UNIVERSITY OF BIRMINGHAM

The Investigation of Cell Free DNA and Genetic Causes in the Evaluation of Miscarriage

By

Emily Charlotte Colley

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

Metabolism and Systems Research

College of Medical and Dental Sciences

University of Birmingham

September 2020

Abstract

Approximately one in four pregnancies result in pregnancy loss, with 1 % of couples suffering from recurrent miscarriage. Genetic investigations are recommended after three consecutive miscarriages and require pregnancy tissue. This thesis investigates the use of cell-free fetal DNA from maternal blood to identify chromosomal abnormalities in recurrent miscarriage. 123 cell-free DNA samples were collected from first trimester miscarriages (5+1 to 11+6 weeks) and 62 % of samples had a corresponding routine cytogenetic result. 54 % of genetic aberrations were correctly identified using single end sequencing on a HiSeq2500 and the WISECONDOR pipeline, and 68 % of genetic aberrations were correctly identified using paired end sequencing on a NextSeq and the Illumina pipeline. This thesis also investigates other genetic causes of miscarriage in euploid pregnancies and includes a published systematic review and whole exome sequencing in seven euploid miscarriage trios. The systematic review identified 50 studies which included variants thought to be causative of miscarriage, and the whole exome sequencing cohort demonstrated the difficulty in interpreting variants without human phenotype ontology or genes without a reported disease-association. However, identification of genes/ variants that are causative of or predisposing to miscarriage will be of significant individual patient impact with respect to counselling and treatment.

Dedication

I would like to dedicate this thesis to my family, especially my Nanny and Grandad, who always believed I can do anything I set my mind to and never saw me reach the end of this.

Acknowledgements

I would like to say a huge thank you to my supervisors, Dr. Stephanie Allen, Dr. Neil Morgan and Professor Arri Coomarasamy who have supported and helped me throughout my PhD. Especially to Steph who has expanded my genetics knowledge and given me support and friendship.

I would like to thank the amazing Tommy's team at Birmingham Women's Hospital and acknowledge the support they give to grieving families at this hard time. I would like to thank the Tommy's midwives and research nurses, Oonagh Pickering, Debbie Bullen, Tracy Bingham and Rachel Boothe for the support and recruitment through this trial. I would also like to thank the Tommy's doctors, Abby Eapen, Bala Karunakaran, Paul Smith, Justin Chu, Rima Smith, Hoda Harb, Laurentiu Craciunas, Jay Ghosh, Hajra Khattak, James Cheshire, Catherine Dunlop and Ioannis Gallos, who offer their support to families and see patients during the Tommy's miscarriage clinic. I would like to thank Helen Williams and Hannah Jeffries for project coordination support, Adam Devall for ethics and REC support and Chloe Trainor and Melissa Lowndes for admin and HR support.

I would like to thank West Midlands Regional Genetics Laboratory, which has supported this research trial and included it within the laboratory. Special thanks go to Rachel Jennings and Samantha Court who trained me in lab based NGS techniques, to Mohammed Omer who helped with half of the library preparation for the cfDNA sequencing and to Louise brown and Emma Ankers who trained me and carried out tissue preparation of POC samples. I would like to thank Edward Davis and Kate Stephens who received samples within the specimen reception and booked these samples in to the lab management system. I would

like to thank the scientific team at West Midlands Regional Genetics Laboratory who analysed the routine results, with special thanks to Sue Hamilton, Fiona Togneri, Elizabeth Young and Claire Vince. I would also like to say a massive thanks to Same Clokie and Chipo M-Wolfgarten who implemented the Bioinformatics pipeline and completed the bioinformatics analysis.

I would like to thank my sister, Rosanna Colley, who has provided some of the illustrations for the figures included in this thesis and to Aurelio Tobias who provided statistical help and guidance.

I would like to thank Lieve Page-Christians, Philippa Burns, Patrizia Di Pietro, Cosmin Deciu, Sarah Kinnings, Eoin Brown, Victoria Corey, Kathryn Robinson, Sven Schaffer, Suchi Bhatt, Barbara Baggiani, Christina Stahler, Marion Burdin and Jasmin Dehnhardt at Illumina who helped our collaboration to be a success.

I would like to thank Tommy's Baby Charity for the funding to make this PhD possible and the fantastic work and support they give to families to ensure the best possible pregnancy outcomes so families can take home happy, healthy babies. Finally, I would like to thank the families who consented to take part in this trial and altruistically gave samples which will help families in the future not to go through the loss they have. I hope these and all families get their rainbow babies.

Table of Contents

List of illustrations

List of tables

List of definitions and abbreviations

aCGH	Array-comparative genomic hybridisation
ACGS	Association for Clinical Genomic Science
ACMG	American College of Medical Genetics
AMC	Arthrogryposis multiplex congenital
BAC	Bacterial artificial chromosome
βhCG	Beta human Chorionic Gonadotrophin
BGI-Agilent-V6	Sequence provider: BGI, Library prep: Agilent V6 kit
cfDNA	Cell free DNA
CMA	Chromosomal microarray
CNV	Copy number variant
CRL	Crown-rump length
CVS	Chorionic villus sampling
DDD	Deciphering Developmental Disorders
DMR	Differentially methylated regions
DR	Detection rate
EPL	Early pregnancy loss
FADS	Fetal akinesia deformation sequence
FF	Fetal fraction
GSD-IV	Glycogen storage disease-Type IV
HM	Hydatidiform mole
HPO	Human phenotype ontology
IUFD	Intra-uterine fetal demise
LMP	Last menstrual period
LMPS	Lethal multiple pterygium syndrome
MAX-AF	Maximum allele frequency
MCC	Maternal call contamination
MeDIP	Methylated DNA immunoprecipitation
MLID	Multi-locus imprinting disturbances
MPLA	Multiplex ligation-dependent probe amplification
MPSS	Massively parallel shotgun sequencing
NaOH	Sodium hydroxide
NCV	Normalised chromosome values

NGS	Next generation sequencing
NIHF	Non-immune hydrops fetalis
NIPD	Non-invasive prenatal diagnosis
NIPT	Non-invasive prenatal testing
Oligo	Oligonucleotide
OMIM	Online Mendelian Inheritance in Man
PAGE	Prenatal Assessment of Genomes and Exomes
PE-NextSeq-Illumina	Illumina's automated VeriSeq NIPT V2 CE-IVD using paired end sequencing reads on the NextSeq and analysed using Illumina's all chromosome pipeline.
PGD	Pre-genetic diagnosis
PGT	Pre-implantation genetic testing
POC	Products of conception
QF-PCR	Quantitative fluorescent polymerase chain reactivation
RCOG	Royal College of Obstetrics and Gynaecology
RHM	Recurrent hydatidiform mole
RM	Recurrent miscarriage
RPL	Recurrent pregnancy loss
SE-HiSeq-WC	Single end sequencing reads on the HiSeq2500 and analysed using WISECONDOR.
SNP	Single nucleotide polymorphism
TE-Agilent-SS	Sequencing provider: Theragen Etex, Library prep: Agilent Sureselect V6 exome capture kit.
TOP	Termination of pregnancy
UoB-TruSeq	Sequencing provider: University of Birmingham Sequencing Service, Library prep: TruSeq exome kit.
UPD	Uniparental disomy
VOUS	Variant of unknown significance
VUS	Variant of uncertain significance
WES	Whole exome sequencing
WGS	Whole genome sequencing
WISECONDOR	Within-SamplE Copy Number aberration DetectOR
XCI	X-chromosome inactivation

Chapter 1: Introduction

Chapter 1- Introduction

Miscarriage or early pregnancy loss (EPL) is the most common adverse outcome of pregnancy and is defined as the loss of an embryo or fetus before 22 weeks gestation. One in four clinical pregnancies results in pregnancy loss ((NICE), 2012, King et al., 2000) and of these, 1% of couples will go on to suffer from recurrent miscarriage. The Royal College of Obstetrics and Gynaecologists (RCOG) define recurrent miscarriage as the loss of three or more consecutive pregnancies (Regan et al., 2011). Recurrent miscarriage is the most common complication during early pregnancy (Larsen et al., 2013a). This is clearly a very distressing time for the couple and regrettably for many of these cases the cause of the RPL is not identified.

1.1 Epidemiology of miscarriage

Whilst it is difficult to estimate the exact number of pregnancies that end in loss, it is quoted that 1 in 4-5 pregnancies end in miscarriage. Approximately two thirds of conceptions are never recognised as a clinical pregnancy, with 30% of pregnancies lost prior to implantation, and 30% lost following implantation but before the missed period. These are collectively known as preclinical losses. Only 30% of conceptions result in a live birth and 10% end in miscarriage (Grossmann et al., 2004, Ramond et al., 2017). Prospective cohort studies using urinary human chorionic gonadotropin (βhCG) have revealed that the human conception rate is 40% over a 12 months period (Grossmann et al., 2004). The outcomes of human conceptions can be visualised and summarised in the pregnancy loss iceberg shown below in Figure 1.

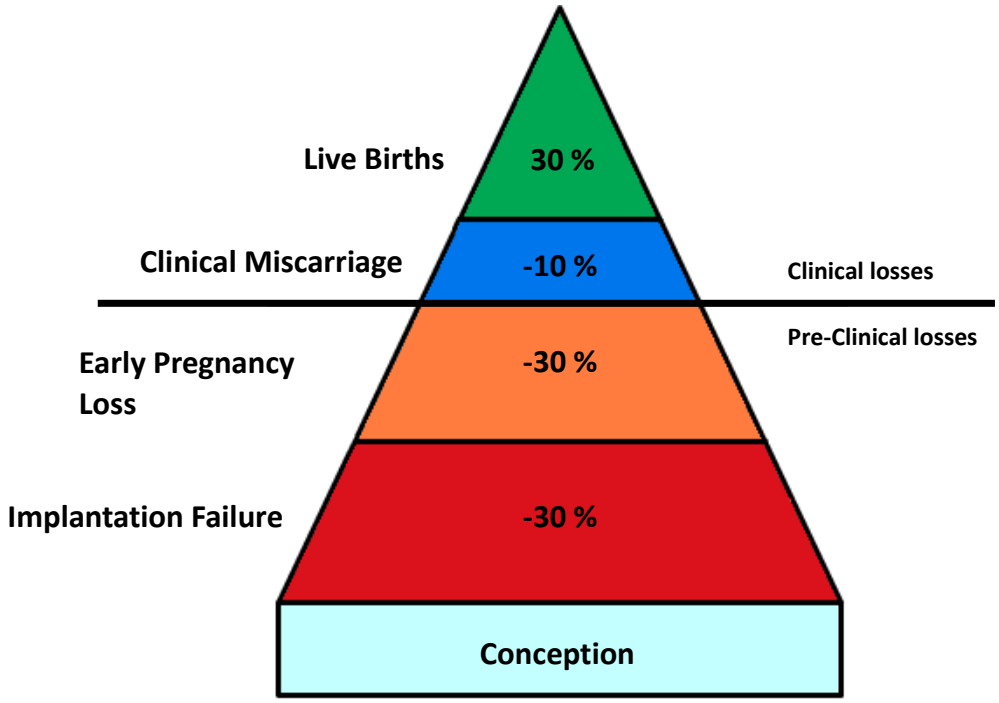


Figure 1- The Pregnancy Loss Iceberg

Pre-clinical and clinical pregnancy losses from conception. Adapted from (Macklon et al., 2002).

1.2 Causes of pregnancy loss

Whilst it is generally accepted that pregnancy loss occurs when the fetus is incompatible with life due to structural malformations or chromosomal aberrations, there are many other risk factors associated with miscarriage. These include lifestyle factors, endocrine factors, immunological causes, infections, thrombophilia, sperm DNA fragmentation and uterine malformations (Larsen et al., 2013a).

1.2.1 Life Style Factors

Life style factors could have a negative impact on pregnancy and certain life styles, such as smoking, caffeine and alcohol consumption should be reduced or stopped during pregnancy. Weight loss should ideally be achieved before conception occurs.

Smoking during pregnancy can cause pregnancy-related problems including miscarriage, fetal abnormalities, placental abruption, premature birth, low birth-weight and sudden unexpected death in infancy, and much advice and guidance is given on how to stop smoking during pregnancy (Richards et al., 2015). Studies have shown that smoking causes a dose-response detrimental effect to miscarriage (Dong et al., 2019, Ellard, 2019). A systematic review has calculated the relative risk of active smoking to be 1.23 (95% CI), with an increased risk with the amount of cigarettes smoked. However, former smokers had no more risk of miscarriage than non-smokers.

The current guideline for pregnant women is to consume no more than 200 mg of caffeine daily. There is increasing evidence that caffeine consumption increases the risk of spontaneous miscarriage with a relative risk of 1.37 (95% CI) when consuming 300 mg of caffeine daily and a relative risk of 2.32 (95% CI), when consuming 600 mg daily (Bender Atik et al., 2018). Therefore reducing caffeine consumption is highly recommended during pregnancy. However there is no known association between caffeine and infertility.

Consumption of alcohol during pregnancy is associated with an increased risk of pregnancy loss (Kesmodel et al., 2002, Windham et al., 1997). A Danish cohort that monitored alcohol consumption during pregnancy reported that 45% of women consumed some alcohol during pregnancy. In women who reported consuming 2-3.5 drinks weekly and >4 drinks, a hazard

ratio of 1.66 and 2.82 (95 % CI) was calculated with pregnancy loss respectively (Andersen et al., 2012).

Obesity is known to cause complications during pregnancy including infertility, miscarriage, and stillbirth, pre-eclampsia and gestational diabetes (Catalano and Shankar, 2017, Mission et al., 2015). Systematic reviews have identified that obesity increases miscarriage rates for both spontaneous pregnancy losses (Metwally et al., 2008) and recurrent pregnancy losses (Metwally et al., 2010). A meta-analysis including over 28,000 women experiencing miscarriage identified a higher spontaneous miscarriage rate of 13.6 % in obese women (BMI ≥28) compared to normal-weight women (BMI <25). A smaller subgroup analysis showed higher recurrent miscarriage rate in the obese group (46 %) vs the normal-weight group (43 %). Therefore it can be concluded that obesity has a higher association with miscarriage (Boots and Stephenson, 2011).

1.2.2 Endocrinological causes

The most prevalent endocrinological causes of miscarriage are polycystic ovary syndrome (PCOS), obesity, hyperinsulinemia and insulin resistance. Elevated levels of follicular phase serum LH levels (Regan et al., 1990), free androgen (Cocksedge et al., 2008) and prolactin during the follicular phase (Bussen et al., 1999), have also all been associated with pregnancy loss. The thyroid dysfunctions, hyperthyroidism (Millar et al., 1994), hypothyroidism (Khalid et al., 2013) and thyroid autoimmunity (Lata et al., 2013) also have an increased risk of miscarriage.

PCOS has been identified as the most common endocrinological condition in women with RPL (Clifford et al., 1994), with miscarriage occurring in 40 % of women with PCOS (Rai et al.,

2000b). However, there may be an underlying condition such as obesity, hyperinsulinemia or hyperandrogenemia associated with PCOS which is causing the pregnancy loss.

1.2.3 Immunology

During feto-placental development, the conceptus comes into close contact with the maternal cellular immune system and must overcome a potential immune response from the mother. During implantation, IL-10, CSF-1 and TGFβ are expressed in uterine cells and during pregnancy natural killer (NK) cells, macrophages and dendritic cells are expressed in the feto–maternal interface (Giakoumelou et al., 2016). Evidence has suggested that immunological factors could be causative of miscarriage with increased or decreased levels of immunological cells being the cause of the pregnancy loss.

Studies have shown that women with RPL have increased peripheral natural killer (NK) cells compared to normal fertile controls (Kwak et al., 1999, Ntrivalas et al., 2001). NK cells have been demonstrated to recognise placental trophoblastic cells (King et al., 2000) and although they don't directly kill trophoblast cells, they can activate cytokines which target the conceptus (Lachapelle et al., 1996).

Recurrent miscarriage has also been associated with a significantly elevated ratio between immune cells, T-helper 1 and T helper 2 (TH1/ TH2 ratio), in women with RPL compared to healthy multiparous controls (Kwak-Kim et al., 2003). While the levels of CD3+ T cells in peripheral blood do not differ from normal fertile controls and non-pregnant women with a history of RPL (Darmochwal-Kolarz et al., 2002), during the first trimester, CD3+ T cell levels are significantly lower in women who miscarried, compared to those who went on to have a live birth (Kwak et al., 1995).

Similarly, T regulator cell levels were not different in non-pregnant women with RPL compared to controls (Kwak-Kim et al., 2003), but T regulator lymphocytes were lower in women with pregnancy losses compared to controls normal fertile controls (Darmochwal-Kolarz et al., 2002).

1.2.4 Infections

Evidence suggests that 15 % of early miscarriages are caused by possibly preventable infections (Giakoumelou et al., 2016). Some bacterial infections, such as bacterial vaginosis (Donders et al., 2009) and Chlamydia (Baud et al., 2011) and viral infections, such as Herpes Simplex Virus (Kapranos and Kotronias, 2009) and Influenza (Bloom-Feshbach et al., 2011), have all been shown to have an increased risk of miscarriage.

As discussed above, the immune system plays a critical role during pregnancy as the tolerance of the semi-allogeneic fetus is maintained. An infection could disrupt normal biological functions such as implantation, placentation or spiral artery invasion during the pregnancy or disrupt the immunological balance and result in a miscarriage. This has been reviewed in detail (Giakoumelou et al., 2016).

1.2.5 Thrombophilia

Thrombophilia can be either inherited or acquired, leading to blood clotting disorders that can cause pregnancy complications. If excessive blood clotting leads to arterial and/or venous thrombosis at the implantation site or in the placental blood vessels it can result in a pregnancy loss (Abu-Heija, 2014). This is often treated with low-dose Aspirin, however there

is no evidence that this improves live birth rates (Rai et al., 2000a, Schisterman et al., 2014) and could also have a negative effect on the pregnancy.

Antiphospholipid syndrome (APS) is an acquired form of thrombophilia which can be caused by either Lupus anticoagulant or anticardiolipin antibodies which are associated with pregnancy loss. The maternal immune system produces antibodies against the phospholipids on the trophoblast cell membrane and the endogenous protein, β2-glycoprotein 1, which is also produced by the trophoblasts (Abu-Heija, 2014, Tong et al., 2015). Low dose heparin is advised as a treatment for pregnant women with Antiphospholipid syndrome using the RCOG, Green-top Guidelines No. 17 (Regan et al., 2011).

1.2.6 Sperm DNA fragmentation

Men with higher levels of sperm DNA damage have been shown to have a significantly higher risk of pregnancy loss (Robinson et al., 2012). Sperm DNA damage affects both nuclear and mitochondrial DNA and is predominantly caused by six mechanisms, reviewed by Sakkas and Alvarez (Sakkas and Alvarez, 2010). These mechanisms include apoptosis during spermatogenesis, DNA strand breakage during the remodelling of sperm chromatin, post-testicular DNA fragmentation by oxygen radicals or endogenous caspases and endonucleases, and DNA damage induced by radiotherapy, chemotherapy or environmental factors such as smoking or air pollutants.

1.2.7 Structural malformations

A potential cause of recurrent miscarriage could be due to maternal uterine malformations. This is caused when there is abnormal development of the Müllerian ducts during embryogenesis. Uterine malformations include a uterine septum, bicorn ate uterus, uterus didephis or unicornate uterus. In a systematic review (Robinson et al., 2012) of over 89,000 women, the prevalence of uterine malformations was 13.3 % (95 % CI) in patients with pregnancy loss compared to 5.5 % (95 5 CI) in an unselected population (Chan et al., 2011).

While the prevalence of RPL and uterine septum has increased, and has been treated using hysteroscopic metroplasty to improve pregnancy outcomes, very little to no evidence is available to suggest a positive outcome (Kowalik et al., 2011). A randomised control trial, "The randomised uterine septum transection trial" (TRUST), is comparing hysteroscopic septum resection and expectant management in women with recurrent pregnancy loss and a diagnosis of uterine septum (Rikken et al., 2018). This will identify whether hysteroscopic septum resection will improve the rate of live births in woman with recurrent miscarriage and a diagnosed septate uterus.

1.3 Genetics of Miscarriage

Genetic factors are the most common cause of pregnancy loss, with the risk of aneuploidy increasing with increasing maternal age (Hassold and Chiu, 1985, Stephenson et al., 2002). Approximately 50 % of miscarriages are caused by chromosomal abnormalities (Choi et al., 2014, Sugiura-Ogasawara et al., 2012) and the majority of these chromosomal abnormalities are *de novo (Carp et al., 2006).* It is considered that chromosomal errors are more common

in spontaneous miscarriage than recurrent miscarriage, and that fetal chromosome errors are more common before 10 weeks of pregnancy and less common between 12 and 22 weeks. Evidence suggests that of these abnormalities, 86 % are numerical abnormalities, 6 % are structural abnormalities and 8 % are from other genetic mechanisms, including chromosomal mosaicism (Goddijn et al., 2004).

1.3.1 Numerical Abnormalities

Numerical Abnormalities include autosomal trisomies (three copies of a particular chromosome), monosomies (one copy of a particular chromosome) and polyploidies (extra set(s) of chromosomes, e.g. Triploidy and tetraploidy). Where these numerical abnormalities are sporadic they most commonly arise from non-disjunction in meiosis during the production of gametes. Non-disjunction can cause either a loss (monosomy) or a gain (trisomy) of a chromosome. If non-disjunction occurs during meiosis I, the gamete will contain chromatids from homologous chromosomes and if non-disjunction occurs during meiosis II, the gamete will have two chromatids from the same homologue (Figure 2).

The majority of aneuploidies are maternally inherited, usually due to an increase in maternal age, although it has been suggested that around 7% of meiotic errors in chromosomes are of paternal origin (Robinson et al., 1999). Suggested explanations are that spermatocytes with synaptic defects undergo apoptosis during the pachytene stage of prophase I or metaphase I whereas oocytes remain viable, or that during oogenesis, oocytes remain in prophase I for an extended amount of time, from birth to ovulation. Therefore this could result in more genetic meiotic errors compared to sperm where spermatogenesis is a continuous process (Kaser, 2018).

