNA-activated protein kinase maps to the site of the XRCC7 gene on chromosome 8. *Proc Natl Acad Sci U S A* 92, 7515-7519.

Smith GC and Jackson SP (1999) The DNA-dependent protein kinase. *Genes Dev* 13, 916-934.

Sonoda E, Takata M, Yamashita YM, Morrison C and Takeda S (2001) Homologous DNA recombination in vertebrate cells. *Proc Natl Acad Sci U S A* 98, 8388-8394.

Spandidos DA, Koumantakis E, Sifakis S and Sourvinos G (1998) Microsatellite mutations in spontaneously aborted embryos. *Fertil Steril* 70, 892-895.

Spits C, De Rycke M, Van Ranst N, Verpoest W, Lissens W, Van Steirteghem A, Liebaers I and Sermon K (2007) Preimplantation genetic diagnosis for cancer predisposition syndromes. *Prenat Diagn* 27, 447-456.

Spits C and Sermon K (2009) PGD for monogenic disorders: aspects of molecular biology. *Prenat Diagn* 29, 50-56.

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. *Hum Reprod* 19, 2849-2858.

Staessen C, Verpoest W, Donoso P, Haentjens P, Van der EJ, Liebaers I and Devroey P (2008) Preimplantation genetic screening does not improve delivery rate in women under the age of 36 following single-embryo transfer. *Hum Reprod* 23, 2818-2825.

Stasiak AZ, Larquet E, Stasiak A, Muller S, Engel A, Van Dyck E, West SC and Egelman EH (2000) The human Rad52 protein exists as a heptameric ring. *Curr Biol* 10, 337-340.

Stevens J, Wale P, Surrey ES and Schoolcraft WB (2004) Is aneuploidy screening for patients aged 35 or over beneficial? A prospective randomized trial. *Fertil Steril* 82 Suppl 2, 249.

Stojic L, Brun R and Jiricny J (2004) Mismatch repair and DNA damage signalling. *DNA Repair (Amst)* 3, 1091-1101.

Strom CM, Levin R, Strom S, Masciangelo C, Kuliev A and Verlinsky Y (2000) Neonatal outcome of preimplantation genetic diagnosis by polar body removal: the first 109 infants. *Pediatrics* 106, 650-653.

Stuckless S, Parfrey PS, Woods MO, Cox J, Fitzgerald GW, Green JS and Green RC (2007) The phenotypic expression of three MSH2 mutations in large Newfoundland families with Lynch syndrome. *Fam Cancer* 6, 1-12.

Sturmey RG, Brison DR and Leese HJ (2008) Symposium: innovative techniques in human embryo viability assessment. Assessing embryo viability by measurement of amino acid turnover. *Reprod Biomed Online* 17, 486-496.

Sturmey RG, Hawkhead JA, Barker EA and Leese HJ (2009) DNA damage and metabolic activity in the preimplantation embryo. *Hum Reprod* 24, 81-91.

Subramanian S, Mishra RK and Singh L (2003) Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. *Genome Biol* 4, R13.

Susse S, Scholz CJ, Burkle A and Wiesmuller L (2004) Poly(ADP-ribose) polymerase (PARP-1) and p53 independently function in regulating double-strand break repair in primate cells. *Nucleic Acids Res* 32, 669-680.

Takata M, Sasaki MS, Sonoda E, Morrison C, Hashimoto M, Utsumi H, Yamaguchi-Iwai Y, Shinohara A and Takeda S (1998) Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO J* 17, 5497-5508.

Taylor EM and Lehmann AR (1998) Conservation of eukaryotic DNA repair mechanisms. *Int J Radiat Biol* 74, 277-286.

Telford NA, Watson AJ and Schultz GA (1990) Transition from maternal to embryonic control in early mammalian development: a comparison of several species. *Mol Reprod Dev* 26, 90-100.

Tesarik J, Kopecny V, Plachot M and Mandelbaum J (1986) Activation of nucleolar and extranucleolar RNA synthesis and changes in the ribosomal content of human embryos developing in vitro. *J Reprod Fertil* 78, 463-470.

Tesarik J, Kopecny V, Plachot M and Mandelbaum J (1987) High-resolution autoradiographic localization of DNA-containing sites and RNA synthesis in developing nucleoli of human preimplantation embryos: a new concept of embryonic nucleologenesis. *Development* 101, 777-791.

Tesarik J, Kopecny V, Plachot M and Mandelbaum J (1988) Early morphological signs of embryonic genome expression in human preimplantation development as revealed by quantitative electron microscopy. *Dev Biol* 128, 15-20.

Thomas DC, Umar A and Kunkel TA (1996) Microsatellite instability and mismatch repair defects in cancer. *Mutat Res* 350, 201-205.

Thomas MR, Sparks AE, Ryan GL and Van Voorhis BJ (2009) Clinical predictors of human blastocyst formation and pregnancy after extended embryo culture and transfer. *Fertil Steril*. In Press.

Thornhill AR, deDie-Smulders CE, Geraedts JP, Harper JC, Harton GL, Lavery SA, Moutou C, Robinson MD, Schmutzler AG, Scriven PN *et al* (2005) ESHRE PGD Consortium 'Best practice guidelines for clinical preimplantation genetic diagnosis (PGD) and preimplantation genetic screening (PGS)'. *Hum Reprod* 20, 35-48.

Tichy ED and Stambrook PJ (2008) DNA repair in murine embryonic stem cells and differentiated cells. *Exp Cell Res* 314, 1929-1936.

Ulrich HD (2005) The RAD6 pathway: control of DNA damage bypass and mutagenesis by ubiquitin and SUMO. *Chembiochem* 6, 1735-1743.

Valerie K and Povirk LF (2003) Regulation and mechanisms of mammalian double-strand break repair. *Oncogene* 22, 5792-5812.

van Brabant AJ, Stan R and Ellis NA (2000) DNA helicases, genomic instability, and human genetic disease. *Annu Rev Genomics Hum Genet* 1, 409-459.

van den Broek WJ, Nelen MR, Wansink DG, Coerwinkel MM, te RH, Groenen PJ and Wieringa B (2002) Somatic expansion behaviour of the (CTG)n repeat in myotonic dystrophy knock-in mice is differentially affected by Msh3 and Msh6 mismatch-repair proteins. *Hum Mol Genet* 11, 191-198.

Van Dyck E, Stasiak AZ, Stasiak A and West SC (2001) Visualization of recombination intermediates produced by RAD52-mediated single-strand annealing. *EMBO Rep* 2, 905-909.

Vanneste E, Voet T, Le Caignec C, Ampe M, Konings P, Melotte C, Debrock S, Amyere M, Vikkula M, Schuit F *et al* (2009) Chromosome instability is common in human cleavage-stage embryos. *Nat Med* 15, 577-583.

Varlet I, Canard B, Brooks P, Cerovic G and Radman M (1996) Mismatch repair in Xenopus egg extracts: DNA strand breaks act as signals rather than excision points. *Proc Natl Acad Sci U S A* 93, 10156-10161.

Varlet I, Radman M and Brooks P (1990) DNA mismatch repair in Xenopus egg extracts: repair efficiency and DNA repair synthesis for all single base-pair mismatches. *Proc Natl Acad Sci U S A* 87, 7883-7887.

Venkitaraman AR (2002) Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell* 108, 171-182.

Verlinsky Y, Cieslak J, Ivakhnenko V, Evsikov S, Wolf G, White M, Lifchez A, Kaplan B, Moise J, Valle J *et al* (1998) Preimplantation diagnosis of common aneuploidies by the first- and second-polar body FISH analysis. *J Assist Reprod Genet* 15, 285-289.

Vilkki S, Tsao JL, Loukola A, Poyhonen M, Vierimaa O, Herva R, Aaltonen LA and Shibata D (2001) Extensive somatic microsatellite mutations in normal human tissue. *Cancer Res* 61, 4541-4544.

Vincze T, Posfai J and Roberts RJ (2003) NEBcutter: A program to cleave DNA with restriction enzymes. *Nucleic Acids Res* 31, 3688-3691.

Vinson RK and Hales BF (2002) DNA repair during organogenesis. *Mutat Res* 509, 79-91.