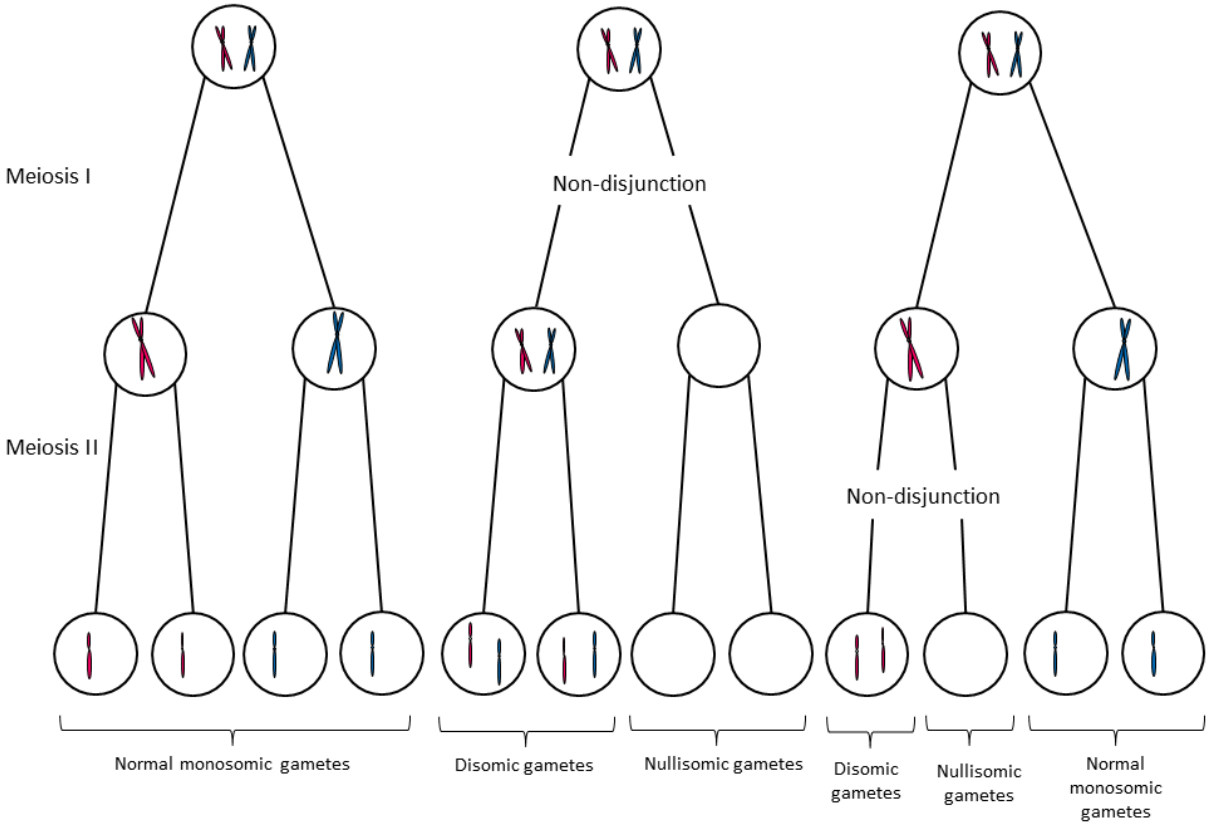


Figure 2- Non-disjunction during gametogenesis

1.3.2 Structural Abnormalities

The majority of numerical abnormalities occur in conceptuses from chromosomally normal parents, however parental structural abnormalities will present in 3-6 % of recurrent miscarriage cases. This gives the fetus a 50% chance of inheriting an unbalanced chromosome translocation (Larsen et al., 2013b, Larsen et al., 2013a). The two main types of chromosomal structural abnormalities are reciprocal translocations (balanced or unbalanced) or Robertsonian translocation (centric fusion of two acrocentric chromosomes), but inversions, deletions, insertions and ring chromosomes can also be identified. Whilst the carriers of balanced translations themselves are usually phenotypically normal, some of the

gametes produced will have unbalanced chromosomes and patients with reciprocal or Robertsonian translocations have an increased risk of RPL or infertility due to transmission of an unbalanced translocation during meiosis.

1.3.2.1 Reciprocal translocations

Reciprocal translocations occur when two non-homologous chromosomes undergo double-strand breakage and exchange segments with ends joining. Translocation carriers are phenotypically normal providing there has been no loss or gain of chromosomal material and the break point has not interrupted a critical gene. During meiosis, reciprocal translocations can form different chromosomal conformations. The chromosomes line up with homologous areas within the chromosome and produce a pachytene quadrivalent. During segregation, this can result in different chromosome imbalances because the chromosomes can separate in different ways.

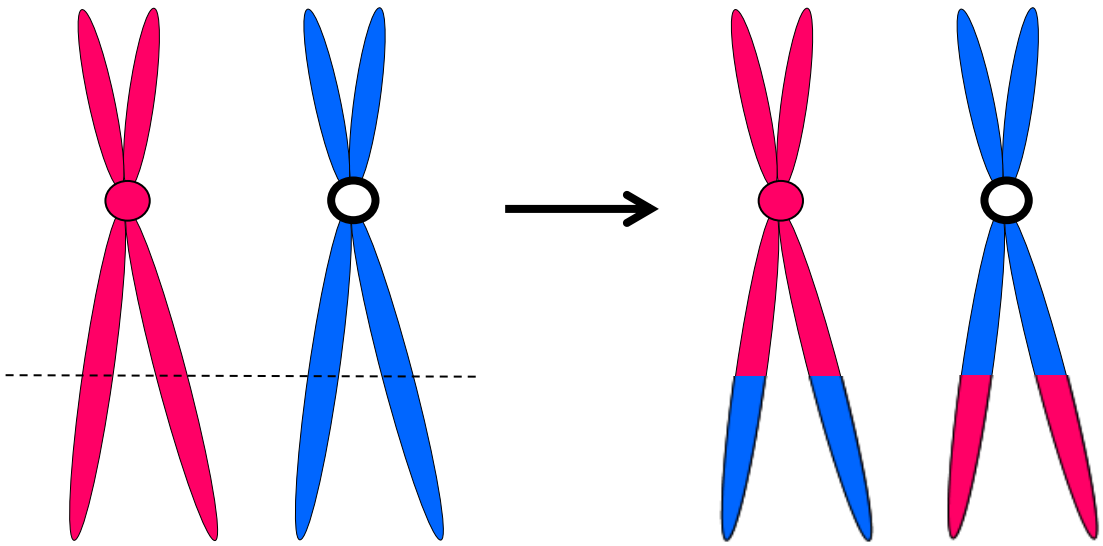


Figure 3- reciprocal translocation

1.3.2.2 Robertsonian translocations

Robertsonian translocations occur when two acrocentric chromosomes (chromosomes 13, 14, 15, 21 or 22) break and recombine at or close to the centromeres, and the two long arms fuse (centric fusion). The short arm (satellite DNA) of each chromosome is lost. However, as the short arms contain genes for ribosomal RNA only and there are multiple copies on acrocentric chromosomes, they are considered as a functionally balanced chromosome. The prevalence of Robertsonian translocations is 1 in 1000, with 13q14q being the most common.

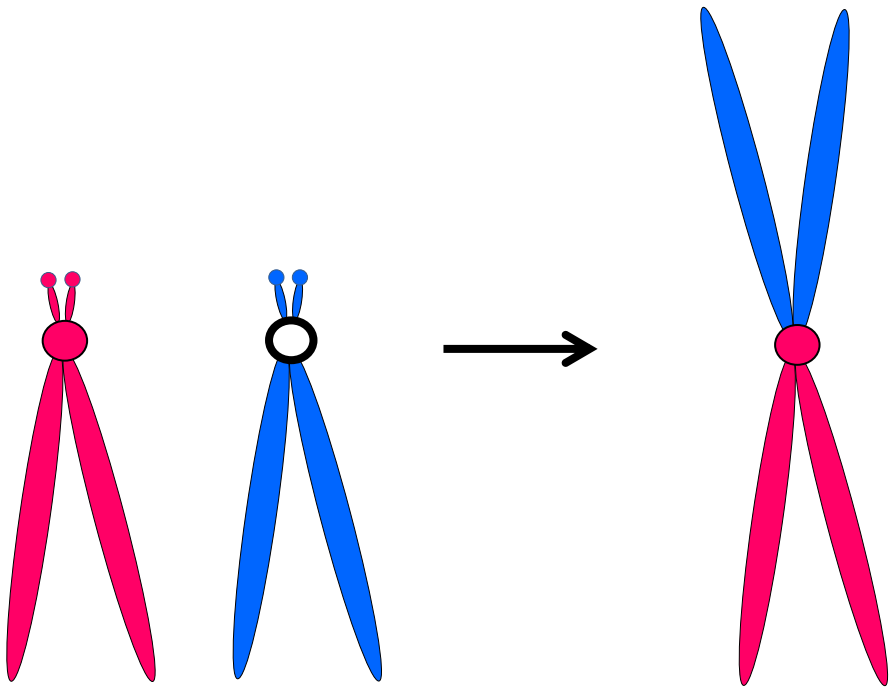


Figure 4- Robertsonian translocation

1.3.2.3 Deletions, Duplications, Insertions, Inversions and Ring Chromosomes

Other structural chromosomal abnormalities include deletions, duplications, insertions and inversions. A deletion is when part of the chromosome is lost and causes a monosomy for that part of the chromosome only and a duplication or insertion is where an additional part of the chromosome is duplicated or inserted.

If a chromosomal segment is inserted into another chromosome, this is known as an insertion. Insertions can remain balanced within a karyotype, but there is a 50 % chance of gametes inheriting either a deletion or insertion.

Inversions occur when a single chromosome has a two break arrangement and the chromosomal segment rotates when it recombines. These can be either pericentric inversions when the centromere is involved or paracentric inversions only one chromosome arm is involved. Although inversions are balanced rearrangements, they can cause problems if an important gene is disrupted, which can lead to chromosomal imbalances in gametes depending on where cross-overs occur in meiosis I.

If a break at each end of the chromosome occurs and the two ends join they produce a ring chromosome. Inevitably, the distal portions of the chromosomes are lost with varying consequences.

1.3.3 Mosaicism

If an individual has two or more cell lines with a different genetic constitution, derived from a single zygote they are defined as mosaic. Mosaicism is usually the result of a mitotic non-disjunction error in the developing conceptus. The timing of when the mitotic error occurs

will reflect in the degree of mosaicism. For example, the earlier the mitotic error the more likely each cell line will be equal. The mosaicism can occur entirely throughout the fetus and placenta, known as generalised mosaicism or it is also possible for the mosaicism to occur after divergence of the placenta and fetus, with the placenta having a different karyotype from the fetus. This is known as placental mosaicism (Kalousek and Dill, 1983, Ariel et al., 1997) and the mosaicism can also be confined to the fetus. Figure 5 below illustrates the three different kinds of mosaicism.

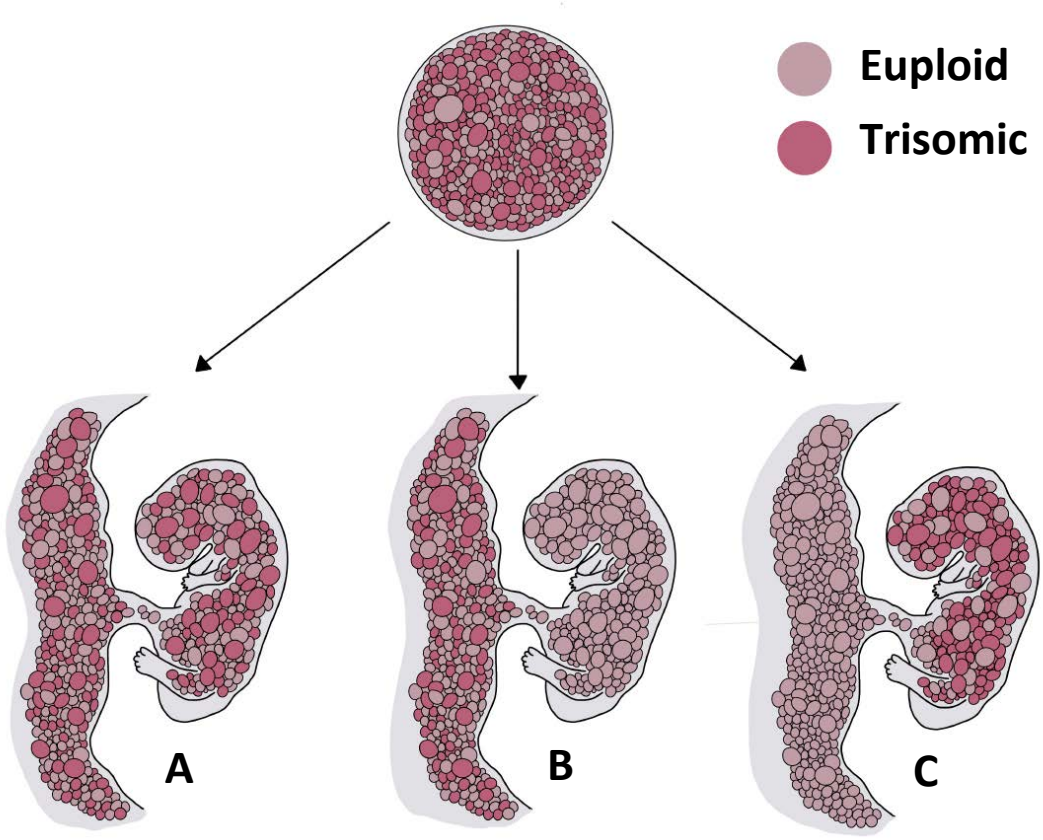


Figure 5- Examples of mosaicism

A, Generalized. *B*, Confined to the placenta. *C*, Confined to the embryo. Illustration provided by Rosanna Colley.

1.3.5 Genetic Abnormalities causing miscarriage

As discussed above, 50 % of miscarriages are due to a chromosomal abnormality. These can be broken down into different anomalies, including but not limited to, trisomies, triploidies, monosomies, tetraploidies and structural chromosome anomalies. Reported abnormality rates can vary between published studies (Table 1). This could be due to gestational age, ascertainment bias, maternal cell contamination (MCC) or tissue culture failure (Hardy and Hardy, 2015).

Table 1– Summary of abnormalities found by cytogenetic analysis of pregnancy loss tissue.

	No. of samples	No. of obtainable results	No. of abnormal results	Single trisomy	Multiple trisomy	Monosomy X	Triploidy	Tetraploidy	Structural rearrangements
Soler et al., 2017(Soler et al., 2017)	1119	1011 (90.3%)	711 (70.3%)	461 (64.9%)	29 (4.2%)	74 (10.4%)	93 (13.1%)	10 (1.4%)	37 (5.2%)
Jenderny et al., 2014*(Jendern y, 2014)	534	390 (73.0%)	237 (60.1%)	125 (52.7%)	17 (7.2%)	18 (7.6%)	29 (12.2%)	25 (10.5%)	10 (4.2%)
Wang et al., 2014(Wang et al., 2014)	5457	4092 (75.0%)	1872 (45.7%)	1236 (66.0%)	not recorded	243 (13.0%)	262 (14.0 %)	75 (4.0%)	37 (2.0%)
Menasha et al., 2005(Menasha et al., 2005)	2180	1920 (88.1%)	1099 (57.2 %)	721 (65.6%)	67 (6.1%)	96 (8.7%)	116 (10.6%)	18 (1.6%)	46 (4.2%)
Eiben et al., 1990(Eiben et al., 1990)	983	750 (76.3%)	380 (50.7%)	229 (60.3%)	7 (1.8%)	40 (10.5%)	47 (12.4%)	17 (4.5%)	18 (4.7%)
Average (%)	**n/a**	**80.5%**	**56.8%**	**61.9%**	**4.8%**	**10.0%**	**12.5%**	**4.4%**	**4.1%**

*including cases where the aberration was mosaic.

It is reported in the literature that Trisomy 16 is the most common trisomy not compatible with life, followed by trisomy 22, which has rarely been reported to have a very brief survival. Trisomies 21, 15 and 18 are reported to be slightly less common, with only trisomy 15 not being observed at term. Trisomy 1 and Trisomy 19 are the rarest with very little documentation of them reported in the literature (Hardy and Hardy, 2015).

1.3.6 Cytogenetic Analysis in Recurrent Miscarriage

Conventional cytogenetic analysis uses whole genome techniques to detect abnormalities in the fetus's entire genome. Although karyotyping is classed as the 'gold standard', it is time intensive and often has limited resolution in detecting micro- deletion and duplication syndromes. Furthermore subtle rearrangements and the origin of small supernumerary marker chromosomes are not detected (van den Berg et al., 2012). The extent of fetal genetic abnormalities and submicroscopic imbalances in miscarriages is unknown.

G-banded-karyotyping of cultured cells often has a high failure rate due to the poor quality of tissues received (McClelland et al., 2011). Thus, laboratories have implemented alternative DNA-based diagnostic approaches (Bruno et al., 2006, Carvalho et al., 2010, Diego-Alvarez et al., 2007, Donaghue et al., 2010). In the West Midlands Regional Genetics Laboratory, chromosomal abnormalities in recurrent miscarriage samples are detected using Quantitative fluorescence polymerase chain reaction (QF-PCR) and chromosomal microarray (CMA) based methods (McClelland et al., 2011). QF-PCR has the advantage of a rapid result for chromosomes 13, 18, 21 and sex chromosomes, whilst also detecting triploidy and whether there is any maternal cell contamination. CMA tests for deletions, duplications and chromosome number abnormalities across all autosomal chromosomes. Both of these methods have a lower failure rate, higher throughput and are less labour intensive than conventional karyotyping, and also allow DNA to be stored for future analysis. This is more cost-effective and allows a higher diagnostic yield.

Currently, genetic testing for the cause of recurrent miscarriage is completed on pregnancy tissues, which comprises of placental villi, fetal skin and umbilical cord, referred to as

products of conception (POC). This tissue needs to be fresh and unfixed so that the tissue can either be cultured or have DNA extraction performed. The problem with using POC is that POC is not always available after a miscarriage and the collection of tissue is very stressful to the woman and sometimes difficult to obtain. Collectable POC is not always present, for example, in cases of very early pregnancy loss, or if the POC is lost down the toilet. POC is often fixed in formaldehyde for histological examinations and therefore rendered unusable for genetic testing as aldehyde fixatives degrade the DNA. Alternatively, the POC is not returned by the patient to allow genetic testing in a laboratory setting. Additionally, POC is often heavily contaminated with maternal material due to the close proximity of the placenta to the maternal circulation, and the subsequent trauma caused during miscarriage. For all of these reasons, another easier, less stressful and reliable way of obtaining the fetal genetic material needs to be addressed.

1.4 Cell Free Fetal DNA

Cell free DNA (cfDNA) was first identified by Dennis Lo (Lo et al., 1997) who demonstrated that male cfDNA can be obtained in ongoing pregnancies from the plasma of pregnant women. It was later discovered that differential methylation of fetal DNA using paternally inherited single nucleotide polymorphisms (SNPs) can be used to detect cfDNA independent of fetal sex (Poon et al., 2002).

1.4.1 What is cell free fetal DNA?

cfDNA is small fragments of DNA which represent the entire fetal genome (Lo et al., 2010). Paired-end sequencing has determined cfDNA to be approximately 160-180 bp in length (Fan et al., 2010). This is considerably smaller than maternally derived DNA (Li et al., 2004).

Although cfDNA is termed as fetal, it is actually derived from the trophoblast cells of the placenta (Alberry et al., 2007) and is present as early as day 22 in singleton pregnancies (Guibert et al., 2003). Trophoblasts are fetally derived cells covering the placental villi. Syncytiotrophoblasts are the outermost layers which come into direct contact with the maternal blood and the cytotrophoblasts are the inner layer which lies on the basement membrane (Figure 6).

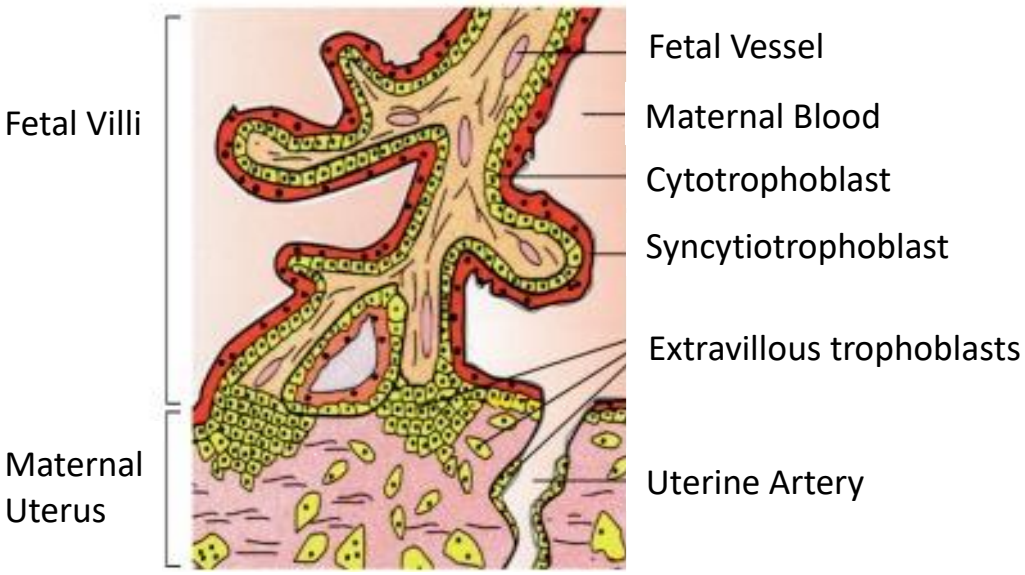


Figure 6- Schematic of Human Placental Villi.

The syncytiotrophoblasts are the outermost layer of the placental villi and the cytotrophoblasts the inner layer of the placental villi. Adapted from (Lavialle et al., 2013).

It was initially thought that cfDNA was derived from fetal blood. However, cfDNA can be detected before the fetal-placental circulation occurs at approximately 28-30 days meaning cfDNA is identifiable before the fetal circulation is established (Guibert et al., 2003). The fetal and placental blood is always separated by the feto-placental barrier. From 12 days post conception, during early implantation, syncytiotrophoblasts invade the maternal endometrium causing disintegration of the maternal endometrial vessels, and fetal capillaries are observable in the mesenchyme of the placental villi between 18 and 20 days. This supports the evidence that cfDNA is derived from the trophoblasts. As the placenta undergoes the natural process of apoptosis, the placentally derived cfDNA is released into the maternal circulation (Figure 7).

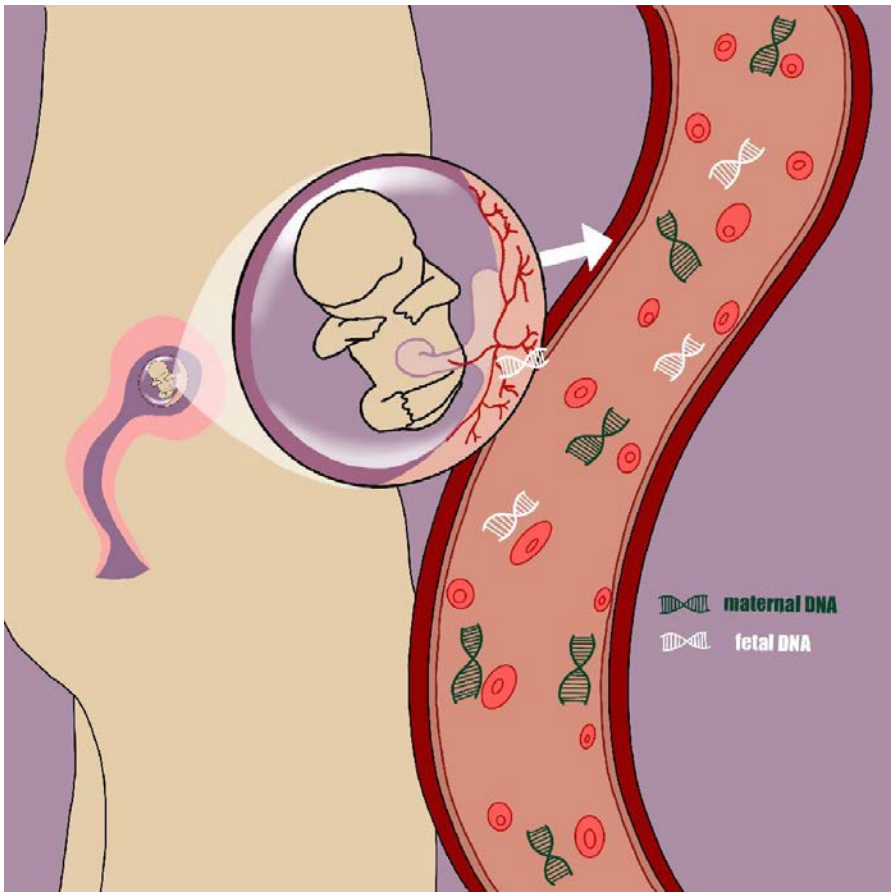


Figure 7- Release of cfDNA in the maternal circulation

Illustration provided by Rosanna Colley.

1.4.2 Clinical applications of cfDNA

CfDNA can be clinically identifiable in maternal serum from as early as 7 weeks gestation and the concentration of cfDNA increases as the pregnancy progresses (Lo et al., 1998). At 11-13 weeks gestation, the average cfDNA in the maternal blood has been assessed at 10%, although it can be affected by many factors, such as maternal weight, fetal crown-rump length, serum free βhCG, serum pregnancy-associated plasma protein-A, smoking and trisomies (Ashoor et al., 2013). The identification of cfDNA in maternal blood and major advances in genomics have led to many different clinical applications in prenatal diagnosis, including fetal sex determination, Rhesus D genotyping, chromosomal aneuploidy detection and diagnosis of monogenic disorders. This means that invasive tests, such as chorionic villus sampling and amniocentesis, in many instances are being replaced by non-invasive tests which don't carry the 1 % miscarriage risk and allows easier and earlier testing during pregnancy (Salomon and Sotiriadis, 2019).

1.4.2.1 Fetal Sex determination

Fetal sex determination was the first application of cfDNA to be applied in a clinical setting. Y chromosome sequences are identified in the maternal plasma and can be amplified to identify a male fetus (Devaney et al., 2011). This is now a widely used test across England (and other countries including the Netherlands, Spain and France) and is reliable (Devaney et al., 2011) and cost effective (Hill et al., 2011).

A systematic review by Devaney et al. has shown that determining fetal sex using the presence of the Y chromosome has 95.4% sensitivity and 98.6% specificity. However,

gestational age and amplification technique could affect the test performance (Devaney et al., 2011). The use of fetal sex determination is important in the management of pregnancies with fetal genital ambiguity, congenital adrenal hyperplasia or at risk of an X-linked condition. In conditions where only males are affected, non-invasive determination of fetal sex reduces the need for invasive testing by as much as 50 % (Daley et al., 2014).

1.4.2.2 Rhesus D (RhD) status

The use of cfDNA analysis has allowed for the non-invasive prediction of fetal rhesus D status to target anti-D prophylaxis. As RhD sequences are absent in maternal D-genomes, the detection of RhD sequences in maternal plasma can determine the rhesus D status of the fetus. Initially the fetal rhesus D status was assessed in male pregnancies using real-time PCR to identify RhD and SRY sequences, with rhesus D males identified correctly (Zhong et al., 2001, Finning et al., 2002). The fetal Rhesus D status can also be correctly identified in female pregnancies (Aykut et al., 2013, Ahmadi et al., 2016). NICE guidelines now recommend high-throughput non-invasive prenatal testing (NIPT) for fetal RhD genotype to reduce the unnecessary use of antenatal anti-D prophylaxis in the prevention of anti-D immunization in RhD-negative patients ((NICE), 2016).

1.4.2.3 Chromosomal aneuploidies

The identification of chromosomal aneuploidies during pregnancy is known as non-invasive prenatal testing (NIPT). NIPT allows for the detection of common fetal aneuploidies 13, 18, and 21 (and the sex chromosomes X and Y) and is now offered in a clinical setting for detection of fetal trisomies (Drury et al., 2016). Invasive testing (amniocentesis and CVS

sampling) carries a small miscarriage risk as a result of the invasive procedure (Salomon et al., 2019). NIPT does not carry this miscarriage risk. However, NIPT for fetal aneuploidy is a screening test, and positive/ high chance results should always be followed-up by an invasive diagnostic test, such as amniocentesis, if termination of pregnancy is being considered. A recent meta-analysis (Mackie et al., 2017) has evaluated the sensitivity and specificity of NIPT using cfDNA in singleton pregnancies. The sensitivity and specificity of the different aneuploidies are shown in Table 2.

Table 2- Sensitivity and specificity of detection of fetal aneuploidy using NIPT

The Sensitivities and specificities of NIPT for trisomy 13, 18, 21 and monosomy X, with a 95% confidence interval (Mackie et al., 2017).

	Sensitivity	**Specificity**
Trisomy 13	0.906 (0.823-0.958)	1.00 (0.999-0.100)
Trisomy 18	0.977 (0.952-0.989)	0.999 (0.998-1.000)
Trisomy 21	0.994 (0.983-0.998)	0.999 (0.999-1.000)
Monosomy X	0.929 (0.741-0.984)	1.00 (0.999-0.100)

NIPT has been has been developed using different technologies such as Next generation sequencing (NGS), microarray analysis and detection of epigenetic modifications, for example using methylation DNA immunoprecipitation (MeDIP).

1.4.2.3.1 Next Generation Sequencing NIPT

There are several different approaches to NIPT using NGS. These include massively parallel shotgun sequencing, targeted sequencing and SNP-based targeted sequencing. Massively

parallel Shotgun Sequencing (MPSS) is a NGS technique which reads a representative of short fetal DNA sequences from the entire fetal genome. CfDNA is extracted from maternal plasma and a library preparation, with each sample indexed, is made from the DNA fragments (Ehrich et al., 2011). Samples are sequenced using MPSS for single-end reads of 36 bp and tens of millions of sequence reads are mapped to a reference genome (Fan et al., 2008). Aneuploidies (trisomy 13, 18 and 21 and monosomy X) are then detected using statistical methods such as, normalised chromosome values (NCVs), where the chromosome of interest is normalised against reference chromosomes (Sehnert et al., 2011). This means that this method cannot be used to detect full triploidy.

Targeted sequencing amplifies and sequences specific target regions of the genome. Unlike MPSS, which generates large amounts of data from randomly analysed genomic fragments across the whole chromosome, targeted sequencing only sequences specific genomic regions and therefore more efficient use of sequencing data. Digital analysis of selected regions (DANSR™) selects loci unique to the individual chromosomes that do not have known polymorphisms and copy number variants to target each chromosome. Each sequence for a selected genomic location is counted and any aneuploidies calculated using a standard Z-test on normalized sequence counts (Sparks et al., 2012). Like MPSS, this cannot detect full triploidy.

Analysis using SNP-based targeted sequencing uses targeted PCR amplification and sequencing of SNPs. This allows the detection of triploidy and uniparental disomy which counting based methods cannot determine (Ryan et al., 2016). Using expected distributions for the possible alleles for the SNP set on the targeted chromosome, the observed allele distribution can determine the aneuploidy status (Zimmermann et al., 2012). This allows for

detection of full trisomies. NGS technologies have the potential to detect small chromosome aberrations but they do require complicated statistical and bioinformatics methods.

1.4.2.3.2 Epigenetic-based NIPT

Methylated DNA immunoprecipitation (MeDIP) uses 5' methyl cytidine specific antibodies to target differentially methylated sites between fetal and maternal cfDNA. There have been more than 2000 differentially methylated regions (DMR) found between maternal and placental DNA using MeDIP (Papageorgiou et al., 2009), and using MeDIP in combination with real time qPCR, cfDNA is hypermethylated compared to maternal DNA. Detection of trisomy 21 can be achieved by comparing the fetal-specific methylated regions (Papageorgiou et al., 2011).

The methylation profile of maternal and placental cfDNA can also be analysed using genomewide bisulfite sequencing. Whole-genome sequencing is completed on bisulphite-converted libraries to analyse DNA methylomes from cfDNA. An algorithm then predicts the placenta methylation profiles which have been associated to pregnancy conditions to non-invasively predict trisomy 21 (Lun et al., 2013).

NIPT using epigenetic-based methods is a low cost and simple method which directly enriches the fetal DNA. However, full conversion of the fetal DNA is rarely achieved and the DNA is prone to degradation (Papageorgiou and Patsalis, 2012). There is also a risk of bias introduced by the assay, and inherent natural variability in methylation between individuals and tissues, and during development.

1.4.2.3.3 Commercially available NIPT Testing

There are currently several NIPT tests available, either through the NHS or privately. These include MaterniT21 PLUS (Sequenom), VeriFi and VeriSeq prenatal tests (VERINATA/Illumina), Harmony prenatal test (ARIOSA diagnostics/ Roche) and Panorama test (NATERA Inc.). The West Midlands Regional Genetics laboratory currently offers the Verifi test, called Lucina, privately or for patients with an increased risk (≥1:50) of trisomy 13, 18 or 21 on combined screening.

Sequenom laboratories offer MaterniT21 using MPSS for whole genome sequencing with enhanced sequencing series for their MaterniT21 NIPT, to report on chromosomes 13, 18 and 21, fetal sex and sex chromosome abnormalities. In addition, Sequenom can also offer the detection of microdeletions and the first genome-wide NIPT (Sequenom, 2017). Verinata Health (Illumina since 2014) offer Verifi prenatal test for trisomies 13, 18 and 21, sex chromosomes and the option for all chromosomes by using MPSS and normalised chromosome values (NCV), for all chromosomes (Illumina, 2017). Harmony prenatal NIPT by ARIOSA diagnostics (Roche since December 2014) is a targeted NIPT test which uses microarray and digital analysis of selected regions (DANSR) to targets specific regions of the chromosome (Roche, 2016).

The Panorama test requires an additional DNA sample from the mother to compare SNPs across all chromosomes between the mother and the fetus. The algorithm, next generation aneuploidy test using SNPs (NATUS) is used to determine the fetal genetic status (Natera, 2016).

1.4.2.4 Monogenic disorders

As well as cfDNA being clinically used for fetal chromosomal aneuploidy in NIPT, fetal monogenic disorders can be identified using cfDNA, known as non-invasive prenatal diagnosis (NIPD). These single gene disorders are identified by first identifying the parental haplotype. The cfDNA isolated from the maternal plasma then undergoes targeted sequencing and the ratios of parental haplotypes present in the cfDNA identify the fetal inheritance of the monogenic disorder. Identification of many monogenic disorders has been demonstrated by NIPD including beta thalassaemia (Chiu et al., 2002, Lam et al., 2012), cystic fibrosis (Gonzalez-Gonzalez et al., 2002), congenital adrenal hyperplasia (Ma et al., 2014, New et al., 2014), congenital deafness (Meng et al., 2014) and Duchenne muscular dystrophy (Xu et al., 2015, Yoo et al., 2015).

1.5 Fetal Fraction

The proportion of cfDNA that is fetally derived is known as the fetal fraction (FF). The level of FF in a sample is important to give accurate results as insufficient FF may give false-negative results. There is limited evidence of the lower threshold of FF needed to produce a result, and many laboratories do not even check the FF. However it has been suggested that the FF should be at least 4 % to minimise the risk of a false-negative risk (Peng and Jiang, 2017).

1.5.1 Estimating Fetal Fraction

There have been huge advances in estimating the FF and there are now many different approaches which are discussed below (Peng and Jiang, 2017). However, methods to calculated FF may be quite variable as FF is difficult to measure accurately, especially at low levels.

1.5.1.1 Y based chromosome approach

The earliest used measurement of FF relies on using Y chromosome based approaches. This method utilizes the Y chromosome present in the male fetus but not present in the mother. Genetic markers, *DYS14* and *SRY,* can be identified using real-time PCR and quantified against reference genomic DNA to determine FF (Zimmermann et al., 2005). The proportion of Y chromosome material can also be used to calculate FF using the percentage of Y chromosome. This is calculated using bioinformatics algorithms and a reference set of euploid pregnancies, in data obtained by massively parallel genomic sequencing (Chiu et al., 2011). Using the Y chromosome is a simple and accurate way of determine FF however it cannot be used for female foetuses.

1.5.1.2 SNP-based methods

SNP-based methods are an accurate way to measure FF for both male and female foetuses. This method relies on SNPs where both the mother and father are homozygous for different genotypes and the fetus is heterozygous for the same SNP. The FF can be quantified by sequencing the cfDNA and using the ratio of fetal alleles to the total number of alleles in the plasma DNA sample (Chu et al., 2010). Although this is a direct and accurate method,

additional analysis is required and both maternal and paternal DNA is required, and may not always be feasible if the biological father's DNA is not available. A similar method, *FetalQuant*, can be used to calculate FF using maternal plasma, however it does require high depth massively parallel sequencing. The FF is calculated using maximum likelihood estimations, assuming that allelic counts at each SNP locus follow a binomial distribution (Jiang et al., 2012).

1.5.1.3 SeqFF

A method to estimate both male and female FF, known as SeqFF, uses single-end, shallow depth sequencing allowing the use of the same data generated during NIPT. This is based on the assumption that maternal and fetal cfDNA have slightly different sequencing behaviour and certain regions have different representation. While these differences are subtle and hard to detect in an individual sample, when these differences are analysed in a large cohort they can be determined. Discrete regions of fetal DNA are overrepresented in some areas of the genome and the more fetal DNA the larger and more evident the difference. Once this relation is known, a statistical model can be trained to quantify the FF. This statistical model aggregates normalised counts of autosomal single-end sequence reads aligned to 50 kb continuously partitioned regions of the human reference genome (hg19) to determine model coefficients using elastic net (Enet) and reduced-rank regression. This model can be applied to other data sets to predict FF (Kim et al., 2015).

1.5.1.4 DNA methylation

DNA methylation profiles vary between maternal and fetal cfDNA (Poon et al., 2002) allowing for FF to be estimated using the different methylation regions in the *RASSF1A* gene (Chan et al., 2006) and by bisulfite sequencing (Lun et al., 2013). Using a methylation-sensitive restriction enzyme digests the hypo-methylated maternal derived *RASSF1A* promoter sequences leaving the methylated fetal derived sequences. By comparing the total copies of DNA against the fetal methylated DNA after restriction enzyme digestion, the FF can be calculated (Chan et al., 2006). Alternatively, the placenta has a hypomethylated profile compared to other tissues and the methylation of distinct genomic loci is specific to placental tissues. cfDNA correlates with the placental methylome (Lun et al., 2013). Using massively parallel bisulfite sequencing, cfDNA can be correlated to the methylation profile of the placenta (Sun et al., 2015), a term referred to as plasma DNA tissue mapping, and thus the fetal fraction can be calculated. However, these methods are expensive and technically complex (Peng and Jiang, 2017).

1.5.1.5 DNA fragment sizes

Fetally derived cfDNA is shorter than maternally derived cfDNA (Chan et al., 2004). Fetal fraction estimation has been developed based on the ratio between DNA fragment sizes as a diagnostic parameter. Using paired-end massively parallel sequencing and a linear regression model, the fetal fraction can be deduced from the size of the DNA fragment size, where 100-150 bp represents the fetal DNA and 163-169 bp represents the maternal DNA (Yu et al., 2014).

1.5.2 Fetal Fraction influences

It is well known that FF increases with gestational age (Song et al., 2016, van Beek et al., 2017). There have been no reports that maternal age influence the FF, but FF does decrease with increasing maternal BMI (Hudecova et al., 2014, Wang et al., 2013, Zhou et al., 2015). It has also been suggested that some trisomies influence the FF. For example, trisomies 13 and 18 decreases (Suzumori et al., 2016) FF but trisomy 21 increases FF (Ashoor et al., 2013).

1.6 Aims

This thesis will explore the genetics of miscarriage over three chapters.

Firstly, a clinical trial to identify if there is a genetic cause of pregnancy loss using cfDNA from maternal blood rather than the standard procedure of completing cytogenetic analysis on pregnancy tissue. The aims of this chapter are to assess whether cfDNA in early pregnancy can reliably detect fetal aneuploidies and other chromosome abnormalities and whether the results obtained are comparable to the results obtained using routine testing on pregnancy tissue when there is a sample available. This chapter will also assess the levels of cfDNA in early pregnancy losses and whether the levels of cfDNA correlate to the levels of βhCG.

Secondly, a published systematic review to identify potential genetic causes in miscarriage. It is known that approximately 50 % of pregnancy losses are caused by genetic abnormalities and the remainder have apparently euploid karyotypes. However, it is plausible that there are cases of pregnancy loss with genetic aberrations that are not currently detected routinely. The systematic review identifies studies that have recorded monogenic genetic contributions to pregnancy loss.

Thirdly, whole exome sequencing was used to identify variants in genes which could be lethal during early development and could be contributing to recurrent miscarriage within the trial cohort.

Chapter 2: The use of cfDNA in the investigation of miscarriage

Chapter 2- The use of cfDNA in the investigation of miscarriage

Work from this Chapter has been published (Appendix 1) in Journal of Clinical Medicine under the following reference:

> Colley E, Devall AJ, Williams H, Hamilton S, Smith P, Morgan NV, Quenby S, Coomarasamy A, Allen S. Cell-Free DNA in the Investigation of Miscarriage. J Clin Med. 2020 Oct 26;9(11):3428. doi: 10.3390/jcm9113428. PMID: 33114508.

Authors roles: E.C.—Data analysis and writing; A.J.D.—Study design and ethical approval; H.W.—Study design, ethical approval, grant administration and editing; S.H.—Data analysis and critical appraisal; P.S.—Study design and critical appraisal; N.V.M.—Supervision and critical appraisal; S.Q.—Critical appraisal; A.C.—Study design, supervision and critical appraisal; S.A.—Study design, supervision, critical appraisal, data analysis, writing and editing.

2.1 Introduction

2.1.1 Miscarriage Genetics

Early pregnancy loss prior to 20 weeks gestation is the most common complication during pregnancy. One in five early pregnancies end in miscarriage and 50 % of these spontaneous pregnancy losses are due to chromosomal abnormalities. It is important to identify whether a chromosomal abnormality was causative of the pregnancy loss because this may have an indication for the prognosis of future pregnancies. If a sporadic chromosomal abnormality is

the cause of the pregnancy loss, the prognosis for future pregnancies is better than if the chromosome complement is normal in which case there may be another, non-chromosomal, reason for the miscarriage. If there is an unbalanced rearrangement in the pregnancy loss, it could mean that one of the parents carries a balanced chromosomal rearrangement. This would mean that future pregnancies would be susceptible to the same or other unbalanced rearrangement. In these cases, it is important to obtain blood samples for parental karyotyping for assessment of recurrence risk.

2.1.2 Cytogenetic analysis

The current Royal College of Obstetricians & Gynaecologists (RCOG) Green-top Guidelines No. 17 (Regan et al., 2011) recommends cytogenetic analysis of pregnancy tissue after the third and subsequent consecutive pregnancy losses, or karyotyping of parental samples if pregnancy tissue is not available. Traditionally, cell culture and G banded chromosome analysis were used to detect abnormalities in the pregnancy tissue. However, there is often a high failure rate, due to the poor quality of tissue received and the difficulty with culturing cells from such tissues and a limited resolution in detecting micro- deletion and duplication syndromes. Molecular-based approaches, such as quantitative fluorescent PCR (QF-PCR), subtelomere multiplex ligation-dependent probe amplification (MLPA) and array-comparative genomic hybridisation (array CGH) have therefore been implemented across laboratories. In the West Midland's Regional Genetics Laboratory, chromosomal abnormalities in pregnancy tissues are detected using QF-PCR and CMA based methods (McClelland et al., 2011). This increased the success of POC testing from 70-80 % to 92 %.

Currently, genetic testing for miscarriage is completed on pregnancy tissue, which comprises of fetal skin, umbilical cord and placental chorionic villi, referred to as products of conception (POC). This tissue needs to be fresh and unfixed so that the fetal tissues can be identified and have DNA extraction or cell culture performed. This comes with the risk of potential maternal cell contamination (MCC) which could lead to misdiagnosis of the sample. The POC samples contain maternal cells including red blood cells and decidua which are intertwined with the fetal tissues. These maternal cells can be carried over during the selection of fetal tissues resulting in maternal DNA during DNA extraction or an overgrowth of maternal cells during cell culture.

One of the other major problems with genetic testing of POC is that it is not always available after a pregnancy loss, or the tissue is not returned by the patient to allow for genetic testing in a laboratory setting. The management of miscarriage can be conservative, medical or surgical. Conservative management is when the miscarriage occurs naturally without any intervention and medical management involves taking medications which cause the cervix to open and the uterus to contract causing the pregnancy tissue to pass. During conservative and medical management the pregnancy tissue is often passed at home. Surgical management involves removing the pregnancy tissue using a suction catheter under general anaesthetic. This is completed in a hospital setting but comes with some risks such as bleeding, infection, perforation, cervical trauma and intrauterine adhesions. The collection of tissue can be very stressful for the patient and difficult to obtain. For example, collectable POC is not always present in the cases of very early pregnancy loss or the POC is lost down the toilet. POC is often fixed in formaldehyde for histological examinations and therefore rendered unusable for genetic testing as formaldehyde fixatives degrade the DNA. For all

these reasons, another easier, less stressful and reliable way of obtaining the fetal genetic material needs to be addressed.

2.1.3 cfDNA

The identification of cell free DNA (cfDNA) by Dennis Lo (Lo et al., 1997) and major advances in genomics has allowed the use of cfDNA to rapidly evolve. cfDNA is small fragments of placental DNA, derived from trophoblast cells of the placenta which can be obtained from the plasma of pregnant women. cfDNA has a half-life of just 16 minutes, meaning cfDNA is cleared from plasma quickly following delivery and ensures that there is no residual cfDNA from previous pregnancies. This has led to the development of non-invasive prenatal testing (NIPT), non-invasive prenatal diagnosis (NIPD), fetal sex determination and fetal rhesus determination. This means invasive tests such as chorionic villus sampling (CVS) and amniocentesis are increasingly being replaced by non-invasive tests which don't carry the ~1% miscarriage risk associated with invasive tests (Daley et al., 2014, Salomon and Sotiriadis, 2019). However, very little work has been completed using cfDNA for miscarriage genetic testing.

As discussed in the introduction, cfDNA is detectable as early as 22 days of gestation in singleton pregnancies and is clinically utilisable from seven weeks gestation. Different applications of cfDNA testing include fetal sex determination, Rhesus D (RhD) status, NIPT and NIPD.

2.1.4 cfDNA in miscarriages

Although cell-free fetal DNA is being utilized for prenatal diagnostic use in a clinical setting, very little work has been done in nonviable pregnancies. Only two (Yaron et al., 2020, Clark-Ganheart et al., 2015) to date have evaluated the use of cfDNA in a miscarriage setting. cfDNA would be extremely useful in miscarriages to examine for any genetic causes of the miscarriage by a simple blood test upon diagnosis. This would be far less stressful for the patient than collecting the POC and would ensure a sample for genetic testing is always available. POC is not always available because it is not present, not returned through the correct pathway or is fixed too soon in formaldehyde.

One prospective cohort study (Clark-Ganheart et al., 2015), analysed cfDNA from 50 patients who had either a missed abortion or fetal demise where the fetal tissue/ placenta was still *in situ.* Cytogenetic analysis, from CVS, amniocentesis or post-mortem tissue, was available in 18/50 (36 %) of pregnancies. The mean gestational age by ultrasound scan was 15.5 weeks (6.1–38.4 weeks) and the median fetal fraction was evaluated according to ultrasound scan findings in Table 3.

Table 3- Median fetal fractions according to ultrasound scan findings (Clark-Ganheart et al., 2015)

Ultrasound scan findings	Fetal Fraction (%)
Gestational sac only	4.7
Yolk sac without fetal pole	2.6
Fetal pole present	4.0
Second trimester conceptus	11.2
Third trimester conceptus	25.7

There was 14/16 (87.5 %) concordant results between cytogenetic analysis and cfDNA analysis, including trisomy 18, trisomy 13 and two trisomy 21. In two cases, the cfDNA result was not concordant with the cytogenetic result. The cytogenetic testing identified a monosomy X and a 1.79 MB microdeletion on chromosome X, cfDNA testing identified these results as euploid and trisomy 7 respectively. Although this was not confirmed, these discrepant results were suggested to be due to confined placental mosaicism.

In total, Clark-Ganheart et al. demonstrated 38/ 50 samples yielded results from cfDNA analysis and had fetal fractions between 3.7 and 65 %. Fetal fractions <3.7 % or >65% were considered to be non-results. Results were obtained in 87.9 % of samples with an ultrasound scan gestational age of eight weeks or greater and 52.9 % of samples with an ultrasound scan gestational age of less than eight weeks. This indicates that after miscarriage, providing the pregnancy tissue is still *in situ*, cfDNA remains in the maternal plasma.

An additional prospective diagnostic test study, (Yaron et al., 2020) has used cfDNA to analyse 109 patients with early pregnancy loss at less than 14 weeks by ultrasound scan. Ultrasound scans were performed to evaluate the gestation and a blood sample for βhCG and cfDNA. Chorionic villus samples (CVS) were taken for each patient for cytogenetic analysis by QF-PCR and karyotyping. Complete non-mosaic cytogenetic results were available for 86 patients (80.7 %), consisting of 55 pregnancies (64.0 %) with genetic aberrations and 31 euploid pregnancies. There was a mean ultrasound gestational age of 9.6 ± 1.9 week and the mean fetal fraction was 5.2% ± 3.6% (range <1% to 28%).

2.1.5 Tommy's National Centre for Miscarriage Research

The Tommy's National Centre for miscarriage research opened in 2016 and is based in Birmingham Women's Hospital, University Hospital Coventry, Queen Charlotte's & Chelsea Hospital and St Mary's Hospital, London. Tommy's miscarriage centre aims to find causes for pregnancy loss, gives parents to-be answers for their pregnancy loss and to stop it from happening again (https://www.tommys.org/our-organisation/our-research/research-miscarriage).