Wang H and Hays JB (2002a) Mismatch repair in human nuclear extracts. Quantitative analyses of excision of nicked circular mismatched DNA substrates, constructed by a new technique employing synthetic oligonucleotides. *J Biol Chem* 277, 26136-26142.

Wang H and Hays JB (2002b) Mismatch repair in human nuclear extracts. Time courses and ATP requirements for kinetically distinguishable steps leading to tightly controlled 5' to 3' and aphidicolin-sensitive 3' to 5' mispair-provoked excision. *J Biol Chem* 277, 26143-26148.

Wang Y, Friedl W, Lamberti C, Jungck M, Mathiak M, Pagenstecher C, Propping P and Mangold E (2003) Hereditary nonpolyposis colorectal cancer: frequent occurrence of large genomic deletions in MSH2 and MLH1 genes. *Int J Cancer* 103, 636-641.

Wang Y, Puscheck EE, Lewis JJ, Trostinskaia AB, Wang F and Rappolee DA (2005) Increases in phosphorylation of SAPK/JNK and p38MAPK correlate negatively with mouse embryo development after culture in different media. *Fertil Steril* 83 Suppl 1, 1144-1154.

Waters SB and Akman SA (2001) A new assay to quantify in vivo repair of G:T mispairs by base excision repair. *Mutat Res* 487, 109-119.

Watson AJ and Barcroft LC (2001) Regulation of blastocyst formation. *Front Biosci* 6, D708-D730.

Watson P and Lynch HT (1993) Extracolonic cancer in hereditary nonpolyposis colorectal cancer. *Cancer* 71, 677-685.

Wells D, Bermudez MG, Steuerwald N, Malter HE, Thornhill AR and Cohen J (2005a) Association of abnormal morphology and altered gene expression in human preimplantation embryos. *Fertil Steril* 84, 343-355.

Wells D, Bermudez MG, Steuerwald N, Thornhill AR, Walker DL, Malter H, Delhanty JD and Cohen J (2005b) Expression of genes regulating chromosome segregation, the cell cycle and apoptosis during human preimplantation development. *Hum Reprod* 20, 1339-1348.

Wells D and Delhanty JD (2001) Preimplantation genetic diagnosis: applications for molecular medicine. *Trends Mol Med* 7, 23-30.

Wells D and Patrizio P (2008) Gene expression profiling of human oocytes at different maturational stages and after in vitro maturation. *Am J Obstet Gynecol* 198, 455-459.

White PJ, Borts RH and Hirst MC (1999) Stability of the human fragile X (CGG)(n) triplet repeat array in Saccharomyces cerevisiae deficient in aspects of DNA metabolism. *Mol Cell Biol* 19, 5675-5684.

Wijnen J, van der KH, Vasen H, Khan PM, Menko F, Tops C, Meijers HH, Lindhout D, Moller P and Fodde R (1998) MSH2 genomic deletions are a frequent cause of HNPCC. *Nat Genet* 20, 326-328.

Wilton L (2002) Preimplantation genetic diagnosis for aneuploidy screening in early human embryos: a review. *Prenat Diagn* 22, 512-518.

Wilton L, Thornhill A, Traeger-Synodinos J, Sermon KD and Harper JC (2009) The causes of misdiagnosis and adverse outcomes in PGD. *Hum Reprod* 24, 1221-1228.

Wimmer K and Etzler J (2008) Constitutional mismatch repair-deficiency syndrome: have we so far seen only the tip of an iceberg? *Hum Genet* 124, 105-122.

Wirthner R, Wrann S, Balamurugan K, Wenger RH and Stiehl DP (2008) Impaired DNA double-strand break repair contributes to chemoresistance in HIF-1 alpha-deficient mouse embryonic fibroblasts. *Carcinogenesis* 29, 2306-2316.

Wong ML and Medrano JF (2005) Real-time PCR for mRNA quantitation. *Biotechniques* 39, 75-85.

Wood RD, Mitchell M and Lindahl T (2005) Human DNA repair genes, 2005. *Mutat Res* 275-283.

Wood RD, Mitchell M, Sgouros J and Lindahl T (2001) Human DNA repair genes. *Science* 291, 1284-1289.

Wu W, Wang M, Mussfeldt T and Iliakis G (2008) Enhanced use of backup pathways of NHEJ in G2 in Chinese hamster mutant cells with defects in the classical pathway of NHEJ. *Radiat Res* 170, 512-520.

Wu X, Wilson TE and Lieber MR (1999) A role for FEN-1 in nonhomologous DNA end joining: the order of strand annealing and nucleolytic processing events. *Proc Natl Acad Sci U S A* 96, 1303-1308.

Yu Z, Chen T, Hebert J, Li E and Richard S (2009) A mouse PRMT1 null allele defines an essential role for arginine methylation in genome maintenance and cell proliferation. *Mol Cell Biol* 29, 2982-2996.

Zeng F, Baldwin DA and Schultz RM (2004) Transcript profiling during preimplantation mouse development. *Dev Biol* 272, 483-496.

Zhang QM and Dianov GL (2005) DNA repair fidelity of base excision repair pathways in human cell extracts. *DNA Repair (Amst)* 4, 263-270.

Zheng P, Patel B, McMenamin M, Paprocki AM, Schramm RD, Nagl NG, Jr., Wilsker D, Wang X, Moran E and Latham KE (2004) Expression of genes encoding chromatin regulatory factors in developing rhesus monkey oocytes and preimplantation stage embryos: possible roles in genome activation. *Biol Reprod* 70, 1419-1427.

Zheng P, Schramm RD and Latham KE (2005) Developmental Regulation and In Vitro Culture Effects on Expression of DNA Repair and Cell Cycle Checkpoint Control Genes in Rhesus Monkey Oocytes and Embryos. *Biol Reprod* 1359-1369.

Ziebe S, Lundin K, Loft A, Bergh C, Nyboe AA, Selleskog U, Nielsen D, Grondahl C, Kim H and Arce JC (2003) FISH analysis for chromosomes 13, 16, 18, 21, 22, X and Y in all blastomeres of IVF pre-embryos from 144 randomly selected donated human oocytes and impact on pre-embryo morphology. *Hum Reprod* 18, 2575-2581.

Zou GM, Luo MH, Reed A, Kelley MR and Yoder MC (2007) Ape1 regulates hematopoietic differentiation of embryonic stem cells through its redox functional domain. *Blood* 109, 1917-1922.

APPENDIX

A: Introduction

Table A.1: DNA glycosylases and endonucleases involved in human base excision repair (Christmann *et al.*, 2003; Ronen and Glickman, 2001)

Enzyme	Gene Symbols	Chromosome	Mammalian Protein Size	Substrate Specificity in DNA
Uracil-DNA glycosylases	*UNG1*	12	227 aa, 25.8 kDa	Uracil
3-Methyladenine-DNA glycosylase	*MPG*	16	294 aa, 32 kDa	3-Methylpurines, hypoxanthine
Thymine mismatch-DNA glycosylase	*MBD4*	12q24	410 aa, 46 kDa	Uracil and thymine paired with guanine
MYH	*MYH, MUTYH*	1p34.3-p32.1 16 exons	429-546 aa, 47-60 kDa	Adenine paired with 8-oxoguanine, Adenine paired with guanine (Binds thymine and guanine paired with 8-oxoguanine without apparent catalysis)
NEIL1	*hMMH1*	-	-	Formamidopyrimidines oxidised pyrimidines (e.g thymine glycol)
NEIL2	*NEIL2*	-	-	5-Hydroxyuracil; 5-hydroxycytosine
NEIL3	*NEIL3*	-	-	Fragmented and oxidised pyrimidines
NTH1 or NTH1	*NTH1*, *OCTS3*	16p13.3 6 exons (8.9 kb)	304/312 aa, 33.5/34.4 kDa	Ring-saturated, oxidised and fragmented pyrimidines
8-Oxoguanine DNA glycosylase	*OGG1*	3p26.2 6 exons (7.8 kb)	345/351/424 aa, 39/47 kDa	8-Oxoguanine paired with cytosine, thymine and guanine
Single-strand selective monofunctional uracil DNA glycosylase	*SMUG1*	12	270 aa, 29.8 kDa	Uracil, high efficiency and preference for ssDNA
Thymine DNA glycosylase	*TDG*	12q24.1 10 exons	410 aa, 46 kDa	Guanine paired with uracil is the preferred substrate. Removes thymine from G:T, C:T and T:T and uracil from G:U (but not A:U) mispairs. Does not remove thymine or uracil from ssDNA.
AP endonuclease	*HAP1, APEX*	14q11.2-12	35.5 kDa	Apurinic and apyrimidinic sites