The four main research focuses in Tommy's are:

- Why did it happen?
- Will it happen again?
- Can we prevent it?
- How can we care for women emotionally?

This chapter of my thesis covers research which is one of thirteen work packages from Tommy's miscarriage centres and focuses on the genetic cause of 'why did it happen?' Within Tommy's we aim to reduce the number of couples who are given no reason or explanation for their miscarriage.

2.2 Aims

The aims of this chapter are to investigate the use of cfDNA from maternal blood sample as a diagnostic test following pregnancy loss, in particular:

- What the levels of cfDNA are in early miscarriages and whether it can be used to reliably detect fetal aneuploidies and other chromosome abnormalities.
- To address whether the results obtained using cfDNA are comparable to the results achieved using routine testing of products of conception when a sample is available.
- Whether the levels of cfDNA correlate with the levels of beta human chorionic gonadotrophin (βhCG).

This research was completed at Birmingham Women's Hospital. Patients were recruited through the Tommy's recurrent miscarriage clinic and laboratory tests were completed in West Midlands Regional Genetics Laboratory.

West Midlands Regional Genetics Laboratory (WMRGL) currently uses Illumina's Verifi technology for non-invasive Prenatal Testing (NIPT). Verifi uses WGS and identifies pregnancies with a high chance of trisomy 13 18 or 21 and is offered to patients with a high first trimester screening risk (>1 in 150) of Edward's, Patau or Down's syndrome following biochemical screening. As Verifi was already well established at WMRGL and uses WGS (Togneri et al., 2019), it was decided to trial this method for cfDNA analysis in pregnancy loss for all chromosomes. The closed pipeline using Verifi only identifies trisomy of chromosomes 13, 18 and 21. Therefore an additional bioinformatics pipeline was needed for the sequencing data to determine genetic aberrations in all chromosomes.

There are multiple bioinformatics pipelines which have been developed for NIPT. These include WIthin-SamplE Copy Number aberration DetectOR (WISECONDOR) (Straver et al., 2014b), Reliable Accurate Prenatal non-Invasive Diagnosis R package (RapidR) (Lo et al., 2014), Illumina bioinformatics pipelines and Tree-based consistency objective function for alignment evaluation (T-COFFEE) (Taly et al., 2011).

WISECONDOR is freely available and was developed in the Netherlands for NIPT. Using low coverage NGS, WISECONDOR detects small aberrations in the fetal DNA (Straver et al., 2014a). WISECONDOR works by splitting the genome into fixed sizes, known as bins. The frequency of each read is mapped to each bin and normalized for the GC-content. The reads are then compared across bins within the same sample and against a set of reference samples. We decided to evaluate WISECONDOR as the pipeline was published and freely available, the pipeline was validated using a HiSeq2500 which was also available in West Midlands Regional Genetics Laboratory and was based on low coverage NGS with 10 to 14 million single end reads which was similar to the read depth to that achieved by routine NIPT at West Midlands Regional Genetics Laboratory.

In addition, after the trial had commenced an opportunity to collaborate with Illumina arose. Samples which had consent to be sent to an external laboratory also had cfDNA testing using Illumina's VeriSeq technology and analysis using Illumina's all chromosome pipeline. This additional laboratory work and analysis was completed by Illumina Cambridge and final results were returned.

2.3 Methods

2.3.1 Ethics and approvals

This study was completed at Tommy's National Centre for miscarriage Research, IRAS project ID, 215646, has received Research Ethics Approval (REC) approval and Health Research Authority (HRA) approval.

2.3.2 Recruitment and Sample collection

2.3.2.1 Patient Consent

Informed consent was obtained from the donors in order to use their samples and data in this research. This explicitly covers consent to work with the patient's POC and genetic material. The samples were collected as donations to medical research and the tissue(s) were used according to the approved ethics and in accordance with the Human Tissue Act (HTA). The donor(s) maintained their ability to withdraw consent for further use but did not retain any rights to the samples after acquisition.

Consent was obtained in a clinical setting by someone trained in accordance with Good Clinical Practice (GCP) and on the trial delegation log. The consent forms explained the trial and how the samples would be used. A copy of the signed consent forms was retained for the sample collection log, a copy retained by the patient and a copy filed in the patient's medical notes.

Consent was either obtained through Tommy's Global (TG), Tommy's Pregnant (TP) or Tommy's Sporadic (TS). TG consent was obtained outside of the pregnancy during the Tommy's Recurrent Miscarriage clinics. TP consent was only used when the couple was

currently pregnant and had not been previously consented. The couple were consented if they were high risk of pregnancy loss or experiencing a miscarriage. The couple must have been at least 16 years of age and have a medical history of recurrent miscarriage and be experiencing early pregnancy complications. The couples were excluded from the study if either of the parents were not genetic parent(s) of the conceptus. Figure 8 below shows the consent flow. In addition, TS was introduced to collect samples from couples without recurrent miscarriage when experience their first or second pregnancy loss.

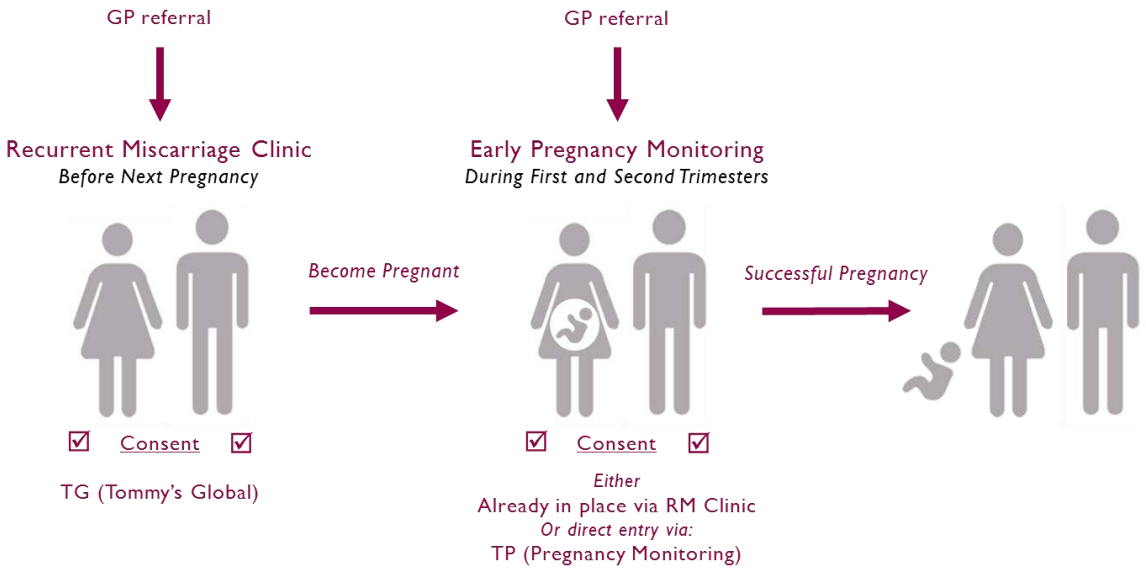


Figure 8- Consent was either obtained during the recurrent miscarriage clinic or during early pregnancy monitoring.

2.3.2.2 Samples

The patient recruitment was loosely divided into three groups, serial blood samples, threatened miscarriage and diagnosed miscarriage. There was some fluidity between these groups as patients could move between groups when they went from threatened to

confirmed miscarriage. The pregnancy tissue was collected where available when the pregnancy was no longer viable.

Serial blood samples were taken from women who had a history of three or more miscarriages and were being monitored during early pregnancy. A blood sample for cfDNA testing and βhCG was taken every week after a positive pregnancy test for six weeks or until the pregnancy ended. The serial samples were started as early as five weeks of gestation from LMP and the gestation confirmed by ultrasound scan. If unfortunately the women miscarried during the serial sampling monitoring, the unfixed POC was collected and analysed.

Threatened miscarriage samples were taken to ascertain whether cfDNA could be used as an early predictor of miscarriage. Knowing the genetics of the threatened pregnancy could determine the future clinical management of the pregnancy (e.g. prescribing progestogen to 'rescue' threatened euploid pregnancies). These women had a history of recurrent miscarriage and experiencing bleeding or pain during a viable pregnancy. A blood sample was taken for cfDNA analysis and βhCG levels. In cases where the patient went on to have a miscarriage, an additional cfDNA and βhCG blood sample was collected and the pregnancy tissues collected for cytogenetic testing, where available.

Diagnosed miscarriage samples were taken to ascertain whether cfDNA was comparable with routine POC testing. Women were at the point of miscarriage diagnosis, but still had some or all of the pregnancy tissue *in situ.* Blood samples were taken for cfDNA and βhCG levels. Where available the POC was obtained for cytogenetic testing.

Blood was collected for βhCG levels at the same time as cfDNA blood collection. This was to determine whether the levels of cfDNA correlate with βhCG levels.

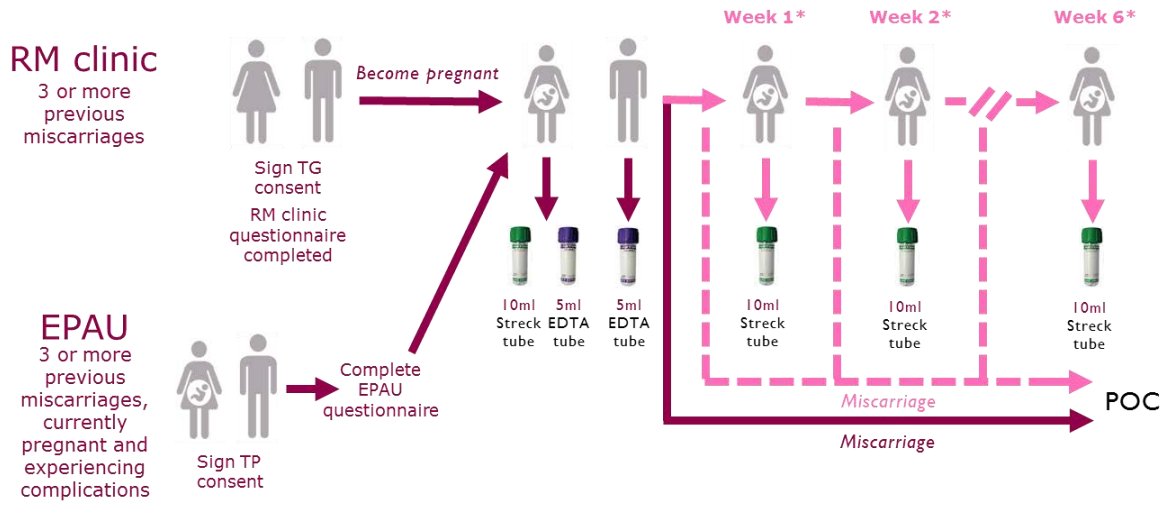


Figure 9- Samples are obtained following threatened or diagnosed miscarriage

2.3.2.3 Clinical Information

Upon recruitment to the trial, a patient questionnaire was completed by both male and female partners. This collected the clinical information which was relevant to the study including a history of previous pregnancies and/ or miscarriages, the gestation of the miscarriage(s), contraceptive and fertility treatments and whether the relationship was between blood relatives. Access to the patient's notes was also available to collect any additional clinical information.

During the sample collection scan findings were collected, where available. These include date of ultrasound scan and estimated gestation according to ultrasound scan, whether gestational sac, yolk sac and fetal pole were seen, crown-rump length measurements and fetal heart activity.

2.3.2.4 Anonymisation

Samples were anonymised using Tommy's identifiers which comprises of the consent number, sample type and a randomly assigned colour, e.g. 0001.MAT.Red. Streck samples for threated, diagnosed miscarriage and serial samples were anonymised at the time of collection. POC samples were assigned an anonymised Tommy's identifier and the data analysed anonymously for the study. However, they were still linked to patient identifiers so a report could be issued.

2.3.3 Sample preparation

2.3.3.1 Plasma isolation

Up to 10 ml of maternal blood was collected for cfDNA in Streck tubes. These are specialised glass blood tubes which keep cell free DNA stable in blood for up to seven days at room temperature. The plasma was isolated from whole blood using double centrifugation according to manufacturer's instructions and transferred into Eppendorf DNA LoBind tube in 1 ml aliquots. These aliquots were stored at -80^{o}C until use. This separated the plasma from red blood cells and buffy coat and removed any white blood cells which could lyse during the freeze thaw and release additional DNA into the sample.

2.3.3.2 βhCG

Blood samples to collect βhCG levels were ta ken at the same time as a Streck sample and allowed comparison of cfDNA levels with the βhCG level. These blood samples were collected in lithium heparin blood tubes and labelled with the patient identifiers. The βhCG

levels were tested by Birmingham Women's and Children's NHS Trust Biochemistry service and the results obtained using Birmingham Women's and Children's NHS Trust Telepath results service. The βhCG range was measured from between 0- 211134 mIU/mL.

2.3.3.3 POC

In cases where the pregnancy loss had been confirmed, where available, the products of conception (POC) were collected using patient identifiers. These were collected as routine clinical samples and QF-PCR and chromosomal microarray analysis (CMA) was completed on POC after the third and subsequent consecutive miscarriages(s) according to the RCOG Green-top Guideline No. 17 (Regan et al., 2011). QFPCR and CMA testing was completed at West Midlands Regional Genetics Laboratory.

Initially fetal tissue (fetal skin > cord > chorionic villi) was identified from the POC sample under a microscope using F-10 Nutrient Mixture (Ham's) Media (Gibco) and any maternal tissue was identified and removed. A small amount of fetal tissue was sampled into a Precellys Lysing Kit tube (Bertin) with Dulbecco's phosphate buffered saline (Gibco). DNA from the tissue was extracted using Precellys 24 Lyser/Homogeniser followed by EZ1 automated extraction according to West Midlands Regional Genetics Laboratory guidelines.

A quantitative Fluorescence-Polymerase Chain Reaction (QF-PCR), trisomy screen was first performed on DNA from the tissue extraction. This is a rapid screen to test for trisomies 13, 18 or 21, triploidy and sex chromosome aneuploidy. If the QF-PCR was abnormal it was reported, if this was normal Microarray testing was initiated. Microarray testing was completed using OGT CytoSure 8x60k Constitutional v3 design; exon/gene level resolution of

~500 DDD/ClinGen curated developmental genes and syndromic regions; tiered backbone resolution ~120-500kb; analysis in build GRCh37 using CytoSure v4.9 and CBS algorithm. The microarray analysis was expected to detect copy number imbalances >1 Mb, however would not have been able to detect balanced rearrangements or UPD and is limited in detecting mosaicism

2.3.3.4 cfDNA extraction

cfDNA was extracted from the 900 µl plasma samples using QIAamp DNA Blood Mini Kit 250 according to the standard protocol. The QIAvac 24 Plus Vacuum Manifold System was assembled according to Figure 10 and the cfDNA extracted from the plasma lysate using the QIAamp MiniElute Column. The DNA elute was transferred to 0.5 ml screw cap elution tubes and frozen to -80^{o}C until used for library preparation.

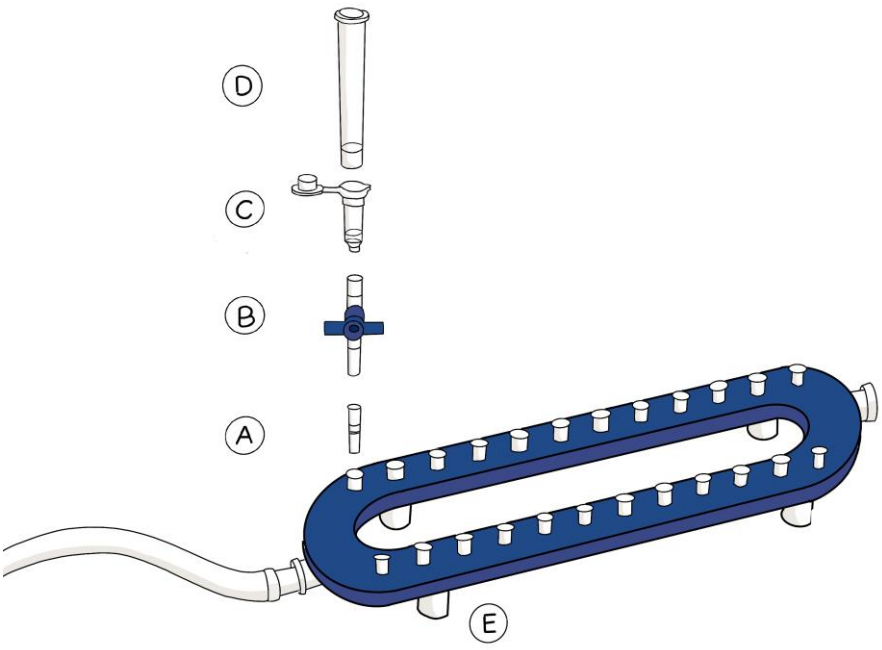


Figure 10- QIAvac 24 Plus Vacuum Manifold System

A, vacuum connector; B, vacuum valve; C, QIAamp mini column; D, extension tube; QIAvac 24 vacuum manifold. Illustration provided by Rosanna Colley.

2.3.4- SE-HiSeq-WC

All samples were analysed within West Midlands Regional Genetics Laboratory using single end sequencing reads on the HiSeq2500 and analysed using WISECONDOR (SE-HiSeq-WC). Laboratory work was completed equally by Mohammed Omer and myself. The extracted DNA was enriched into a 16 sample library, including a normal control, and a no DNA template control (Dulbecco's Phosphate-Buffered Saline), using the Illumina® TruSeq® Nano DNA sample prep kits (low-throughput), box A and box B, and additional Enhanced PCR Mix (EPM) tubes (Illumina) according to manufacturer's guidelines.

In brief, the sample library was made by converting each DNA fragment into a blunt end using the end repair mix and the 3' ends adenylated. A unique DNA adapter was ligated to each sample and the products cleaned up using PCR Sample Purification beads. This method was PCR based and the samples underwent PCR amplification using a BioRad thermal cycler and the final products cleaned up again using PCR Sample Purification beads.

The final libraries were quantified using Qubit dsDNA HS Assay Kit according to the manufacturer's guidelines. The concentration of each sample is calculated using the following calculation:

[Qubit reading ng/ml] x [40 dilution factor] x [1000 ml/L] x [bp mol/660g] x [1/290 bp] = nM

The library was normalised to a concentration of 1.7 nM and the DNA denatured using equal volumes of 0.1 M NaOH. A further diluted 8.5 pM library was shallow sequenced on the HiSeq2500 using 50 cycle rapid SBS kit (boxes 1 and 2) (Illumina) and Rapid SR cluster kit (Illumina), rapid run flow cell (Illumina), HiSeq Rapid SBS kit v2 (Illumina) and HiSeq Rapid SR

Cluster kit v2 (Illumina), according to standard manufacturer's guidelines. Run parameters were selected on the HiSeq500 to use single read indexing and 36 cycles.

Once the HiSeq2500 run had finished, the shallow sequencing data was run on WISECONDOR. The fetal fraction for the SE-HiSeq-WC method was calculated using the Illumina VeriFi pipeline.

2.3.4.1 The WIthin-SamplE Copy Number aberration DetectOR (WISECONDOR)

The WIthin-SamplE Copy Number aberration DetectOR (WISECONDOR) (Straver et al., 2014a) algorithm was downloaded from https://github.com/VUmcCGO/wisecondor and implemented as recommended, using a custom pipeline written in Python using the Ruffus framework. The custom pipeline works by removing low quality data and removing adapters added during the library preparation. The trimmed data files are aligned to the human genome (hg19 build) using the BWA *aln* algorithm to generate an intermediate SAI (Suffix Array Indexes) file. Reads from both lanes of the HiSeq were aligned to the reference genome independently, as this is more conveniently and computationally efficient, and the SAI file was converted to a SAM (Sequence Alignment and Mapping) file. The files were then converted into a BAM file (a compressed binary equivalent of the SAM file) with the lane identifier within the read group of each BAM, sorted into chromosome order and the two lanes merged together. The merged BAM files were run through WISECONDOR *convert* tool to create a WISECONDOR-specific (NPZ) file that records the number of reads in 50,000 base bins across the genome. The WISECONDOR *test* tool then compared reads across bins within each sample and to a previously generated set of reference samples that are from

aneuploid-negative samples. A binary output file (in NPZ format) is produced, converted into a readable .txt file and plotted along with G banding information. The WISECONDOR reference set was generated following authors instructions (Straver et al., 2014a) on samples previously assessed in house using Illumina's Verifi NIPT assay. An equal male to female ratio of samples were chosen that did not contain values, on any chromosome, with a normalised count value (NCV) of greater than four; which is the agreed threshold for calling a trisomy.

Results were scored from the WISECONDOR plots. 'Called regions' in an 'un-callable region' were considered to be artefacts. Sample results where ≥3 called chromosome regions were considered failed results. Low fetal fraction or low βhCG results were not used as cut-offs, to fail a result. However fetal fraction and βhCG results are taken into account in the results section (2.4). If there were no regions called as under- or over-represented in the sample then these pregnancies were considered to be euploid.

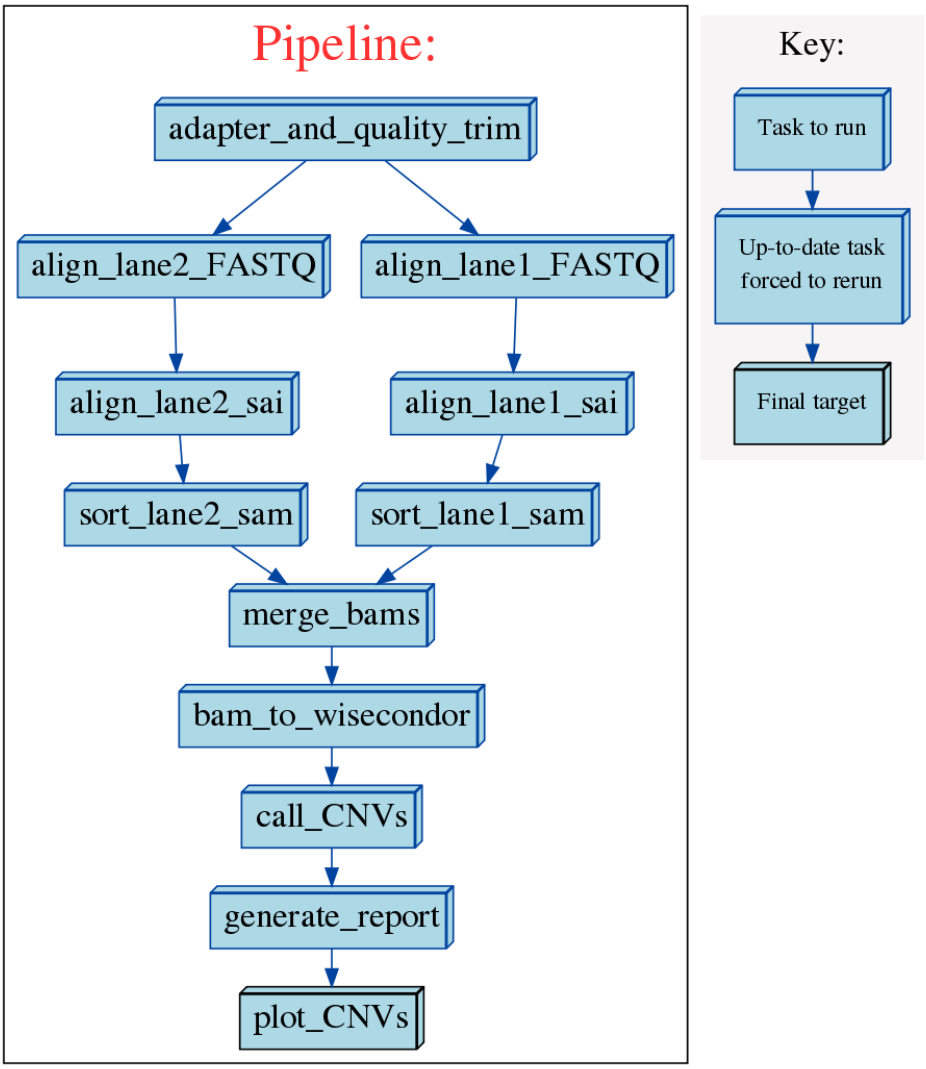


Figure 11- WIthin-SamplE COpy Number aberration DetectOR (WISECONDOR) Bioinformatics Pipeline.

2.3.5 PE-NextSeq-Illumina

During the trial the opportunity arose to collaborate with Illumina. At least 1 ml of plasma sample, from patients which had been consented to be tested by an external laboratory were sent to an Illumina Laboratory in Cambridge. These samples were run using Illumina's automated VeriSeq NIPT v2 CE-IVD using paired end sequencing reads on the NextSeq and analysed using Illumina's all chromosome pipeline (PE-NextSeq-Illumina).

The samples were processed in a 24-sample batch through a modified Illumina VeriSeq NIPT solution v2 workflow as previously described (Pertile et al., 2020, Andrews et al., 2020), but using the latest analysis platform (Leinfinger et al., 2020) and with a few small modifications. This protocol is PCR free and whole genome sequencing is used to detect aneuploidies, partial duplications and deletions, sex chromosome aneuploidy and calculate fetal fraction. Table 4 below illustrates the two sequencing and analysis protocols used in the study.

Table 4- Sequencing and analysis protocols used in the study.

SE-HiSeq-WC	PE-NextSeq-Illumina
Manual	Automated
PCR based	PCR free
Single end reads	Paired end reads
HiSeq	NextSeq
WISECONDOR	Illumina custom pipeline

2.4 Results

2.4.1 PAGE Results

To investigate the limit of detection using SE-HiSEq-WC, an initial analysis was carried out on plasma aliquots collected from pregnancies recruited to the Prenatal Assessment of Genomes and Exomes (PAGE) study (Lord et al., 2019). These were pregnancies with structural abnormalities where genetic aberrations had been detected by CMA analysis. Three PAGE samples were selected with five varying sizes of deletions and duplications to assess the detection capabilities of WISECONDOR. Table 5 shows the detection of aberrations in PAGE samples ranging in size from 2.6Mb to 71Mb.

Table 5- Detection of aberration in PAGE samples by WISECONDOR.

Gestation	Scan findings	Fetal sex	Aberrations	Deletion/ Duplication size	Chromosome	WISECONDOR detection
Not recorded	Hypoplastic left heart syndrome, short long bones, bright bowel	Male	22q11.21 micro-duplication	2.6 Mb duplication	22	Not detected
22+6	Encephalocele	Female	46,XX,der(13)t(3:13)(q28;q22.1)	8.3 Mb duplication	3	Not detected
				41.2 Mb deletion	13	Detected
13+1	Nuchal pad/ nuchal thickening (6.7 mm)	Female	46,XX,der(6)t(3;6)(q21;q25)mat	71 Mb duplication	3	Detected
				10.5 Mb deletion	6	Detected

The 22q11.21 micro-duplication and the 8.3 Mb duplication were not detectable by WISECONDOR, but the 41.2 Mb deletion, 71 Mb duplication and 10.5 Mb deletion were all detected. It was concluded from this initial analysis that in general aberrations under 10 Mb were unlikely to be detectable by WISECONDOR but aberrations above 10 Mb are more likely to be detectable. However this is dependent on several factors including fetal fraction and gestation. The plasmas from the PAGE were taken from pregnancies in the second trimester, whereas pregnancy losses <10 weeks are likely to have lower FFs. The CMA result and WISECONDOR MPS/WGS analysis for two PAGE cases are shown in Figure 12.

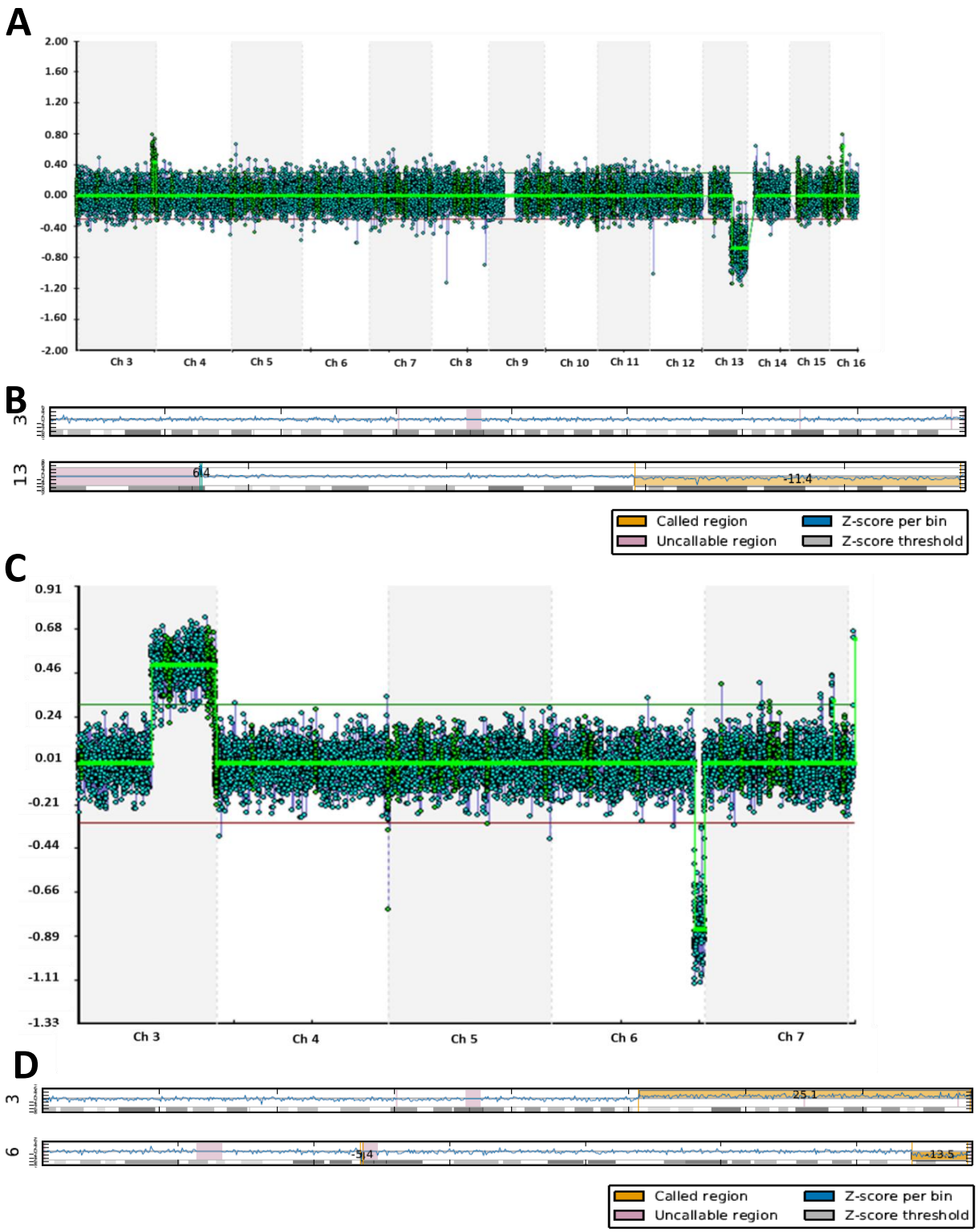


Figure 12- CMA and WISECONDOR analysis of PAGE Samples

A) CMA of POC showing a duplication of 8.3 Mb on chromosome 3 and a 42.2 Mb deletion on chromosome 13. 46,XX,der(13)t(3:13)(q28;q22.1) B) WISECONDOR WGS of the corresponding plasma sample showing a called region at the end of chromosome 13 but no called region on chromosome 3. C) CMA of POC of 46,XX,der(6)t(3;6)(q21;q25)mat, showing a duplication of 71 Mb on chromosome 3 and a 10.5Mb deletion on chromosome 6. D) WISECONDOR analysis of the corresponding plasma sample showing a called region at the end of chromosome 3 and 6.

2.4.2 Sample Collection

In total, 219 maternal blood samples were collected in Streck blood tubes for cfDNA analysis. This comprised of 72 serial samples, 30 threatened miscarriage samples and 123 confirmed miscarriage samples (of which 6 were from serial patients).

2.4.2.1 Serial Samples

A total of 72 serial samples were collected from 15 patients. A maternal blood sample was collected approximately every week and an early pregnancy ultrasound scan was performed on the patient to assess the gestational age and health of the conceptus. The gestations by ultrasound scan ranged from four to 13 weeks with a mean of 6 weeks, the average number of previous pregnancy losses was three and the average maternal age was 34 (Figure 13). Nine of these pregnancies resulted in live births and six of them miscarried (Table 6). cfDNA testing was completed with SE-HiSeq-WC. The βhCG measurement was calculated for 66 samples and the fetal fraction calculated using the NIPT VeriFi pipeline during the cfDNA testing.

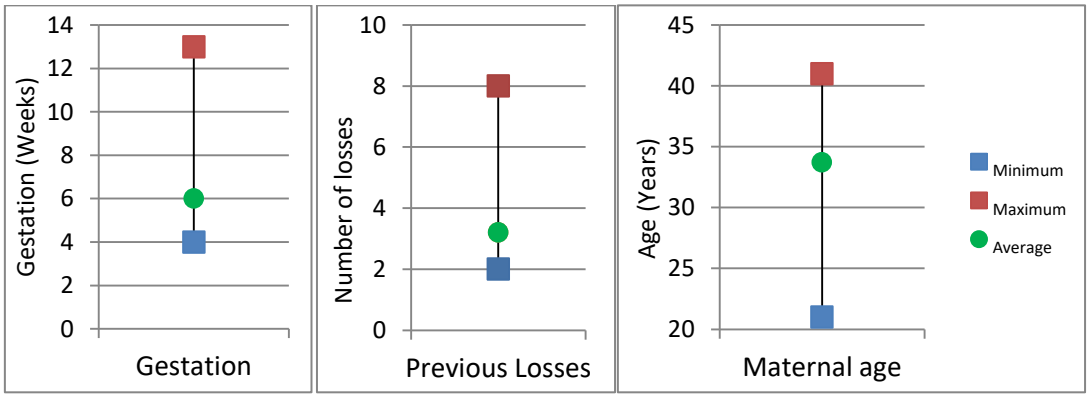


Figure 13- Baseline Characteristics for Serial Samples

Graph shows the maximum, minimum and average values for the gestation by ultrasound scan, number of previous pregnancy losses and maternal age of the serial sample cohort.

Table 6- Serial Sample Collection

The gestation of the pregnancy was calculated by ultrasound scan. The red samples represent when the miscarriage was diagnosed. The sample in purple represents an early fetal death with complete infraglottic laryngeal atresia complicated by severe pulmonary hypoplasia. a, b etc. represents when more than one pregnancy was sampled from the same patient.

Patient	Previous losses	Age	Sample 1 Gestation	Sample 2 Gestation	Sample 3 Gestation	Sample 4 Gestation	Sample 5 Gestation	Sample 6 Gestation
4	2	40	5+0	6+2	7+2			
34b	3	41	6+0	7+0	8+2	9+6	10+4	11+0
73	3	37	6+0	7+0	8+0	9+0	10+0	11+0
82	8	39	5+5	6+4	8+0			
91	2	35	5+5	6+4	7+0	9+0	10+2	11+5
96	3	33	5+3	5+0				
98	3	38	5+0	6+4	7+1	8+0	9+4	10+2
121	5	32	5+0	-	8+3	9+4	10+6	11+5
151	3	26	5+0	6+0	7+0	8+0	9+0	11+0
153	3	38	5+0	6+0	7+0	8+0	9+0	10+0
168	3	38	4+3	No scan	5+6	6+0	6+0	
170a	4	28	4+0	5+0	7+0	8+0		
198	2	31	6+0	6+2				
218	2	21	5+6	6+0	7+0	8+0	9+0	11+5
234	2	29	7+0	8+2	9+0	10+0	11+0	13+0

2.4.2.2 Threatened miscarriage Samples

In total 26 samples from 25 different patients were collected as “threatened miscarriage” samples where there was pain or bleeding during the pregnancy. The gestations ranged from 5+2 to 12+4 weeks, the average number of previous pregnancy losses was three and the average maternal age was 32. The pregnancies were followed up and 16 resulted in live birth and eight resulted in miscarriage. Four of the miscarriages (Sample IDs 27, 34a, 107 and 242) also had an additional sample taken at the point of miscarriage diagnosis, which is included in the confirmed miscarriage cohort. CfDNA testing was completed using SE-HiSeq-WC. The βhCG measurement was taken for the samples and the fetal fraction calculated using the Illumina VeriFi pipeline during the cfDNA testing.

2.4.2.3 Confirmed Miscarriage Samples

In total 123 blood samples were collected for confirmed miscarriages, 113 of these samples were collected only for the confirmed miscarriage cohort, 4 samples also had threatened miscarriage samples collected (sample IDs 27, 34a, 107 and 242) and 6 samples (sample IDs 4, 82, 96, 168, 170a and 198) were from the serial sample cohort at the point of diagnosis. The gestations ranged from 5+0 to 11+6 weeks, the average number of previous pregnancy losses was four and the average maternal age was 34. Seventy-nine (64.2 %) samples received cytogenetic results from the POC, of which 37 (46.8 %) had no significant copy number imbalance detected and 42 (53.2 %) had a genetic aberration.

2.4.3 Linear Regression

Linear regression analysis compares the direct relationship between two variables and was completed to examine the relationship between fetal fraction, gestation by scan and βhCG value. Statistical analysis was completed on gestation, βhCG and fetal fraction using linear regression calculated for:

- Serial samples
  - βhCG on gestation
  - Fetal fraction (VeriFi) on βhCG
  - Fetal fraction on gestation

  Calculated difference between live birth and miscarriage samples.
- Threatened miscarriage samples
  - βhCG on gestation.
  - Fetal fraction (Verifi) on βhCG

  Calculated difference between live birth and miscarriage samples.
- Confirmed miscarriage samples
  - βhCG on gestation
  - Fetal fraction (Verifi) on βhCG
  - Fetal fraction (VeriSeq) on βhCG

The combined correlation of fetal fraction on gestation for serial samples from pregnancies resulting in either live birth or pregnancy loss was calculated. This showed that for each additional week of gestation the fetal fraction increased by 1% in pregnancies that resulted in livebirth and by 0.2 % for pregnancies that resulted in miscarriage (P=0.09).

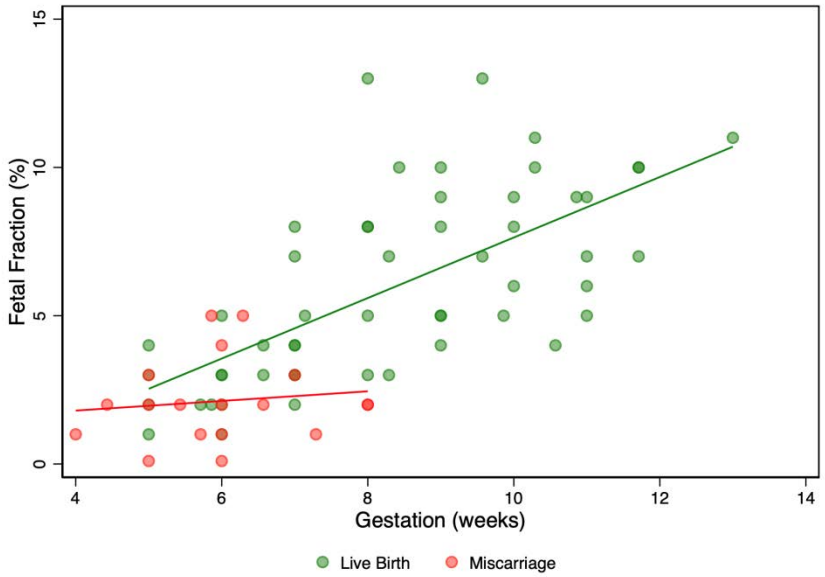


Figure 14- correlation of fetal fraction on gestation for serial samples from pregnancies resulting in either live birth or pregnancy loss

Linear regression for fetal fraction on βhCG analysis, for all groups, showed that there was a statistically significant correlation between fetal fraction per 10,000 mIU/mL rise of βhCG for all samples (Figure 15/Table 7). When the pregnancies resulting in livebirth were compared to the pregnancies resulting in miscarriage (from the serial and threatened cohorts combined), there was no observed difference in the correlation of fetal fraction with βhCG (p = 0.426)

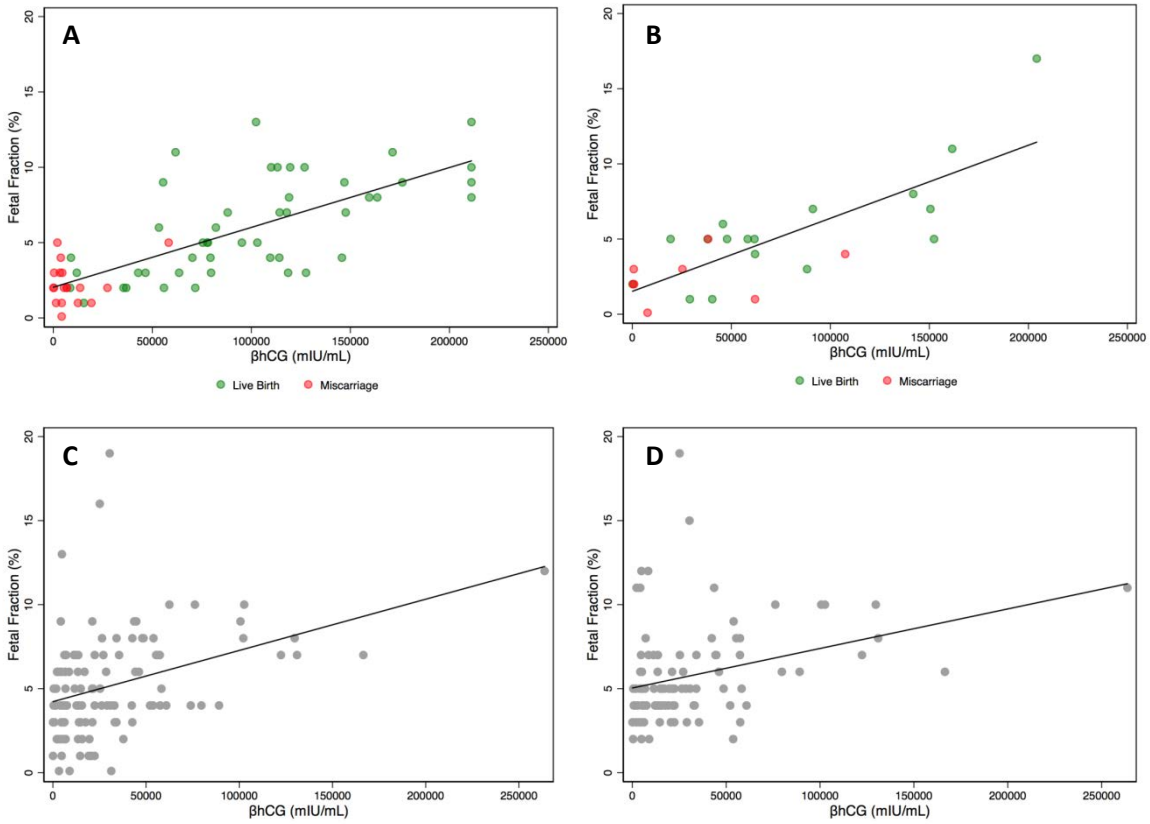


Figure 15- Linear regression graphs for fetal fraction on βhCG
A) Serial, B) Threatened, C) Confirmed (Verifi), D) Confirmed (VeriSeq).

Table 7- Linear Regression analysis for Fetal fraction on βhCG
Where coef. is the coefficient per 10,000 mIU/mL rise of βhCG, se is the standard error.

		coef.	(se)	p	R2 (%)
βhCG	**Serial**	0.40	(0.04)	<0.001	56.0
	Threatened	0.49	(0.08)	<0.001	63.0
	Confirmed (VeriFi)	0.30	(0.07)	<0.001	14.6
	Confirmed (VeriSeq)	0.24	(0.07)	0.001	10.6

The linear regression analysis showed that there was a statistically significant increase in βhCG for each additional week of gestation for pregnancies in the "serial" and "threatened" groups. As illustrated in the graphs, this was largely due to the contribution of pregnancies that went on to result in a live birth represented in green. The increase of βhCG for each additional week of gestation for the samples where a miscarriage had been confirmed was not statistically significant (Figure 15/ Table 8).

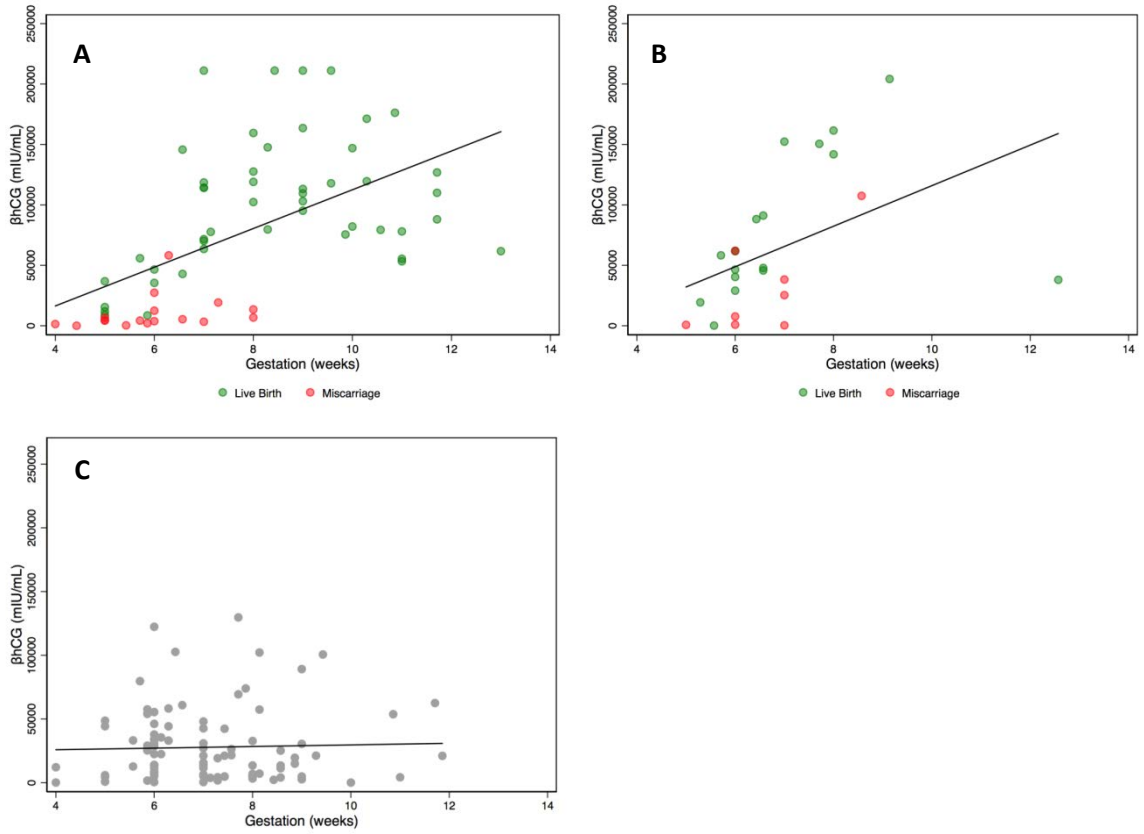


Figure 16- Linear regression graphs for βhCG on gestation
A) Serial, B) Threatened, C) Confirmed

Table 8- Linear Regression analysis for βhCG on gestation

Where coef. is the coefficient for each additional week of gestation, se is the standard error.

		coef.	(se)	P	R2 (%)
Gestation	**Serial**	16023.0	(3033.7)	<0.001	31.0
	Threatened	16772.9	(6709.4)	0.019	20.0
	Confirmed	632.6	(1961.6)	0.748	0.1

2.4.4 Assessment of test performance in miscarriage samples

In total, 123 samples were collected once a miscarriage had been confirmed. 80 pregnancies had a corresponding cytogenetic result from POC analysis. All 123 samples were analysed using SE-HiSeq-WC and 101 samples had additional analysis using PE-NextSeq-Illumina (Figure 17). The βhCG measurement was taken for 120 of the samples and the fetal fraction for both SE-HiSeq-WC and PE-NextSeq-Illumina was calculated during the cfDNA testing.