Table A.2: Expression of DNA repair genes in mammalian oocytes and preimplantation embryos (from mRNA analysis)

DNA repair pathway	Gene	Oocyte GV-MII	1 to 4-cell* stage embryo	4 to 10-cell stage embryo	Morula	Blastocyst	Hatched blastocyst	Specie	Reference
Mismatch Repair	*MLH1*	Low Medium	Low	Low	Low	Low	Low >	Rh. Monkey Human	(Zheng *et al.*, 2005) (Menezo, 2007)
	MSH2	Low High	Low <	Low >	Low	Low <	Low <	Rh. Monkey Human	(Zheng *et al.*, 2005) (Menezo, 2007)
	MSH3	High High	High >	High >	High >	High >	High <	Rh. Monkey Human	(Zheng *et al.*, 2005) Menezo, 2007
	MSH6	Low High	Low <	Low >	Medium	Medium >	Medium >	Rh. Monkey Human	(Zheng *et al.*, 2005) (Menezo, 2007)
	PMS2	Low Medium	Low >	Low <	Low	Low	Low >	Rh. Monkey Human	(Zheng *et al.*, 2005) Menezo, 2007
Nucleotide Excision Repair	*ERCC2*	Low n.d.	Low <	Low	Low	Low	Low >	Rh. Monkey Human	(Zheng *et al.*, 2005) (Menezo, 2007)
	XPC	Low High	Low	Low <	Low <	Low	Low	Rh. Monkey Human	(Zheng *et al.*, 2005) (Menezo, 2007)
Base Excision Repair	*APEX*	Medium High	Medium <	Medium	Medium	Medium	Medium >	Rh. Monkey Human	(Zheng *et al.*, 2005) Menezo, 2007
	MBD4	Low Medium	Low >	Low <	Low	Low	Low >	Rh. Monkey Human	(Zheng *et al.*, 2005) (Menezo, 2007)
	UNG	Low High	Low <	Low	Low <	Low	Low <	Rh. Monkey Human	(Zheng *et al.*, 2005) (Menezo, 2007)
	XRCC1	Low High+	Low	Low <	Low <	Low >	Low	Rh. Monkey Human	(Zheng *et al.*, 2005) (Menezo, 2007)
	OGG1	n.d. Medium	n.d.	n.d.	n.d.	n.d.	n.d.	Rh. Monkey Human	(Zheng *et al.*, 2005) (Menezo, 2007)

Table A.2 (continued): Expression of DNA repair genes in mammalian oocytes and preimplantation embryos (from mRNA analysis)

DNA repair pathway	Gene	Oocyte GV-MII	1 to 4-cell* stage embryo	4 to 10-cell stage embryo	Morula	Blastocyst	Hatched blastocyst	Specie	Reference
Double Strand Break Repair	*BRCA1*	Medium	Medium	Medium <	Medium >	Medium >	Medium <	Rh. Monkey	(Zheng *et al.*, 2005)
		Low	n.d. / Low	High	Low	Low >	Low	Human	(Wells *et al.*, 2005b)
		High						Human	(Menezo, 2007)
	BRCA2	Medium	Low	Low >	Low >	Medium	High	Human	(Wells *et al.*, 2005b)
		Medium						Human	(Menezo, 2007)
	G22P1	Low	Low <	Low >	Low	Low >	Low <	Rh. Monkey	(Zheng *et al.*, 2005)
		n.d.						Human	(Menezo, 2007)
	MRE11A	Low	Low <	Low <	Low	Low	Low >	Rh. Monkey	(Zheng *et al.*, 2005)
		High						Human	(Menezo, 2007)
	NBS1	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	Rh. Monkey	(Zheng *et al.*, 2005)
	RAD50	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	Rh. Monkey	(Zheng *et al.*, 2005)
		Medium						Human	(Menezo, 2007)
	RAD51	Medium	Medium <	Medium >	Medium >	Medium >	High	Rh. Monkey	(Zheng *et al.*, 2005)
		Low						Human	(Menezo, 2007)
	RAD54	Low	Low	Low >	Low <	Low >	Low >	Rh. Monkey	(Zheng *et al.*, 2005)
		Medium						Human	(Menezo, 2007)
Post-Replication Repair	*UBE2A*	Medium	Medium <	Medium <	Medium	Low	Low	Rh. Monkey	(Zheng *et al.*, 2005)
		n.d.						Human	(Menezo, 2007)
	UBE2B	Medium	Medium	Medium >	Medium >	Medium >	Medium <	Rh. Monkey	(Zheng *et al.*, 2005)
		n.d.						Human	(Menezo, 2007)

Table A.2 (continued): Expression of DNA repair genes in mammalian oocytes and preimplantation embryos (from mRNA analysis)

DNA repair pathway	Gene	Oocyte GV-MII	1 to 4-cell* stage embryo	4 to 10-cell stage embryo	Morula	Blastocyst	Hatched blastocyst	Specie	Reference
Cell Cycle Checkpoint Control	*APC*	High	Low	Low >	Low	Medium	Medium >	Human	(Wells *et al.*, 2005b)
	BUB1	Low	n.d./ Low	Low >	Low	Low >	High	Human	(Wells *et al.*, 2005b)
	CDKN1B	Low	Low	Low >	Low >	Low >	Low >	Rh. Monkey	(Zheng *et al.*, 2005)
	MAD2	High	Low	Low >	Low <	Low >	high	Human	(Wells *et al.*, 2005b)
	MDM2	Low	Low	Low <	Low >	Low	Low <	Rh. Monkey	(Zheng *et al.*, 2005)
	MTBP	Low	Low >	Low <	Low <	Low <	Low >	Rh. Monkey	(Zheng *et al.*, 2005)
	PLK1	Low	Low <	Low <	Low >	Low	Low	Rh. Monkey	(Zheng *et al.*, 2005)
	PLK3	Low	Low <	Low >	Low <	Low	Low <	Rh. Monkey	(Zheng *et al.*, 2005)
	RB1	Low	Low <	Low >	Low	Low >	High	Human	(Wells *et al.*, 2005b)
	TP53	Low	Low >	Low >	Low <	Low >	Low <	Rh. Monkey	(Zheng *et al.*, 2005)
		Low	Low <	Low >	Low >	Medium	High	Human	(Wells *et al.*, 2005b)
		Medium						Human	Menezo, 2007
DNA Damage Sensors	*ATM*	High	High	High >	High <	High >	High >	Rh. Monkey	(Zheng *et al.*, 2005)
		Low	Low <	Low >	Low	Medium	High	Human	(Wells *et al.*, 2005b)
	ATR	Low	Low <	Low >	Low	Low	Low >	Rh. Monkey	(Zheng *et al.*, 2005)
		High						Human	(Menezo, 2007)
	BLM	Low	Low	Low <	Low >	Low >	Low >	Rh. Monkey	(Zheng *et al.*, 2005)
	CHEK1	Medium	Medium <	Medium <	Medium	Medium <	Medium >	Rh. Monkey	(Zheng *et al.*, 2005)
	CHEK2	Low	Low <	Low	Low <	Low >	Low	Rh. Monkey	(Zheng *et al.*, 2005)
		High						Human	(Menezo, 2007)
	RFC1	n.d./Low	n.d.	Low	Low >	Low >	Low >	Rh. Monkey	(Zheng *et al.*, 2005)
	PCNA	High	High <	High <	High <	High <	High >	Rh. Monkey	(Zheng *et al.*, 2005)
	TOPBP1	Low	Low <	Low >	Low >	Medium	Medium >	Rh. Monkey	(Zheng *et al.*, 2005)

Gene expression levels are relative and comparable only for same specie / study.