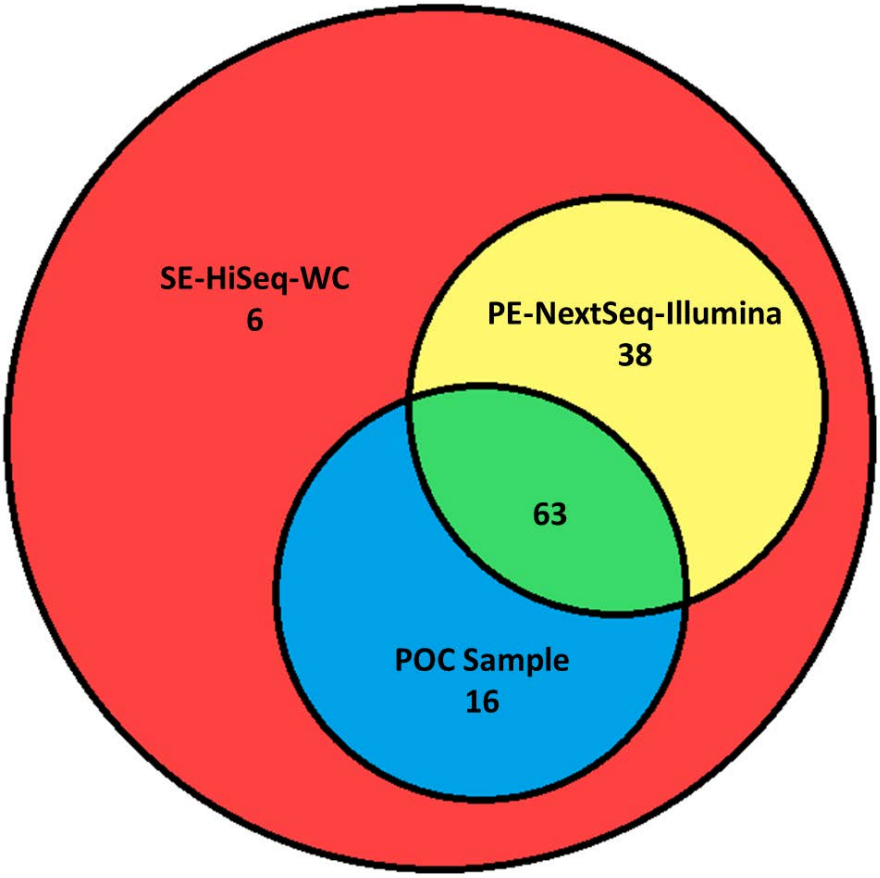


Figure 17- Venn diagram showing the total number of confirmed miscarriage samples with corresponding POC samples.

Six Samples only had analysis SE-HiSeq-WC, 16 samples had SE-HiSeq-WC with a known cytogenetic result, 38 samples had SE-HiSeq-WC analysis and PE-NextSeq-Illumina without a known cytogenetic result, 63 samples had SE-HiSeq-WC analysis and PE-NextSeq-Illumina with a known cytogenetic result.

The analysis from cfDNA was compared to the cytogenetic results from the POC samples (Table 9), where available. Using SE-HiSeq-WC, 57.5 % of samples were correctly identified in comparison to POC results and using PE-NextSeq-Illumina, 72.4% of samples were correctly identified. Triploid results cannot be identified by either pipeline, so have been excluded.

Table 9- All miscarriage samples analysed by cfDNA using SE-HiSeq-WC and PE-NextSeq-Illumina and compared to genetic analysis of POC (CMA).

Includes all samples from the miscarriage cohort and additional samples from the threatened and serial cohorts where the sample was taken at the point of miscarriage diagnosis. Note, samples 529-UHC, 264 and 433-UHC appear twice as they fall into two subgroups. No sample- sample was not submitted for analysis, no fetal material- sample was received but no fetally derived tissue could be identified (including chorionic villi), maternal cell contamination- sample received but QF-PCR results showed too much of the sample tested was maternally derived and therefore a genetic result was not available. The gestation of the pregnancy was calculated by ultrasound scan. *represents the confirmed miscarriage sample from serial patients. a, b etc. represents when more than one pregnancy was sampled from the same patient. ii represents where a previous threatened sample was taken. NR indicates not recorded results.

	Patient	POC results (Male/ Female)	POC tissue analysed	Gestation (Weeks)	Fetal fraction (%)	βhCG (mIU/mL)	No. of previous losses	Maternal age (Years)	SE-HiSeq-WC outcome	PE-NextSeq-Illumina outcome
Normal POC, both tests concordant	26	No significant copy number imbalance (M)	Placental villi	7+4	5	21572	6	29	No anomaly detected	No anomaly detected- XY
	88	No significant copy number imbalance (F)	Placental villi	10+0	3	69	4	42	No anomaly detected	No anomaly detected- XX
	107ii	No significant copy number imbalance (F)	Placental villi	5+0	8	48612	3	24	No anomaly detected	No anomaly detected- XX
	131	No significant copy number imbalance (F)	Placental villi	9+0	4	89142	2	30	No anomaly detected	No anomaly detected- XX
	155	No significant copy number imbalance (F)	Placental villi	7+5	6	129704	3	23	No anomaly detected	No anomaly detected- XX
	161	No significant copy number imbalance (M)	Fetal skin	9+0	19	30465	2	20	No anomaly detected	No anomaly detected- XY
	170bii	No significant copy number imbalance (F)	Fetal skin	8+4	7	11296	5	28	No anomaly detected	No anomaly detected- XX
	191	No significant copy number imbalance (M)	Placental villi	11+0	9	4229	4	40	No anomaly detected	No anomaly detected- XY
	215a	No significant copy number imbalance (M)	Placental villi	8+4	4	4067	3	35	No anomaly detected	No anomaly detected- XY
	217	No significant copy number imbalance (F)	Placental villi	8+4	16	25177	6	24	No anomaly detected	No anomaly detected- XX
	242ii	No significant copy number imbalance (F)	Placental villi	7+0	2	15670	2	25	No anomaly detected	No anomaly detected- XX

	263	No significant copy number imbalance (M)	Placental villi	6+0	7	122350	NR	29	No anomaly detected	No anomaly detected- XY
	308	No significant copy number imbalance (F)	Placental villi	7+3	2	4888	4	31	No anomaly detected	No anomaly detected- XX
	98-UHC	No significant copy number imbalance (F)	Placental villi	NR	10	76188	4	34	No anomaly detected	No anomaly detected- XX
	198-UHC	No significant copy number imbalance (F)	Placental villi	NR	9	43599	4	33	No anomaly detected	No anomaly detected- XX
	243-UHC	No significant copy number imbalance (M)	Placental villi	6+3	10	102631	2	28	No anomaly detected	No anomaly detected- XY
	306-UHC	No significant copy number imbalance (M)	Placental villi	5+6	4	57511	4	31	No anomaly detected	No anomaly detected- XY
	351-UHC	No significant copy number imbalance (F)	Placental villi	6+4	4	60847	4	31	No anomaly detected	No anomaly detected- XX
	420-UHC	No significant copy number imbalance (M)	Placental villi	NR	4	7526	14	34	No anomaly detected	No anomaly detected- XY
	544-UHC	No significant copy number imbalance (M)	Fetal skin	6+1	7	35524	5	30	No anomaly detected	No anomaly detected- XY
	620-UHC	No significant copy number imbalance (M)	Fetal skin	10+6	4	53747	5	31	No anomaly detected	No anomaly detected- XX
	669-UHC	No significant copy number imbalance (F)	Placental villi	NR	7	130993	7	39	No anomaly detected	No anomaly detected- XX
	785-UHC	No significant copy number imbalance (M)	Placental villi	NR	2	2098	3	32	No anomaly detected	No anomaly detected- XX
	796-UHC	No significant copy number imbalance (M)	Placental villi	5+5	4	79671	3	25	No anomaly detected	No anomaly detected- XY
Normal POC, SE-HiSeq-WC Concordant, PE-NextSeq-Illumina not tested	2	No significant copy number imbalance (M)	Fetal skin	7+0	2	6317	4	35	No anomaly detected	-
	57b	No significant copy number imbalance (M)	Placental villi	7+0	8	42635	9	43	No anomaly detected	-
	68	No significant copy number imbalance (M)	Fetal skin	11+5	10	62510	3	35	No anomaly detected	-
	82*	No significant copy number imbalance (M)	Placental villi	8+0	2	6844	8	39	No anomaly detected	-
	83	No significant copy number imbalance (M)	Placental villi	NR	7	55844	3	34	No anomaly detected	-
	144	No significant copy number imbalance (F)	Placental villi	7+0	8	48002	2	34	No anomaly detected	-

Normal POC, both tests discordant	81	No significant copy number imbalance (F)	Fetal skin	NR	6	4648	3	36	chromosome 5 deletion	del(10)(q26.13q26.3)- XX
Normal POC, SE-HiSeq-WC discordant, PE-NextSeq-Illumina not tested	35	No significant copy number imbalance (F)	Fetal skin	9+0	6	2693	2	36	Chromosome 19 insertion	-
Normal POC, SE-HiSeq-WC concordant, PE-NextSeq-Illumina discordant	27-UHC	No significant copy number imbalance (F)	Placental villi	NR	6	16968	4	41	No anomaly detected	Trisomy 16- XY
	850-UHC	No significant copy number imbalance (F)	Placental villi	6+0	6	46177	0	35	No anomaly detected	Trisomy 16- XX
Normal POC, SE-HiSeq-WC discordant, PE-NextSeq-Illumina concordant	67	No significant copy number imbalance (M)	Placental villi	6+0	6	8670	3	33	partial trisomy 9 + partial trisomy 15	No anomaly detected- XY
	141	No significant copy number imbalance (M)	Fetal skin	7+3	9	21160	6	32	Chromosome 7 duplication	No anomaly detected- XY
	229	No significant copy number imbalance (F)	Placental villi	NR	12	263766	3	27	Chromosome 1-deletion + Chromosome 10 deletion	No anomaly detected- XX
Abnormal POC, both tests concordant	133	~5.8 Mb terminal deletion at 7q36.2q36.3 and ~21.14 Mb terminal duplication at 19q13.12q13.43 (F)	Placental villi	NR	7	57348	4	34	Terminal Chromosome 19 duplication	dup(19)(q13.12q13.43)- XX
	164	Trisomy 15 (F)	Placental villi	6+0	4	6774	2	42	Trisomy 15	Trisomy 15- XX
	176	Trisomy 15 (F)	Placental villi	7+0	5	491	2	43	Trisomy 15	Trisomy 15- XX
	279	Trisomy 7 (F)	Placental villi	6+0	NR	8429	4	40	Trisomy 7	Trisomy 7- XX
	519-UHC	Trisomy 21 (M)	Placental villi	9+3	9	100638	6	28	Trisomy 21	Trisomy 21- XY
	529-UHC	Trisomy 21, monosomy X (F)	Placental villi	7+0	5	21171	5	39	Trisomy 21	Trisomy 21-XX
	541-UHC	Trisomy 15 (F)	Placental villi	7+3	4	42333	2	33	Trisomy 15	Trisomy 15- XX

Abnormal POC, SE-HiSeq-WC Concordant, PE-NextSeq-Illumina not tested	12	Trisomy 15 (M)	Placental villi	7+5	-	69345	3	38	Trisomy 15	-
	114	Trisomy 16 (F)	Placental villi	6+0	5	11719	2	36	Trisomy 16	-
	129	Trisomy 22 (M)	Placental villi	8+0	5	6764	5	43	Trisomy 22	-
Abnormal POC, both tests discordant	4*	Trisomy 22 (F)	Placental villi	7+2	1	19247	2	40	No anomaly detected	No anomaly detected- XX
	202	~9.23 Mb terminal duplication at 4q34.3q35.2 and ~29.78 Mb terminal deletion at 5q33.1q35.3 (F)	Placental villi	6+0	4	13819	4	29	No anomaly detected	No anomaly detected- XX
	245	Trisomy 22 (M)	Placental villi	NR	3	14762	3	33	No anomaly detected	No anomaly detected- XY
	303-UHC	Deletion 13q13.3q34 (F)	Placental villi	6+0	3	34087	3	35	No anomaly detected	No anomaly detected- XX
	328-UHC	Trisomy 11 (F)	Placental villi	NR	6	2983	3	42	No anomaly detected	No anomaly detected- XX
	586-UHC	Mosaic Trisomy 17 (F)	Placental villi	6+0	7	22435	2	41	No anomaly detected	No anomaly detected- XX
	816-UHC	Trisomy 18 (F)	Placental villi	5+0	4	5852	3	40	No anomaly detected	No anomaly detected- XX
	965-UHC	Trisomy 15 (M)	Placental villi	7+6	4	73962	3	42	No anomaly detected	No anomaly detected- XY
	99	Trisomy 5 (M)	Placental villi	6+0	3	5111	4	27	Chromosome 8 duplication	No anomaly detected- XY
	260	Trisomy 22 (M)	Placental villi	7+0	4	5194	2	43	Monosomy 10 +, Monosomy 20	No anomaly detected- XY
	228	Monosomy 21	Placental villi	7+0	3	6220	5	41	No anomaly detected	dup(15)(q21.3q23);dup(20)(q11.21q13.12)- XY
	264	Trisomy 13 (F)	Placental villi	6+0	3	14002	NR	33	No anomaly detected	Trisomy 13, Trisomy 16- XX
Abnormal POC, SE-HiSeq-WC discordant, PE-NextSeq-Illumina not tested	51a	Mosaic trisomy 4 (M)	Placental villi	5+4	4	12725	6	34	No anomaly detected	-
	57a	Mosaic trisomy 16 (F)	Placental villi	NR	3	42635	8	41	No anomaly detected	-
	62	Trisomy 4 (F)	Placental villi	6+0	<1	31316	3	36	No anomaly detected	-
	80	Trisomy 17 (F)	Placental villi	NR	4	26179	3	40	No anomaly detected	-
	459-UHC	Trisomy 13 (F)	Placental villi	7+3	1	4736	6	42	No anomaly detected	-

Abnormal POC, SE-HiSeq-WC discordant, PE-NextSeq-Illumina concordant	264	Trisomy 13 (F)	Placental villi	6+0	3	14002	NR	42/80	No anomaly detected	Trisomy 13, Trisomy 16- XX
	287	Trisomy 12 (F)	Placental villi	NR	4	NR	0	42	No anomaly detected	Trisomy 12- XX
	290	Trisomy 16 (F)	Placental villi	5+0	6	44313	0	39	No anomaly detected	Trisomy 16- Undetected
	319-UHC	duplication 7q22.1q36.3 (F)	Fetal skin	7+0	7	13642	6	40	No anomaly detected	dup(7)(q22.1q31.1)
	400-UHC	Trisomy 22 (M)	Placental villi	5+6	4	29052	5	42	No anomaly detected	Trisomy 22- XY
	462-UHC	Trisomy 15 (M)	Placental villi	6+1	4	22429	2	40	Chromosome 11 deletion	Trisomy 15- XY
Monosomy X, POC results	7	Monosomy X (F)		6+0	2	37819	5	35	No anomaly detected	-
	163	Monosomy X (F)		8+4	6	13466	3	24	No anomaly detected	Anomaly detected- XO
	175	Monosomy X (F)		8+0	4	5323	6	34	No anomaly detected	Anomaly detected- XO
	529-UHC	Trisomy 21, monosomy X (F)		7+0	5	21171	5	39	Trisomy 21	Trisomy 21-XX
No POC, both tests concordant- normal result	27ii	No fetal material		6+0	4	415	3	35	No anomaly detected	No anomaly detected- XX
	77	No fetal material		5+4	3	33187	5	37	No anomaly detected	No anomaly detected- XY
	160	No sample		6+0	6	55400	2	36	No anomaly detected	No anomaly detected- XX
	168*	No fetal material		6+0	<1	NR	3	38	No anomaly detected	No anomaly detected- XX
	170a*	No sample		8+0	2	13523	4	28	No anomaly detected	No anomaly detected- XY
	173	No fetal material		8+0	4	32727	5	30	No anomaly detected	No anomaly detected- XY
	198*	No fetal material		6+2	5	58253	2	31	No anomaly detected	No anomaly detected- XY
	215b	Maternal cell contamination		8+1	7	7117	5	36	No anomaly detected	No anomaly detected- XY
	278	No fetal material		7+0	7	27099	4	36	No anomaly detected	No anomaly detected- XY

	288	No fetal material		NR	7	6176	0	24	No anomaly detected	No anomaly detected- XX
	289	No sample		NR	1	22377	0	26	No anomaly detected	No anomaly detected- XX
	16-UHC	No sample		NR	4	52141	4	39	No anomaly detected	No anomaly detected- XX
	97-UHC	Maternal cell contamination		NR	9	44760	2	35	No anomaly detected	No anomaly detected- XX
	116-UHC	No fetal material		6+0	6	28585	2	37	No anomaly detected	No anomaly detected- XX
	430-UHC	No sample		NR	1	20565	4	40	No anomaly detected	No anomaly detected- XX
	432-UHC	No sample		7+4	8	26369	5	25	No anomaly detected	No anomaly detected- XY
	470-UHC	No sample		5+6	5	25336	7	36	No anomaly detected	No anomaly detected- XX
	531-UHC	No fetal material		7+2	6	4289	8	38	No anomaly detected	No anomaly detected- XX
	534-UHC	No sample		9+2	3	21194	4	25	No anomaly detected	No anomaly detected- XX
	564-UHC	No sample		8+6	2	19590	10	28	No anomaly detected	No anomaly detected- XX
	643-UHC	Maternal cell contamination		4+0	7	12128	2	39	No anomaly detected	No anomaly detected- XY
	765-UHC	No sample		5+0	3	789	5	23	No anomaly detected	No anomaly detected- XY
	709-UHC	No fetal material		8+6	5	14931	12	37	No anomaly detected	No anomaly detected- XX
	1022-UHC	No sample		8+3	6	2242	4	26	No anomaly detected	No anomaly detected- XY
	1058-UHC	No sample		6+2	4	33008	2	36	No anomaly detected	No anomaly detected- XX
No POC, both tests concordant-abnormal result	130	No fetal material		7+0	4	30865	9	29	Trisomy 16	Trisomy 16 - XX
	269	No fetal material		NR	NR	2262	NR	34	Trisomy 12	Trisomy 12- XX
	273	No sample		7+0	7	11530	2	33	Trisomy 16	Trisomy 16- XX
	340-UHC	No sample		5+6	8	53981	2	33	Trisomy 13	Trisomy 13- XO
	433-UHC	No fetal material		NR	3	17567	2	43	Trisomy 14	Trisomy 14, Trisomy 18- XX
	825-UHC	No sample		8+1	NR	57373	4	42	Trisomy 20	Trisomy 20- XX

No POC, SE-HiSeq-WC outcome, PE-NextSeq-Illumina not tested	16	No sample		4+0	1	88	4	28	No anomaly detected	-
	78	No fetal material		8+0	6	6649	2	26	No anomaly detected	-
	96*	No sample		5+0	3	4543	3	33	No anomaly detected	-
	111	No fetal material		NR	2	2809	4	43	No anomaly detected	-
	122	No sample		11+6	5	21038	2	22	No anomaly detected	-
	42	No sample		6+2	6	44238	3	42	Partial trisomy 7 + partial monosomy 13- translocation?	-
No POC, discordant results	3	No fetal material		NR	1	14640	3	40	Trisomy 16, Trisomy 17	No anomaly detected- XY
	34a	No sample		degrading	<1	8992	2	41	Monosomy 19	dup(1)(p36.32p31.3);dup(6)(p24.2p12.3)- XX
	154	No sample		7+0	4	15568	NR	32	Monosomy 22	No anomaly detected- XX
	270-UHC	Maternal cell contamination		8+0	<1	3284	9	35	No anomaly detected	Invalidated
	433-UHC	No fetal material		NR	3	17567	2	43	Trisomy 14	Trisomy 14, Trisomy 18- XX
	730-UHC	No fetal material		5+6	4	1658	2	30	No anomaly detected	Trisomy 5- XY
	790-UHC	No fetal material		8+1	8	102157	2	37	Failed	del(6)(p22.3p22.1)- XX
	840-UHC	No fetal material		6+0	8	34148	3	39	Monosomy 18	Trisomy 22-XY
Triploid POC results	93	Triploid (F)		9+0	13	4843	2	29	No anomaly detected	No anomaly detected- XX
	268	Triploid (F)		8+0	9	NR	7	26	No anomaly detected	No anomaly detected- XX
	41-UHC	Triploid (F)		NR	7	166537	2	35	No anomaly detected	No anomaly detected- XX
	313-UHC	Triploid (M)		7+2	5	1780	5	31	No anomaly detected	Trisomy 7- XY
	356-UHC	Triploid (M)		7+1	2	3794	11	34	No anomaly detected	No anomaly detected- XX
	437-UHC	Triploid (M)		7+2	2	3538	14	32	No anomaly detected	Invalidated
	601-UHC	Triploid (M)		NR	1	NR	3	30	No anomaly detected	No anomaly detected- XX

Table 10- Concordance sensitivity and specificity for SE-HiSeq-WC and PE-NextSeq-Illumina cfDNA results vs. known cytogenetic analysis from POC testing
529-UHC has been counted twice for both because trisomy 21 was detected and monosomy X was not detected in both. 264 was counted twice for PE-NextSeq-Illumina because it concordantly identified trisomy 13 but also discordantly identified trisomy 16.

	SE-HiSeq-WC	PE-NextSeq-Illumina
Concordance	57.5 %	72.4 %
Sensitivity	55.6 %	75.0 %
Specificity	58.2 %	71.1 %

123 samples were tested using cfDNA analysis and 79 had corresponding POC genetic analysis. Using SE-HiSeq-WC, cfDNA analysis correctly identified 42/73 (57.5%) samples including 10/36 (27.7%) abnormal and 32/37 (86.5%) normal samples. Table 9 compares POC results and cfDNA results. The following anomalies were detected using cfDNA: Trisomy 7 (1), trisomy 15 (3), trisomy 16 (1), Trisomy 21 (2), trisomy 22 (1) and a terminal duplication of chromosome 19 (1). The positive POC results not identified by SE-HiSeq-WC analysis were: a translocation (1), a deletion (1), a duplication (1), trisomy 4 (1), trisomy 5(1), trisomy 11 (1), trisomy 12 (1), trisomy 13 (3), trisomy 15 (2), trisomy 16 (1), trisomy 17 (1), trisomy 18 (1), monosomy 21 (1), trisomy 22 (4), mosaic trisomy 4 (1), mosaic trisomy 16 (1), mosaic trisomy 17 (1) and monosomy X (4). Please note that sample 529-UHC has been counted twice because trisomy 21 was concordant and monosomy X was discordant.

101 of the same samples were also analysed using PE-NextSeq-Illumina and 63 samples had corresponding POC results. Using PE-NextSeq-Illumina, cfDNA analysis correctly identified 42/58 (72.4%) samples including 15/28 (53.6%) abnormal and 27/30 (90.0%) normal samples. Table 9 compares POC results and cfDNA results. The following anomalies were detected using cfDNA: trisomy 7 (1), trisomy 12 (1), trisomy 13 (1), trisomy 15 (4), trisomy 16 (1), trisomy 21 (2), trisomy 22 (1), a duplication (1), a terminal duplication of chromosome 19 (1) and monosomy X (2). The positive POC results not identified by SE-HiSeq-WC analysis were: a translocation (1), a deletion (1), trisomy 5

(1), trisomy 11 (1), trisomy 13 (1), trisomy 15 (1), trisomy 18 (1), monosomy 21 (1), trisomy 22 (3), mosaic trisomy 27 (1) and monosomy X (1). Please note that sample 529-UHC has been counted twice because trisomy 21 was detected and monosomy X was not detected and sample 264 is also counted twice because PE-NextSeq-Illumina concordantly identified trisomy 13 and discordantly identified trisomy 16.

Both false negative and false positive results were identified using both SE-HiSeq-WC and PE-NextSeq-Illumina. In cases of a false negative result, it may result in some unnecessary follow up e.g. parental karyotyping. However, a false negative result could have a higher impact as the potential need for follow up could be missed. For example, a false negative cfDNA result could miss the need to follow up a balanced translocation in a parent(s). It would be best to avoid a false negative result. However, if POC testing is also available it could mitigate the effect of a false negative result.

In total PE-NextSeq-Illumina had the highest number of samples correctly match the known genetic aberration. The results were also grouped using different gestations (Table 11), βhCG value (Table 12) and fetal fraction (Table 13) cut offs to see if this could improve the result calling between the cfDNA and POC cytogenetic analysis. In all categories, PE-NextSeq-Illumina had a better concordance than SE-HiSeq-WC (Table 14).

Table 11- CfDNA results using Gestation thresholds as cut off

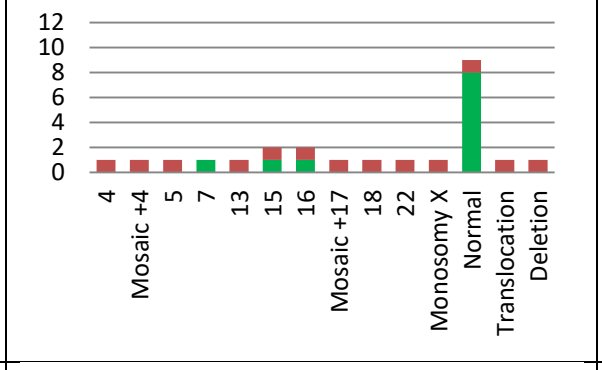
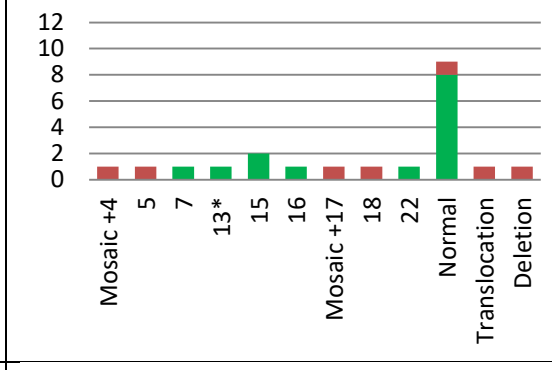
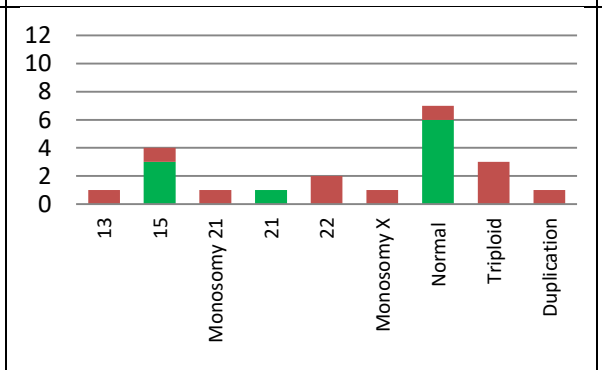
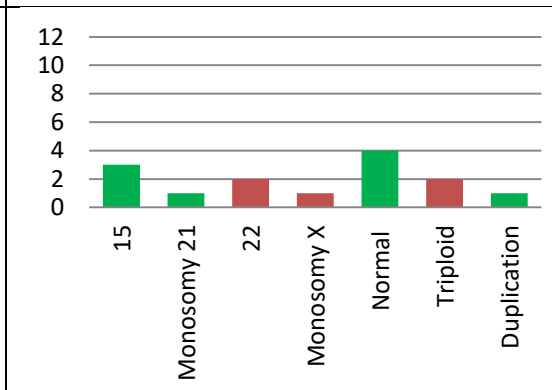
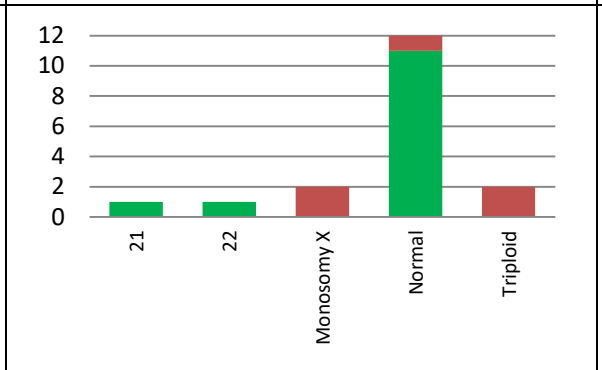
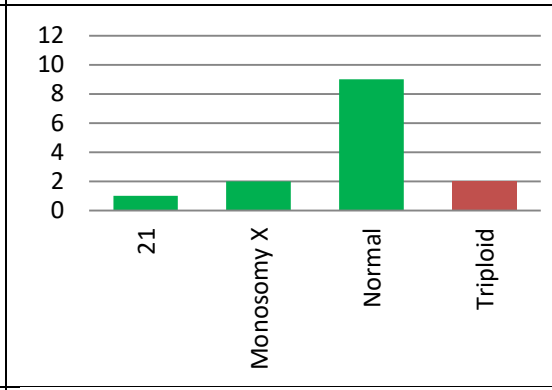
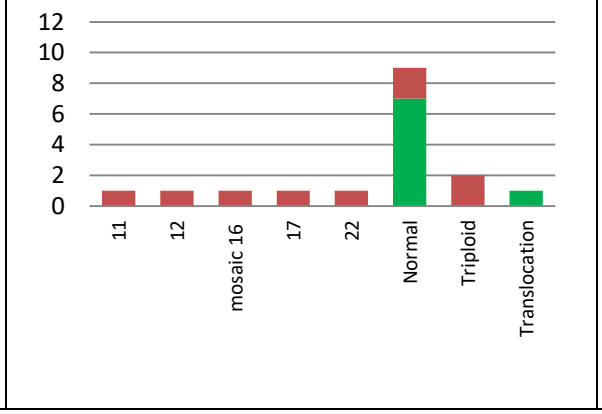
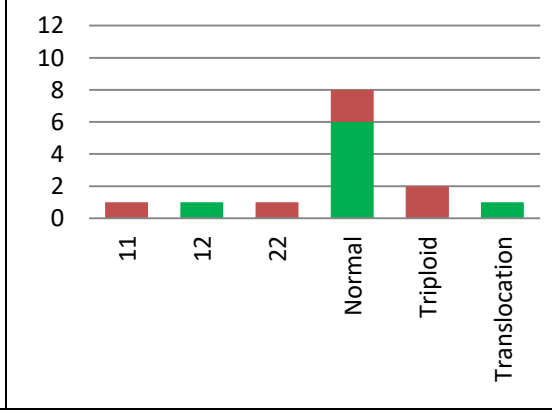
	Verifi + WISECONDOR	VeriSeq + Illumina all chromosome pipeline
<7 weeks		
7<8 weeks		
≥8 weeks		
Unknown		

Table 12- CfDNA results using βhCG thresholds as cut offs

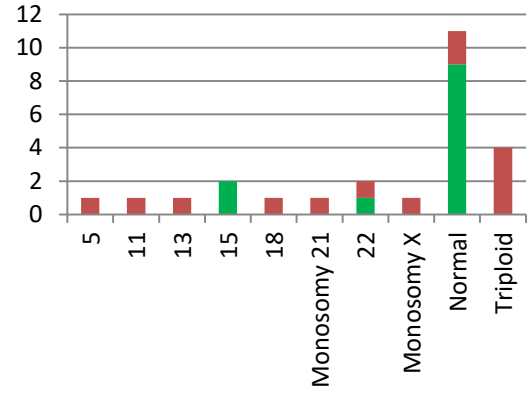
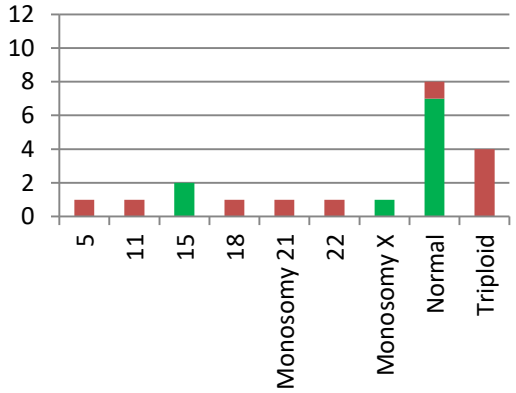
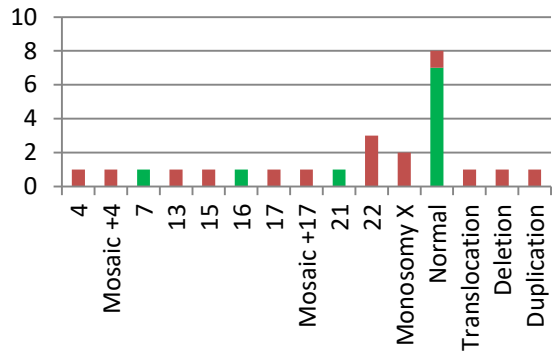
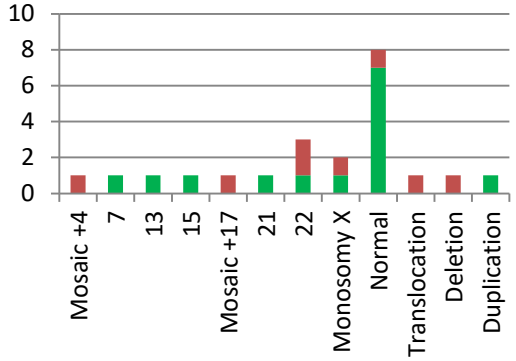
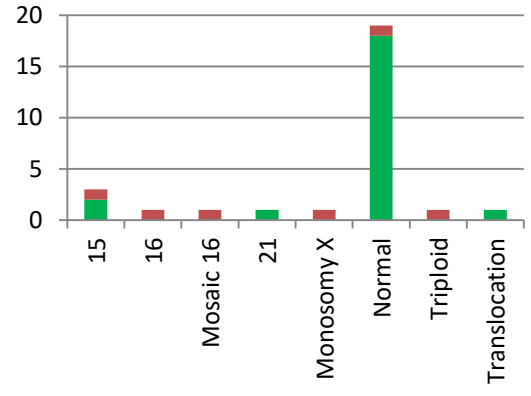
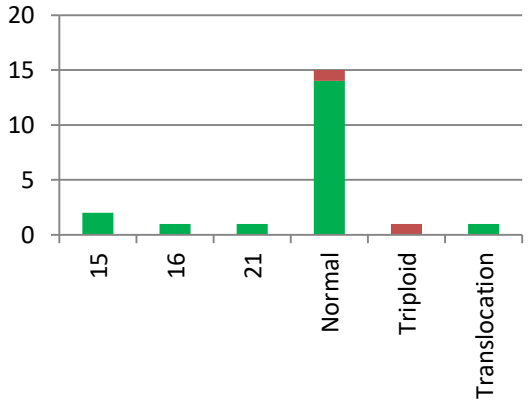
	Verifi + WISECONDOR	VeriSeq + Illumina all chromosome pipeline
<8000		
8000-35000		
>35000		

Table 13- CfDNA results using Fetal Fraction thresholds as cut offs

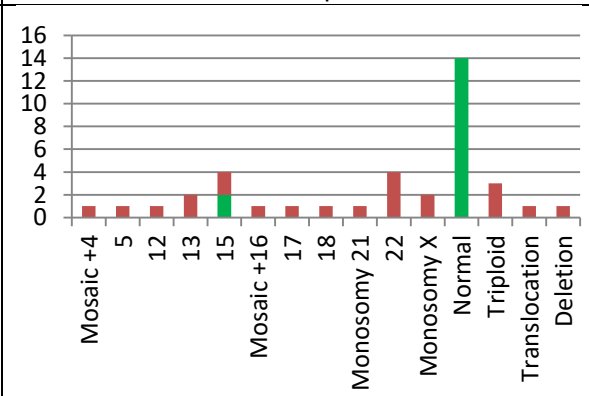
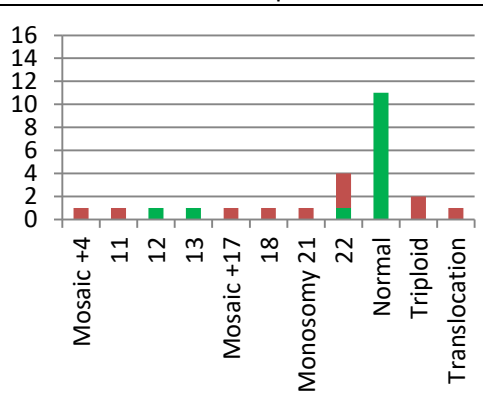
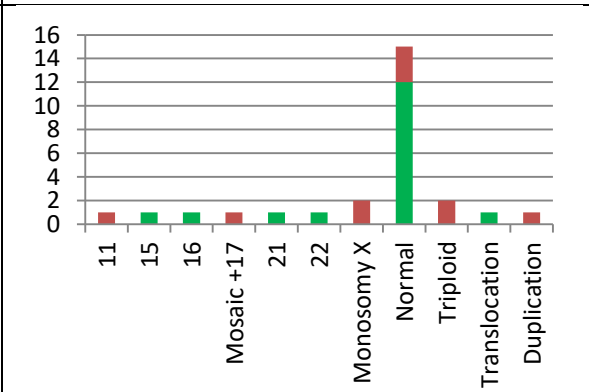
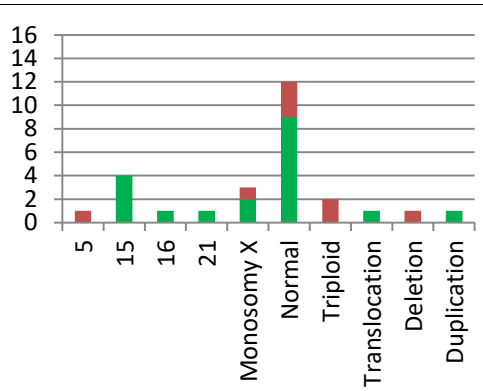
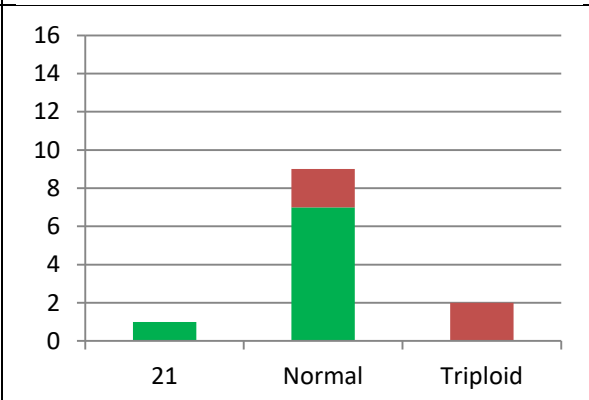
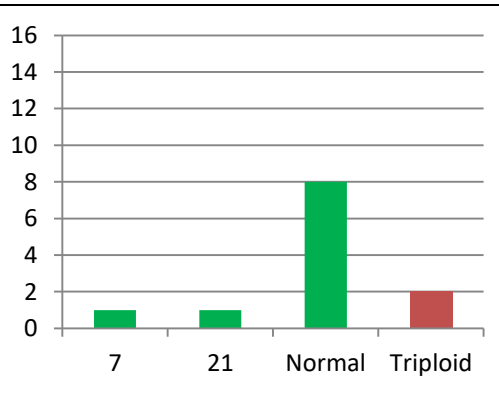
	SE-HiSeq-WC	PE-NextSeq-Illumina
<5 %		
5-8 %		
≥9 %		

Table 14- Concordance with genetic aberrations using SE-HiSeq-WC and PE-NextSeq-Illumina.

		SE-HiSeq-WC			PE-NextSeq-Illumina		
		Total concordance (%)	Abnormal concordance (%)	Normal concordance (%)	**Total concordance (%)**	Abnormal concordance (%)	Normal concordance (%)
Gestation	<7 weeks	**45.8**	20.0	88.9	**66.7**	50.0	88.9
	7<8 weeks	**55.6**	36.4	85.7	**75.0**	62.5	100
	≥8 weeks	**81.3**	50.0	91.7	**100**	100	100
	Unknown	**53.3**	16.7	77.8	**57.1**	50.0	75.0
βhCG	<8000	**57.1**	30.0	81.8	**62.5**	37.5	87.5
	8000-35000	**40.0**	17.6	87.5	**66.7**	53.8	87.5
	>35000	**81.5**	50.0	94.7	**95.0**	100	93.3
Fetal Fraction	<5 %	**45.7**	9.5	100	**60.8**	25.0	100
	5-8 %	**68.0**	50.0	80.0	**76.0**	76.9	75.0
	≥9 %	**80.0**	100	77.8	**100**	100	100

2.4.5 Comparison of miscarriage and live birth samples

Cell free fetal DNA analysis was completed on serial samples using SE-HiSeq-WC. Serial samples were collected approximately every week and the cfDNA results were compared against the pregnancy outcome (Table 15). Where the outcome of the pregnancy was a live birth without physical congenital abnormalities, the cytogenetic result was considered to be normal as it is unlikely for a large chromosomal abnormality to go term without complications. For the six pregnancies which resulted in a pregnancy loss, four had no cytogenetic analysis due to no sample collection or no fetal material in the sample. Of the two miscarriages that did have cytogenetic analysis, one had trisomy 22 and the other had no significant copy number imbalance detected.

Table 15- Serial sample Outcomes from cfDNA using the SE-HiSeq-WC method.

Pregnancies which resulted from live birth were assumed to be euploid.

Patient	Pregnancy Outcome	Genetics	SE-HiSeq-WC					
			Sample 1 Outcome Gestation (Weeks) Fetal fraction (%) βhCG (mIU/mL)	Sample 2 Outcome Gestation (Weeks) Fetal fraction (%) βhCG (mIU/mL)	Sample 3 Outcome Gestation (Weeks) Fetal fraction (%) βhCG (mIU/mL)	Sample 4 Outcome Gestation (Weeks) Fetal fraction (%) βhCG (mIU/mL)	Sample 5 Outcome Gestation (Weeks) Fetal fraction (%) βhCG (mIU/mL)	Sample 6 Outcome Gestation (Weeks) Fetal fraction (%) βhCG (mIU/mL)
4	Miscarriage	Trisomy 22	no anomaly detected 5+0 2 6733	no anomaly detected 6+2 1 12469	no anomaly detected 7+2 1 19247			
34b	Early neonatal death	Unknown	no anomaly detected 6+0 2 35472	Monosomy 19 7+0 3 63527	no anomaly detected 8+2 3 79663	no anomaly detected 9+6 5 75512	no anomaly detected 10+4 4 79352	no anomaly detected 11+0 5 78056
73	Live birth	Assumed euploid	no anomaly detected 6+0 1 NR	no anomaly detected 7+0 2 71692	no anomaly detected 8+0 8 119048	no anomaly detected 9+0 4 109539	no anomaly detected 10+0 8 NR	no anomaly detected 11+0 7 NR
82	Miscarriage	euploid	no anomaly detected 5+5 1 4273	no anomaly detected 6+4 2 5397	no anomaly detected 8+0 2 6844			
91	Live birth	Assumed euploid	no anomaly detected 5+5 2 55859	no anomaly detected 6+4 4 145736	no anomaly detected 7+0 8 >211134	no anomaly detected 9+0 9 >211134	no anomaly detected 10+2 11 171290	no anomaly detected 11+5 10 126870
96	Miscarriage	Unknown	no anomaly detected 5+3 2 325	no anomaly detected 5+0 3 4543				
98	Live birth	Assumed euploid	no anomaly detected 5+0 3 11862	no anomaly detected 6+4 3 42881	no anomaly detected 7+1 5 77680	no anomaly detected 8+0 13 102375	no anomaly detected 9+4 7 117886	no anomaly detected 10+2 10 119603
121	Live birth	Assumed euploid	no anomaly detected 5+0 2 36837		No result 8+3 10 >211134	no anomaly detected 9+4 13 >211134	no anomaly detected 10+6 9 176252	no anomaly detected 11+5 10 110011
151	Live birth	Assumed euploid	no anomaly detected 5+0 1 15449	No anomaly detected 6+0 3 NR	no anomaly detected 7+0 3 118553	no anomaly detected 8+0 3 127584	no anomaly detected 9+0 5 103033	no anomaly detected 11+0 6 53335
153	Live birth	Assumed euploid	no anomaly detected 5+0 4 8904	no anomaly detected 6+0 3 46558	no anomaly detected 7+0 4 114126	no anomaly detected 8+0 8 159513	no anomaly detected 9+0 8 163593	no anomaly detected 10+0 9 146969
168	Miscarriage	Unknown	no anomaly detected 4+3 2 35	no anomaly detected NR 3 404	no anomaly detected 5+6 5 2056	no anomaly detected 6+0 4 3787	no anomaly detected 6+0 <1 NR	

170a	Miscarriage	Unknown	no anomaly detected 4+0 1 1404	no anomaly detected 5+0 <1 4238	no anomaly detected 7+0 3 3318	no anomaly detected 8+0 2 13523		
198	**Miscarriage**	**Unknown**	no anomaly detected 6+0 2 27349	no anomaly detected 6+2 5 58253				
218	**Live birth**	**Assumed euploid**	no anomaly detected 5+6 2 8541	Trisomy 4 6+0 5 NR	no anomaly detected 7+0 4 70222	no anomaly detected 8+0 5 NR	no anomaly detected 9+0 5 95216	no anomaly detected 11+5 7 88048
234	**Live birth**	**Assumed euploid**	no anomaly detected 7+0 7 114288	no anomaly detected 8+2 7 147639	no anomaly detected 9+0 10 113215	no anomaly detected 10+0 6 82105	no anomaly detected 11+0 9 55441	no anomaly detected 13+0 11 61733

Of the 54 samples collected from pregnancies which resulted in a live birth, the SE-HiSeq-WC method detected no anomaly in 51 of the samples. One sample was scored as 'no result' and two samples were false positives (monosomy 19 and trisomy 4).

All the samples which resulted in miscarriage were detected to have no anomaly by the SE-HiSeq-WC method. For one pregnancy loss this was correct as it was known to have no significant copy number balance but one pregnancy loss was known to have trisomy 22, which was not detected by the SE-HiSeq-WC method. These samples had a gestation of 5+0, 6+2 and 7+2 weeks, were detected to have fetal fractions of 2 %, 1% and 1 % respectively, and a βhCG of 6733 mIU/mL, 12469 mIU/mL and 19247 mIU/mL respectively (Figure 18).