* prior to genome activation; Rh. Monkey: Rhesus monkey; n.d.: not detected /expressed

>: Increase in expression level relatively to previous stage; <: Decrease in expression level relatively to previous stage

B: Materials and Methods

General laboratory chemicals were obtained from VWR International (UK) and were of AnalaR grade. This includes all general reagents like ethanol and salts used for the preparation of buffers.

Adenosine 5'-triphosphate disodium salt hydrate (ATP), Beta-mercaptoethanol (Beta-ME), Bovine serum albumin (BSA), Bromophenol blue, Dimethylsulphoxide (DMSO), Dithiothreitol (DTT), Ethidium bromide, Ethylene diamine tetraacetic acid (EDTA), Formamide (deionised), Glucose, Glutathione (reduced), Glycerol, Igepal, Phenol red, Polyvinyl alcohol (PVA), Sodium acetate (3M solution), Sodium dodecyl sulfate (SDS), Tricine and Tween 20 were of molecular grade and supplied by Sigma Chemical Company (UK). All reagents were stored at room temperature except for ATP and glutathione, which were stored at 4ºC, and BSA, which was stored at -20ºC. ExoI and SAP were obtained from New England Biolabs (UK) and PCR grade Proteinase K was obtained from Roche (UK). All enzymes were stored at -20ºC.

A CP323P Sartorius scale was used to weigh the necessary chemicals. All solutions were prepared with deionised Millipure water and sterilised by autoclaving before use.

B.1 Solutions for single cell isolations

PBS/PVA	PBS supplemented with 0.1% (w/v) PVA
PBS/PVA/ RNasin®**	0.75% (v/v) or 0.3U/µl RNasin® in PBS/PVA
Dissociation buffer (DB)*	0.8g NaCl, 0.02g KCl, 0.005g Na H_2PO_4, 0.1g Glucose, 0.1g EDTA, 0.1g $NaHCO_3$ and 0.010g phenol red in 100ml Millipure water **Prior to use, 40µl of 10mg/ml BSA were added to 960µl of dissociation medium
Alkaline lysis buffer (ALB)**	200mM NaOH and 50mM DTT

*stored at 4ºC
**freshly prepared prior to use

B.2 Solutions for DNA extraction from whole blood

TKM1 *(low salt buffer)*	10mM Tris-HCL (pH 7.6), 10mM KCl, 10mM $MgCl_2$ and 2mM EDTA
TKM2 *(high salt buffer)*	10mM Tris-HCL (pH 7.6), 10mM KCl, 10mM $MgCl_2$, 0.4mM NaCl and 2mM EDTA
TE buffer	10mM Tris-HCl and 1mM EDTA (pH 8.0)

B.3 Solutions for Agarose Gel Electrophoresis

10× TBE buffer[+] pH 8.0	90mM Tris-HCl pH 8.0, 90mM Boric acid and 2mM EDTA disodium salt (Na_2EDTA)
Loading buffer	70% (v/v) 1×TBE and 30% Glycerol + 0.25% (w/v) Bromophenol blue
Ethidium Bromide solution	10mg/ml ethidium bromide

[+] 1× TBE was prepared from the dilution of stock 10× TBE buffer

B.4 Solutions for separation of DNA strands using Dynabeads®

2× B&W buffer*	10mM Tris-HCl (pH 7.5), 1mM EDTA and 2M NaCl
1× B&W buffer*	1 in2 dilution of 2× B&W buffer
Sodium Acetate	3M Sodium Acetate solution purchased from Sigma
0.1M NaOH*	1M NaOH was freshly prepared and diluted (1/10) to 0.1M

*stored at 4ºC

**0.1M NaOH can be stored at -20º for up to 3 months.

B.5 DNA Purification using the MinElute PCR Purification kit (QIAGEN, UK)

Protocol:

5 volumes of PB buffer were added to one volume of PCR product (250µl to 50µl amplified DNA) and mixed well by pipetting. The colour was checked for suitable pH≤7.5, if necessary the pH was adjusted with sodium acetate solution (pH 5.0). The mixture was transferred to a MinElute column placed in a 2ml tube (provided) and centrifuged at 13,000 rpm (10,000g) for 30 seconds at room temperature in a bench-top microcentifuge (MSE Microcentaur, Sanyo, UK). The flow-through was discarded and the MinElute column was placed back in the same tube. 750µl of buffer PE were added to the column and the centrifugation step was repeated as above, discarding the flow-through and replacing the MinElute column in the same tube. This was followed by another centrifugation for 1 minute at 13,000 rpm (10,000g). The MinElute column was then transferred to a new collection tube. 10µl of buffer EB were added to the centre of the MinElute column membrane. The new tube with the column was left to stand for one minute and then centrifuged for 1 minute at 13,000 rpm (10,000g). The MinElute column was discarded and the purified DNA (free from primers, dNTPs, enzymes or salts) was obtained in the collection tube. This DNA was immediately used or stored at 4ºC.

Solution composition:

Buffer PB	contains guanidine HCl and isopropanol, add pH indicator
Buffer PE	Add 96-100% ethanol (~80% v/v)
Elution buffer EB	10mM Tris.Cl (pH 8.5)

B.6 Nicking of heteroduplex DNA substrates

1X NE Buffer 2	10mM Tris-HCl, 50mM NaCl, 10 mM $MgCl_2$, 1mM dithiothreitol (DTT) (pH 7.9 at 25°C)

B.7 Solutions for Exposure of Heteroduplex Constructs to Nuclear Extracts

MMR Reaction Mix	6.67ng/μl substrate (heteroduplex construct), 6.67g/l nuclear extract and 50ng/μl BSA
Solution 1	20mM Tris-HCl (ph 7.6), 1.5mM ATP, 1mM Glutathione, 0.1mM of each of the 4 dNTPs, 5mM $MgCl_2$ and 110mM KCl
Stop Solution	25mM EDTA, 0.67% SDS, 90μg/ml Proteinase K
Nuclear Extracts dilution buffer^{++}	20mM Hepes (pH 7.9), 100mM KCl, 1mM MgCl2, 20% glycerol, 0.5mM PMSF and 0.5mM DTT

$^{++}$ Buffer in which the nuclear extracts were provided

C: Results

Figure C.1: ABI Prism™ 3100 panels showing sequences obtained from selected homozygous DNA samples used to prepare G.T and A.C heteroduplexes

a. Homozygous control A/A

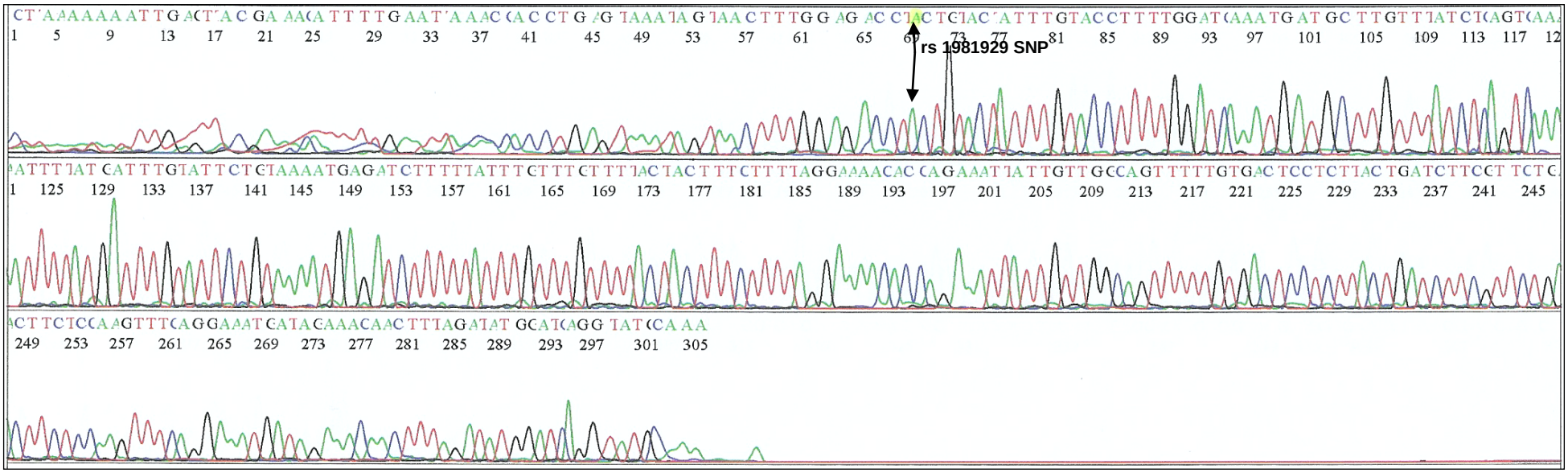


Figure C.1 (continued): ABI Prism™ 3100 panels showing sequences obtained from selected homozygous DNA samples used to prepare G.T and A.C heteroduplexes

b. Homozygous control G/G

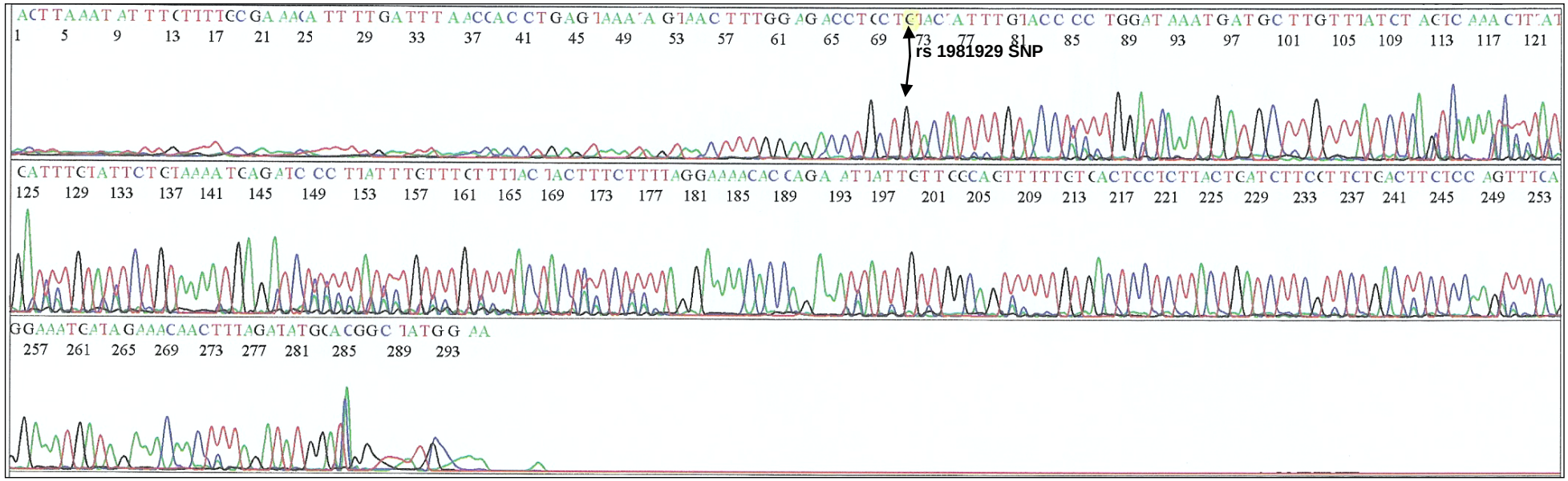
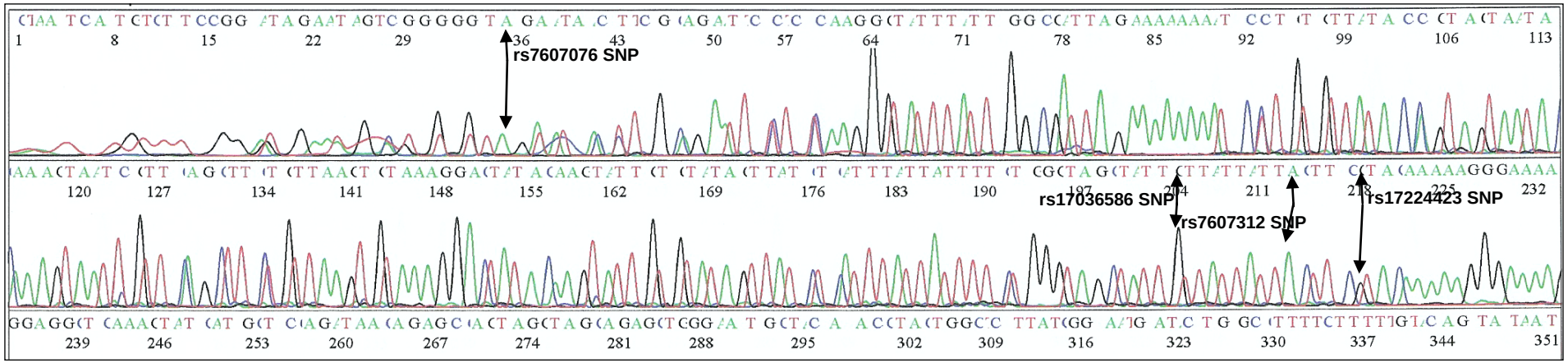


Figure C.2: ABI Prism™ 3100 panels showing the sequences obtained with primer set A (MSH2_4SNPs) for the selection of informative SNPs for the couple undergoing PGD for the *MSH2* exon 8 deletion

a. Female partner



b. Proband (male)

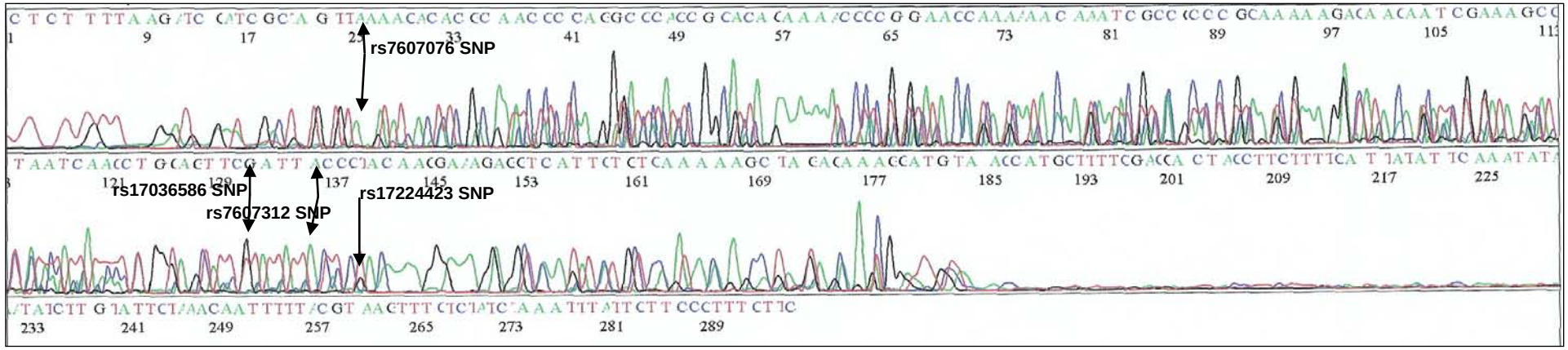


Figure C.2 (continued): ABI Prism™ 3100 panels showing the sequences obtained with primer set A (MSH2_4SNPs) for the selection of informative SNPs for the couple undergoing PGD for the *MSH2* exon 8 deletion

c. Affected relative

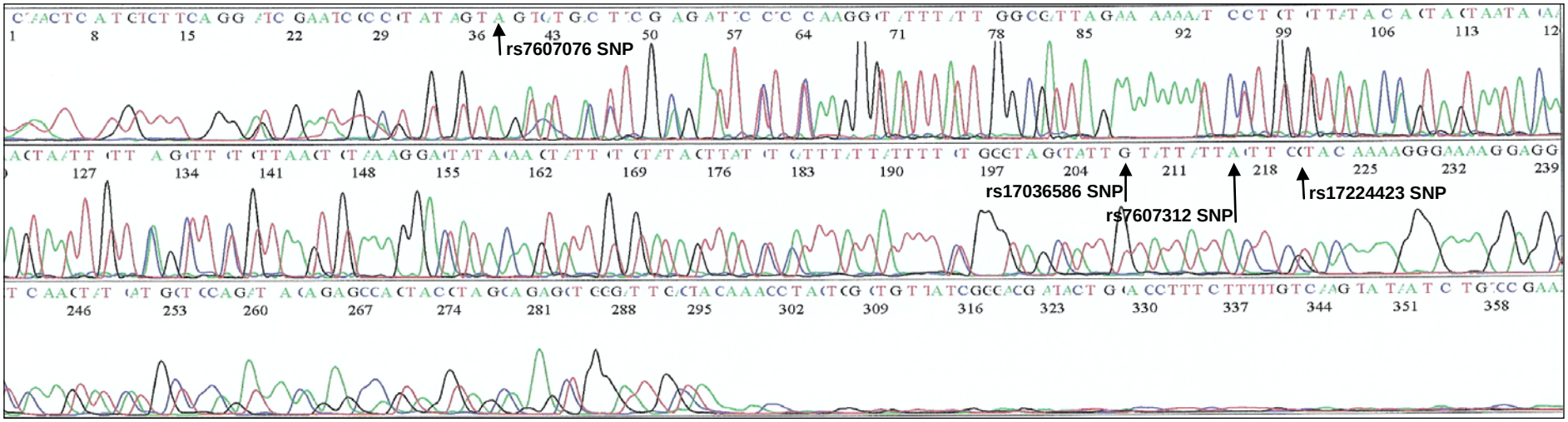
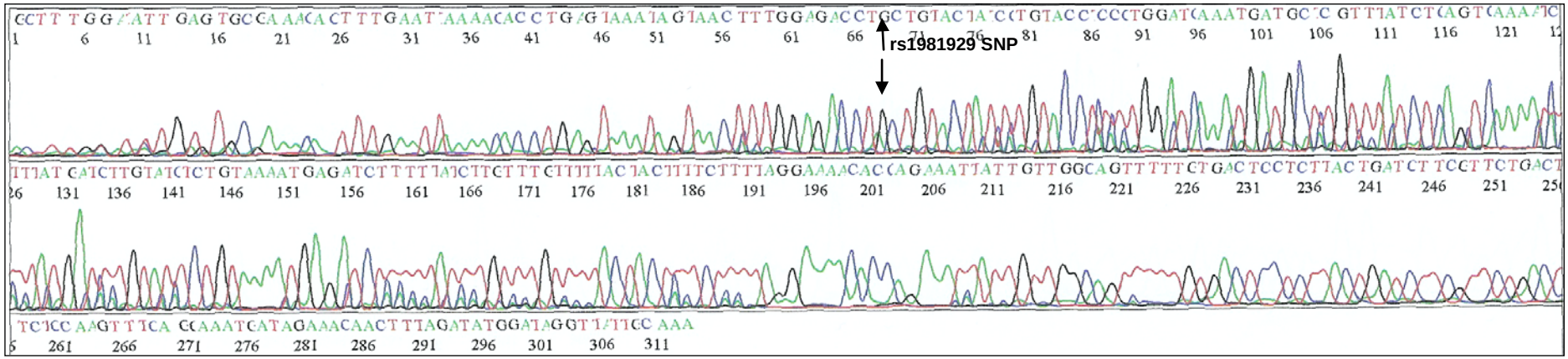


Figure C.3: ABI PrismTM 3100 panels showing the sequences obtained with primer set B (rs1981929) for informativity assessment of the rs1981929 SNP for the couple undergoing PGD for the MSH2 exon 8 deletion

a. Female partner



b. Proband (male)

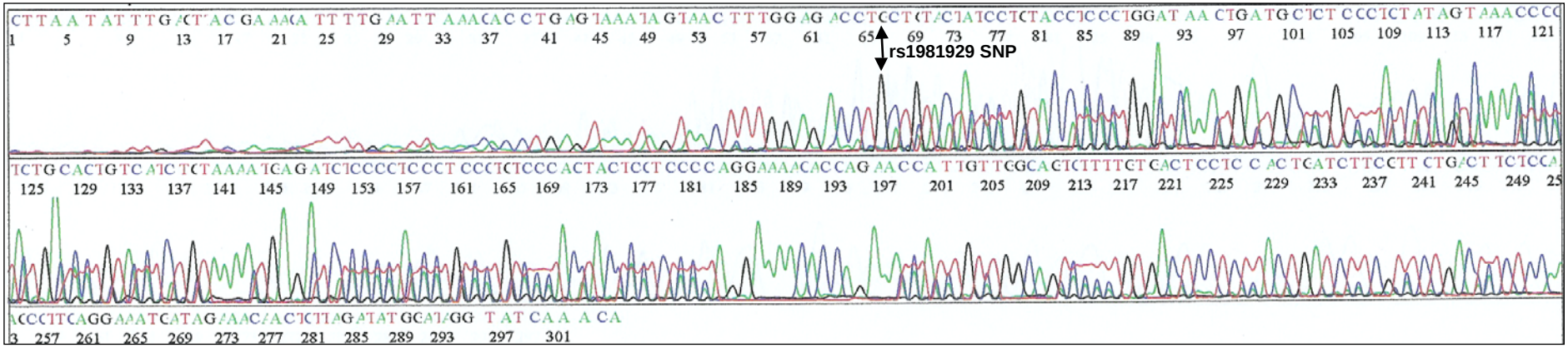


Figure C.3 (continued): ABI PrismTM 3100 panels showing the sequences obtained with primer set B (rs1981929) for informativity assessment of the rs1981929 SNP for the couple undergoing PGD for the MSH2 exon 8 deletion

c. Affected relative

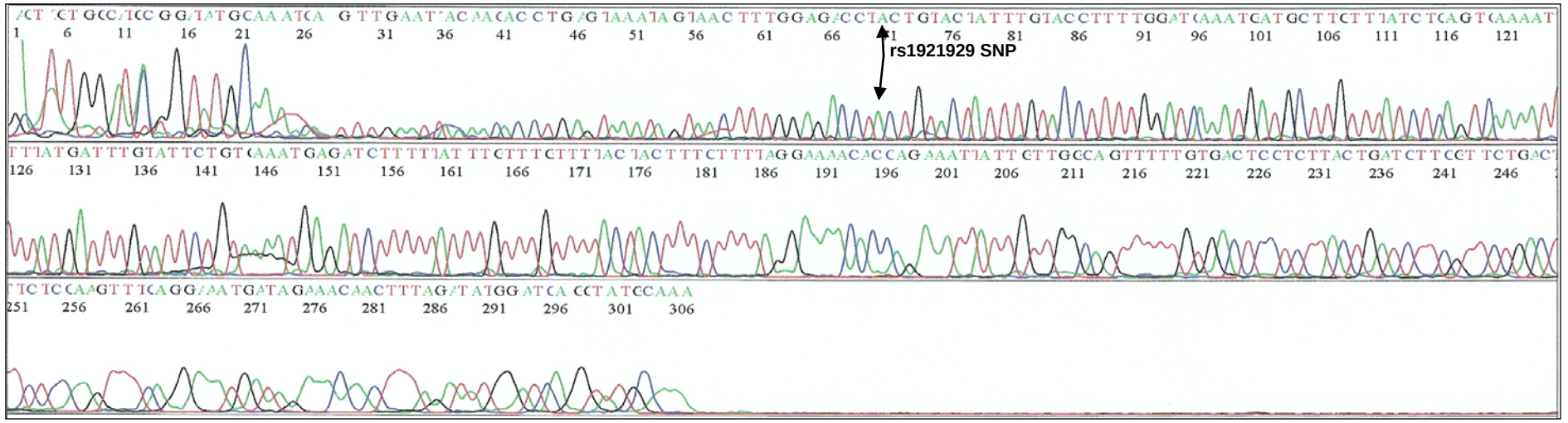
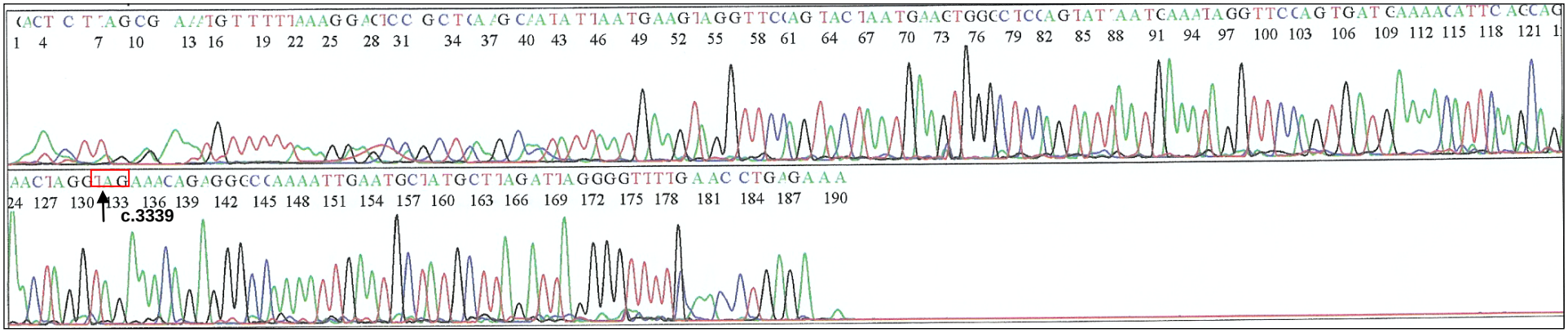


Figure C.4: ABI Prism™ 3100 panels showing the sequences obtained with the *BRCA1* c.3339T>G primer set to confirm the mutation specified in the patient's genetic report

a. Female partner



b. Proband (male)

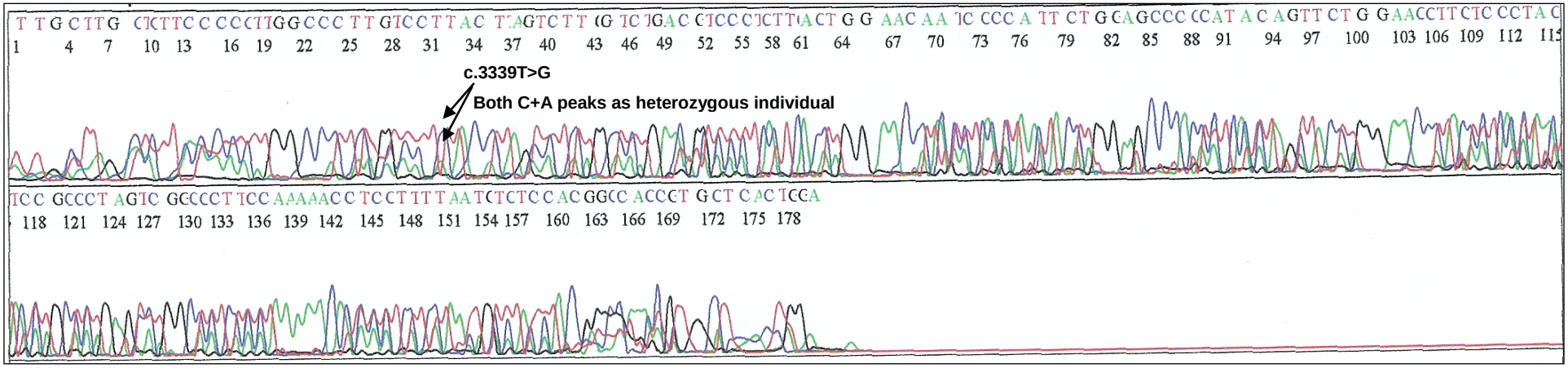


Figure C.4 (continued): ABI Prism™ 3100 panels showing the sequences obtained with the *BRCA1* c.3339T>G primer set to confirm the mutation specified in the patient's genetic report

c. Unaffected relative

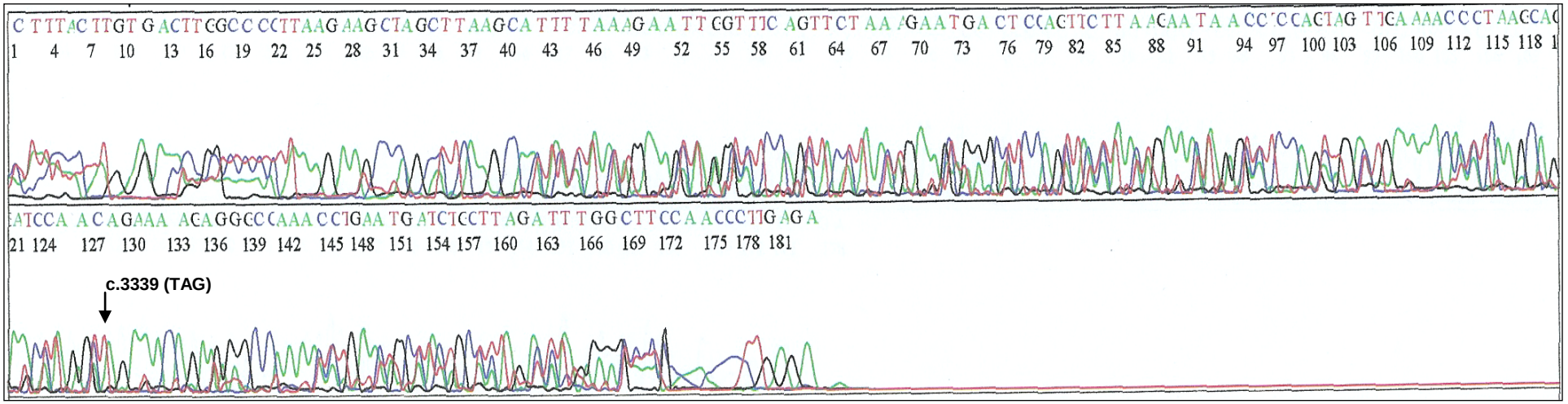


Table C.1: Complete list of R-values and ΔR-values calculated for all heteroduplex samples exposed to nuclear extracts (listed in Table 3.13)

Sample number	Loop size	Nick	Calculation of R-value*	R-value unprocessed substrate	R-value negative control	R-value substrate + N.E.	ΔR post exprosure to nuclear extracts
1	3	no	p.h. 582 strand/p.h. 585 strand	0.96	0.99	0.98	0.01
2	3	long	p.h. nicked 585 strand/p.h. 582 strand	1.25	1.28	1.32	-0.04
3	3	long	p.h. nicked 585 strand/p.h. 582 strand	1	0.92	0.58	0.34
4	3	long	p.h. nicked 585 strand/p.h. 582 strand	1.25	1.28	1.17	0.11
5	3	short	p.h. nicked 582 strand/p.h. 585 strand	0.75	0.72	0.62	0.1
6	3	long	p.h. nicked 585 strand/p.h. 582 strand	0.83	1.34	1.34	0
7	3	short	p.h. nicked 582 strand/p.h. 585 strand	1.22	1.31	1.31	0
8	3	short	p.h. nicked 582 strand/p.h. 585 strand	1.78	1.76	1.73	0.03
9	3	long	p.h. nicked 585 strand/p.h. 582 strand	1	0.92	0.7	0.22
10	3	short	p.h. nicked 582 strand/p.h. 585 strand	0.75	0.72	0.52	0.2
11	3	short	p.h. nicked 582 strand/p.h. 585 strand	1.21	1.28	1.19	0.09
12	3	short	p.h. nicked 582 strand/p.h. 585 strand	1.78	1.74	1.69	0.05
13	3	no	p.h. 582 strand/p.h. 585 strand	1.04	0.93	0.91	0.02
14	3	no	p.h. 582 strand/p.h. 585 strand	1.06	1	1.01	0.01
15	3	short	p.h. nicked 582 strand/p.h. 585 strand	2.69	2.66	2.6	-0.06
16	3	short	p.h. nicked 582 strand/p.h. 585 strand	0.75	0.72	0.65	0.07
17	3	no	p.h. 582 strand/p.h. 585 strand	0.96	0.99	0.98	0.01
18	3	long	p.h. nicked 585 strand/p.h. 582 strand	1.25	1.28	1.28	0
19	3	no	p.h. 582 strand/p.h. 585 strand	1.06	1	0.98	0.01
20	3	short	p.h. nicked 582 strand/p.h. 585 strand	1.04	0.93	0.94	0.01

* The numbers indicate the sizes (in bases) of the two strands in the heteroduplex construct; p.h.: peak height; + N.E.: exposed to nuclear extract

$\Delta R = R_{\text{negative control}} - R_{\text{sample exposed to N.E.}}$, where R = peak height of short or nicked strand/peak height of complementary strand.

Table C.1 (continued): Complete list of R-values and ΔR-values calculated for all heteroduplex samples exposed to nuclear extracts (listed in Table 3.13)

Sample number	Loop size	Nick	Calculation of R-value	R-value unprocessed substrate	R-value negative control	R-value substrate + N.E.	ΔR post exprosure to nuclear extracts
21	21	no	p.h. 561 strand/p.h. 582 strand	0.94	0.86	0.81	0.05
22	21	long	p.h. nicked 582 strand/p.h. 561 strand	1.8	1.76	1.89	-0.13
23	21	long	p.h. nicked 582 strand/p.h. 561 strand	1.8	1.76	1.75	0.01
24	21	short	p.h. nicked 561 strand/p.h. 582 strand	2	1.02	0.95	0.07
25	21	no	p.h. 561 strand/p.h. 582 strand	1.38	1.48	1.36	0.12
26	21	no	p.h. 561 strand/p.h. 582 strand	1.5	1.6	1.62	-0.02
27	21	no	p.h. 561 strand/p.h. 582 strand	1.5	1.6	1.3	0.3
28	21	no	p.h. 561 strand/p.h. 582 strand	0.94	0.86	0.83	0.03
29	21	long	p.h. nicked 582 strand/p.h. 561 strand	1.8	1.76	1.87	-0.11
30	21	long	p.h. nicked 582 strand/p.h. 561 strand	1.68	1.59	1.6	-0.01
31	24	no	p.h. 561 strand/p.h. 585 strand	0.93	0.9	0.81	0.09
32	24	short	p.h. nicked 561 strand/p.h. 585 strand	0.89	0.91	0.87	0.04
33	24	no	p.h. 561 strand/p.h. 585 strand	1.49	1.59	1.62	-0.03
34	24	no	p.h. 561 strand/p.h. 585 strand	1.49	1.59	1.3	0.29
35	24	short	p.h. nicked 561 strand/p.h. 582 strand	1.68	1.59	1.4	0.19
36	24	no	p.h. 561 strand/p.h. 585 strand	1.48	1.6	1.55	0.05
37	24	no	p.h. 561 strand/p.h. 585 strand	1.38	1.48	1.37	0.11
38	24	long	p.h. nicked 585 strand/p.h. 561 strand	1.69	1.59	1.58	0.01

* The numbers indicate the sizes (in bases) of the two strands in the heteroduplex construct; p.h.: peak height; + N.E.: exposed to nuclear extract

$\Delta R = R_{\text{negative control}} - R_{\text{sample exposed to N.E.}}$, where R = peak height of short or nicked strand/peak height of complementary strand

D: Concluding remarks & future work

Table D.1: List of DNA repair genes to be investigated in human oocytes and preimplantation embryos for mRNA expression analysis

Gene	DNA damage response pathway	Expression in preimplantation embryos according to other studies	Specie	Reference
APEX	BER	Moderate	Rh. Monkey	(Zheng *et al.*, 2005)
*OGG1**	BER	Not detected	Rh. Monkey	(Zheng *et al.*, 2005)
*UNG**, *MPG**, *SMUG1**,*TDG**	BER	?	-	-
*XRCC1**	BER	Low	Rh. Monkey	(Zheng *et al.*, 2005)
BUB1	CCCC	Low to high at HB stage	Human	(Wells *et al.*, 2005b)
TP53	CCCC	Low Low to high at HB stage	Rh. Monkey Human	(Zheng *et al.*, 2005) (Wells *et al.*, 2005b)
*PCNA**	DNA polymerase/DNA damage detection	High	Rh. Monkey	(Zheng *et al.*, 2005)
BRCA1	DSBR	Moderate Low overall; High at 4-cell to 10–cell stage	Rh. Monkey Human	(Zheng *et al.*, 2005) (Wells *et al.*, 2005b)
*BRCA2**	DSBR	Low to high at HB stage	Human	(Wells *et al.*, 2005b)
MRE11A	DSBR	Low	Rh. Monkey	(Zheng *et al.*, 2005)
*RAD50**	DSBR	Not detected	Rh. Monkey	(Zheng *et al.*, 2005)
RAD51	DSBR	Moderate; High at HB stage	Rh. Monkey	(Zheng *et al.*, 2005)
*RAD54**	DSBR	Low	Rh. Monkey	(Zheng *et al.*, 2005)
XRCC5 (Ku70)	DSBR	?	-	-
ATM	DSB Repair & DNA damage detection	High Low to high at HB stage	Rh. Monkey Human	(Zheng *et al.*, 2005) (Wells *et al.*, 2005b)
MLH1	MMR	Low	Rh. Monkey	(Zheng *et al.*, 2005)
MLH3	MMR	?	-	-
*MSH2**	MMR	Low	Rh. Monkey	(Zheng *et al.*, 2005)
*MSH3**	MMR	High	Rh. Monkey	(Zheng *et al.*, 2005)
*MSH5**	MMR	?	-	-
MSH6	MMR	Low	Rh. Monkey	(Zheng *et al.*, 2005)
*PMS1**	MMR	?	-	-
PMS2	MMR	Low	Rh. Monkey	(Zheng *et al.*, 2005)
*CDK7**	NER	?	-	-
ERCC1	NER	?	-	-
ERCC2	NER	Low	Rh. Monkey	(Zheng *et al.*, 2005)
*LIG1**	NER	?	-	-
*ATR**	NER & DNA damage detection	Low	Rh. Monkey	(Zheng *et al.*, 2005)

*DNA repair genes found to be differentially expressed in human blastocysts compared to MII oocytes according to our microarray results
?: No data available in the literature; CCCC: Cell Cycle Checkpoint Control; HB: Hatched Blastocyst

POSTERS & PUBLICATIONS FROM THIS PROJECT

A. Published articles:

Jaroudi S, Kakourou G, Cawood S, Doshi A, Ranieri DM, Serhal P, Harper JC and SenGupta SB (2009) Expression profiling of DNA repair genes in human oocytes and blastocysts using microarrays. *Human Reproduction* 24, 2649-2655.

Jaroudi S and SenGupta SB (2007) DNA repair in mammalian embryos. *Mutation Research* 635, 53–77.

B. Conference abstracts:

British Society of Human Genetics Conference, Warwick, UK 31 August-2 September 2009

Kakourou G, Jaroudi S, Gotts S, Doshi A, Serhal P, Harper JC and SenGupta SB (2009) Gene expression profiling of human oocytes and embryo blastocysts. *J Med Genet* 46 Suppl 1, S92, 2.75.

Mammalian DNA Repair Gordon Research Conference, Ventura, CA, USA 8-13 February 2009

Jaroudi S, Kakourou G, Cawood S, Doshi A, Ranieri DM, Serhal P, Harper JC and SenGupta SB. Expression profiling of DNA repair genes in human oocytes and blastocysts using microarrays.

International Society for Prenatal Diagnosis, Vancouver, Canada 1-4 June 2008

SenGupta SB, Dhanjal S, Mamas T, Jaroudi S, Kakourou G, Fordhan K, Doshi A, Gotts S, Serhal P, Harper JC, Delhanty J (2008) Preimplantation Genetic Diagnosis for Cancer Predisposition. *Prenatal Diagnosis* 28 Suppl 3, 2-3.

British Society of Human Genetics Conference, York, UK 17-19 September 2007

Mamas T, Dhanjal S, Jaroudi S, Kakourou G and SenGupta SB (2007) Preimplantation Genetic Diagnosis for inherited cancer predisposition – Different Strategies. *J Med Genet* 44 Suppl 1, S69, 1.62.

European Society of Human Genetics Conference, Nice, France 16-19 June 2007

Jaroudi S, Harper JC and SenGupta SB (2007) Functinal Assessment of DNA loop repair in human nuclear extracts. *Eur Journal Human Genet* 15 Suppl 1, 317, P1276.