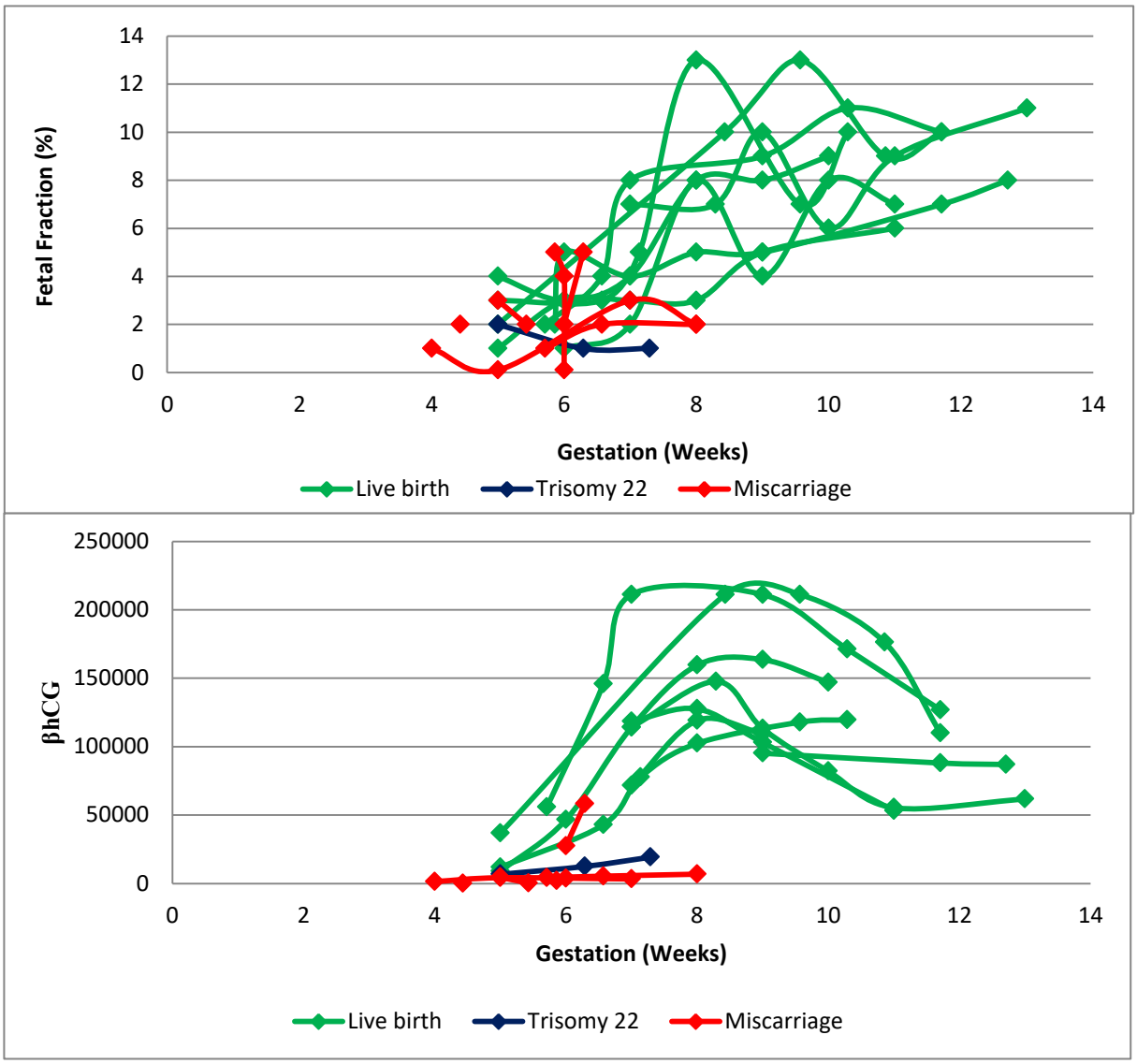


Figure 18- Fetal fraction and βhCG against Gestation for Serial Samples

2.4.6 Threatened Miscarriage Samples

Analysis was completed on threatened miscarriage samples using SE-HiSeq-WC method and compared against the pregnancy outcome. Unfortunately there was not enough maternal plasma aliquoted to complete cfDNA testing on three of the samples, two which resulted in live birth and one that resulted in TOP. Where the threatened pregnancy resulted in a live birth the cytogenetics was considered to be normal. Out of the eight threatened samples which lead to miscarriage, three pregnancies were known to have no significant copy number imbalance. Five of the pregnancies that miscarried had no known cytogenetic outcome.

Out of the 17 threatened samples which resulted in live birth, 16 of the SE-HiSeq-WC results detected no anomaly. One live birth sample was considered to be 'no result' by SE-HiSeq-WC. SE-HiSeq-WC detected no anomaly in the three pregnancies known to have no significant copy number imbalance. In the five pregnancies losses with no known cytogenetic outcome, SE-HiSeq-WC detected no anomaly in three samples, one monosomy 19 and chromosome 12 duplication

Table 16- Samples of patients with viable pregnancies but bleeding and pain ("threatened miscarriage") with follow up outcomes

The gestation of the pregnancy was calculated by ultrasound scan. a, b etc. represents when more than one pregnancy was sampled from the same patient. i represents where a second miscarriage sample was taken.

	Patient	Pregnancy Outcome	POC results (Male/ Female)	POC tissue analysed	Gestation (Weeks)	Fetal fraction (%)	βhCG (mIU/mL)	No. of previous losses	Maternal age (Years)	SE-HiSeq-WC outcome
Miscarriage outcome, Known euploid from POC analysis	51b	Miscarriage	No significant copy number imbalance	Placental villi	8+4	4	107494	7	35	No anomaly detected
	107i	Miscarriage	No significant copy number imbalance	Placental villi	6+0	1	61979	3	24	No anomaly detected
	242i	miscarriage	No significant copy number imbalance	Placental villi	7+0	5	38266	2	25	No anomaly detected
Miscarriage outcome, unknown cytogenetic outcome	11	Miscarriage	Unknown	-	6+0	2	920	3	33	No anomaly detected
	27i	Miscarriage	Unknown	-	7+0	2	346	3	35	No anomaly detected
	170bi	Miscarriage	Unknown	-	6+0	<1	7709	2	41	monosomy 19
	74	Miscarriage	Unknown	-	5+0	3	739	2	39	No anomaly detected
	200	miscarriage	Unknown	-	7+0	3	25316	5	39	Trisomy 12
Live birth	1	Live birth	Assumed euploid	-	5+2	5	19340	6	37	No anomaly detected
	6	Live birth	Assumed euploid	-	6+5	5	58283	2	37	No result
	8	Live birth	Assumed euploid	-	6+3	3	88246	3	38	No anomaly detected
	15	Live birth	Assumed euploid	-	8+0	11	161527	3	24	No anomaly detected
	24	Live birth	Assumed euploid	-	9+1	17	204204	2	27	No anomaly detected
	25	Live birth	Assumed euploid	-	6+0	5	61689	4	41	No anomaly detected
	30	Live birth	Assumed euploid	-	7+0	5	152348	5	27	No anomaly detected
	36	Live birth	Assumed euploid	-	5+4	2	153	5	33	No anomaly detected
	40	Live birth	Assumed euploid	-	7+5	7	150538	3	30	No anomaly detected
	59	Live birth	Assumed euploid	-	8+0	8	141855	2	28	No anomaly detected
	60 ϕ	Live birth	Assumed euploid	-	6+4	5	47852	2	28	No anomaly detected
	60 ϕ	Live birth	Assumed euploid	-	12+4	5	38011	2	28	No anomaly detected
	103	Live birth	Assumed euploid	-	6+0	4	61979	3	22	No anomaly detected
	132	live birth	Assumed euploid	-	6+0	1	29052	4	38	No anomaly detected
	162	live birth	Assumed euploid	-	6+0	7	91182	2	29	No anomaly detected
	209	live birth	Assumed euploid	-	6+4	6	45752	3	26	No anomaly detected
	230	live birth	Assumed euploid	-	6+0	1	40387	3	31	No anomaly detected
Lost to follow up	66	Lost to follow up	unknown	-	6+0	2	10051005	6	24	No anomaly detected

N.B. patients that only had a sample when the pregnancy was threatened and not included in the miscarriage cohort.

2.4.6.1 Threatened to diagnosed samples

Four of the samples collected from the threatened group had an additional sample collected at the time of miscarriage diagnosis; these samples were also included in confirmed cohort. Figure 19 below shows the scan findings for two of these pregnancies.

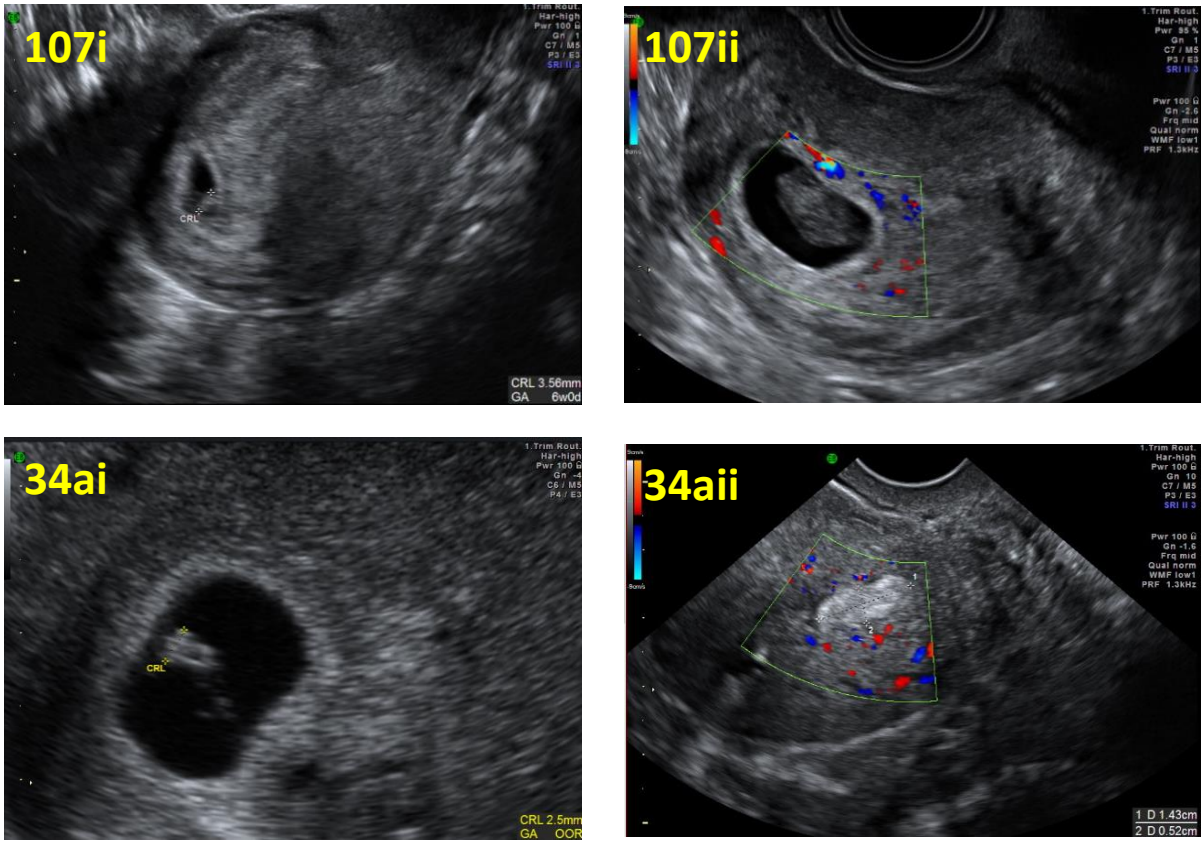


Figure 19- Scan Findings

107i) shows a viable pregnancy at ~6 weeks by ultrasound scan with a CRL of 3.56 mm. This scan was seen at the time the threatened sample was taken. 107ii) shows a non-viable scan at the time of miscarriage diagnosis. Colour Doppler is used to confirm there is no fetal heart. As no vascularity is seen within the conceptus this confirms it is a non-viable pregnancy. Colour Doppler is not used during a healthy pregnancy as the Doppler heat can be detrimental to the pregnancy. 34ai) shows a pregnancy of unknown viability. A fetal pole of 2.5 mm is present but no fetal heart beat is observed. The CRL is too low to predict gestation. 34aii) shows retained POC. The colour Doppler shows no vascularity in the RPOC and confirms it is not a viable conceptus.

2.4.8 No cytogenetic comparison

In total, 44 samples didn't have a corresponding POC result. This was due to either no fetal material in the sample (38.6%), maternal cell contamination (9.1 %) or a POC sample not being received (52.3 %). 37 of these samples were analysed using both SE-HiSeq-WC and PE-NextSeq-Illumina.

Whilst it is difficult to say if the genetic aberrations have been classified correctly, for 23 samples (62.2%) there was no genetic aberration in both SE-HiSeq-WC and PE-NextSeq-Illumina, for 6 samples (16.2 %) SE-HiSeq-WC and PE-NextSeq-Illumina agreed on the same genetic aberration. However, PE-NextSeq-Illumina also detected Monosomy X in the trisomy 13 sample and detected trisomy 18 in the trisomy 14 sample. In 8 samples (21.6 %) without POC results, the aberrations between SE-HiSeq-WC and VeriSeq + PE-NextSeq-Illumina were not concordant (Table 17). There was no significant difference between the βhCG measurement and gestation between concordant and discordant results (Figure 20).

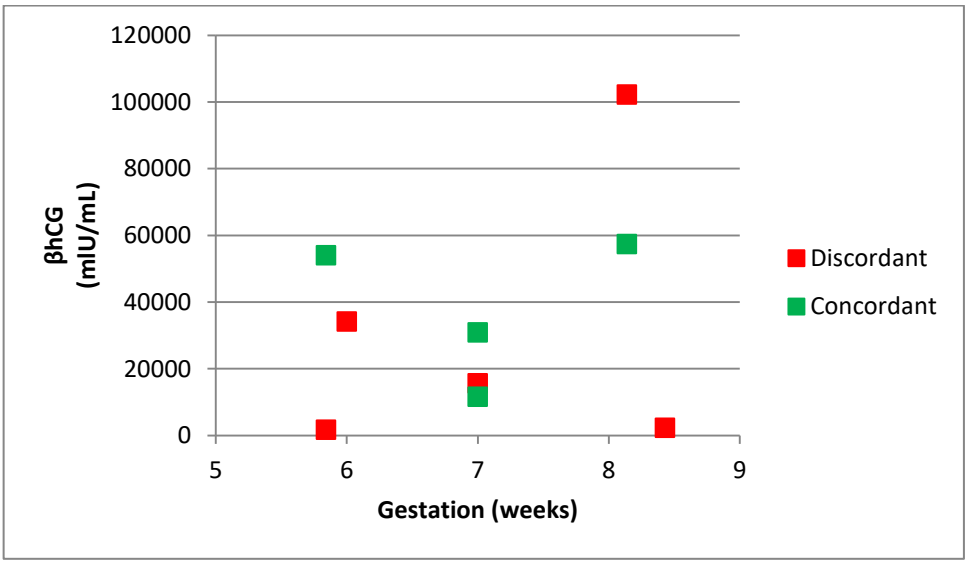


Figure 20- βhCG and gestation for concordant and discordant results were there was no cytogenetic analysis on the POC sample.

Table 17- Difference in results between SE-HiSeq-WC and PE-NextSeq-Illumina.

	βhCG (mIU/mL)	Gestation (weeks)	SE-HiSeq-WC	Fetal Fraction (%)	PE-NextSeq-Illumina	Fetal Fraction (%)
Concordant	30865	7+0	Trisomy 16	4	Trisomy 16	5
	2262	NR	Trisomy 12	NR	Trisomy 12	11
	17567	NR	Trisomy 14	3	Trisomy 14; Trisomy 18	4
	11530	7	Trisomy 16	7	Trisomy 16	5
	53981	5+6	Trisomy 13	8	Trisomy 13, Monosomy X	9
	57373	8+1	Trisomy 20	NR	Trisomy 20	8
Discordant	14640	NR	Trisomy 16 and 17	1	dup(1)(p36.13q31.3);dup(16)(q22.1q24.1)	3
	102157	8+1	Failed	8	del(6)(p22.3p22.1)	NR
	34148	6+0	Monosomy 18	8	Trisomy 22	7
	1658	5+6	No anomaly	4	Trisomy 5	4
	8992	Degrading	Monosomy 19	<1	dup(1)(p36.32p31.3);dup(6)(p24.2p12.3)	2
	15568	7+0	Monosomy 22	4	No anomaly	5
	20565	NR	Chromosome 9 duplication	1	No anomaly	3
	2242	8+3	normal	6	Trisomy 9	4

2.5 Discussion

2.5.1 PAGE Discussion

A limited number of samples from the PAGE study were analysed to evaluate the detection size of SE-HiSeq-WC. From the samples tested, these suggest that the aberration detection size is ~10 Mb. This suggested that WISECONDOR has the potential to pick up relatively small chromosomal aberrations, where the fetal fraction is high enough. As the majority of known chromosome abnormalities associated with miscarriage are aneuploidies and translocations associated with relatively large regions of imbalance, it gave confidence that massively parallel sequencing and analysis using SE-HiSeq-WC would have the potential to detect clinically relevant chromosome abnormalities in the miscarriage samples. WISECONDOR also picked up aberrations which were not confined to Chromosomes 13, 18 and 21.

The PAGE samples were collected from viable pregnancies with fetal structural anomalies detected by ultrasonography and therefore the gestations of these samples were at least 11 weeks gestation, and many in the second trimester. This is much later in pregnancy than samples from patients experiencing miscarriage are likely to be taken. The fetal fraction of the PAGE samples was not confirmed; however it is likely to be higher due to the later gestation at sampling.

2.5.2 cfDNA Discussion

2.5.2.1 Sample recruitment

In total 219 Streck blood samples were collected for the Tommy's study comprising of 72 serial samples, 30 threatened miscarriage samples and 123 confirmed miscarriage samples. Sample collection was less than anticipated so the study was extended to an additional site where samples were collected for diagnosed miscarriages only.

The collection of plasma samples in the case of confirmed pregnancy loss must be completed while the pregnancy is still in situ. This is because cfDNA remains in the maternal plasma as long as the pregnancy remains in situ, potentially due to continued apoptosis/ necrosis of placental cytotrophoblast cells (Tjoa et al., 2006). Once the pregnancy has been expelled the cfDNA does not remain in the maternal circulation for long (Lo et al., 1999) and therefore a maternal blood sample must be taken for cfDNA analysis before the pregnancy has been expelled. This study was limited to patients still having the pregnancy loss *in situ* and therefore the patients who had already passed the tissue were not eligible for the study.

2.5.2.2 Gestations

Tommy's Miscarriage centre is a specialist in the care of families undergoing recurrent miscarriage. These families are very aware of when they first become pregnant and receive careful monitoring during their first trimester. In comparison to other studies (Clark-Ganheart et al., 2015, Yaron et al., 2020) miscarriages are diagnosed earlier by ultrasound scan than the other two studies (Figure 21). This is likely due to the close monitoring of this cohort of Tommy's patients in comparison to the general population.

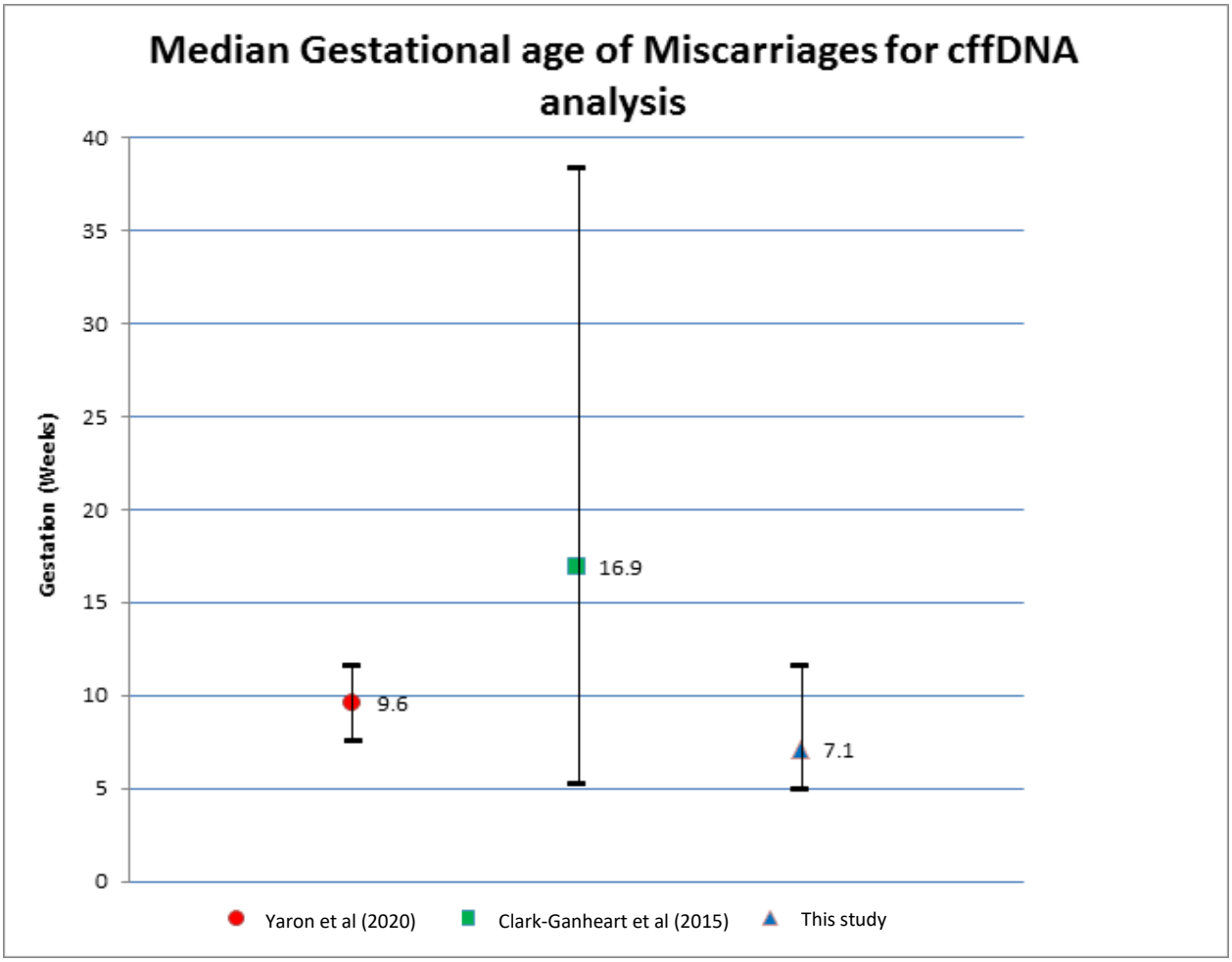


Figure 21- Gestations of pregnacy loss by ultrasound scan (Clark-Ganheart et al., 2015, Yaron et al., 2020)

2.5.2.3 Fetal Fraction

The fetal fraction measures how much cfDNA is in the maternal plasma in comparison to maternal cfDNA. In general it increases with increased gestational age. This is because as the placenta grows more cfDNA is realised into the maternal circulation.

Fetal fraction was measured in all samples tested within this study. In four samples the fetal faction was less than 1 %. Potential reasons for the low fetal fractions could be due to a degrading/ collapsing conceptus, the tissue already having detached from the endometrium

and sitting on the neck of the cervix or from pregnancies with a low gestation and likely to be too soon for there to be sufficient development of the placenta for sufficient detectable levels of cfDNA.

The linear regression shows that in addition to FF, the βhCG measurement increased with increasing gestation age when analysis was completed in the cohort of samples which contained samples resulting in live birth. When the same analysis was completed in miscarriage samples only, there was no significance between the increase of βhCG and the gestation. The fetal fraction was shown to increase with increasing βhCG. The fetal fraction also increased with increasing gestational age. However there was a significant difference for this result between miscarriage and live birth samples. The samples in the serial samples cohort demonstrated that pregnancies which resulted in miscarriage had a fetal fraction increase of 0.2 % per week of gestation compared to pregnancies that resulted in live birth which had a 1 % increase of fetal fraction per week of gestation. Our results suggest that less cfDNA is realised during a miscarriage than in a viable pregnancy as there is a lower fetal fraction. Other literature also suggests there is lower fetal fraction in pregnancies with an adverse outcome (Krishna et al., 2016).

It could be feasible that some women have a lower amount of fetal fraction. This could be because either there is low cfDNA in the maternal circulation or there is an extremely large maternal amount of maternal cfDNA which causes a lower percentage of cfDNA. Factors that affect levels of fetal DNA include gestation, maternal weight/ BMI and aneuploidies (Hestand et al., 2019). There could also be a biological relevance to the health of trophoblast cells which produce a low fetal fraction.

2.5.3 SE-HiSeq-WC and PE-NextSeq-Illumina analysis

We have successfully shown that WISECONDOR detects genetic aberrations over 10 Mb and both SE-HiSeq-WC and PE-NextSeq-Illumina can detect fetal aberrations where there is sufficient fetal fraction. There was a 57.5 % and 72.4 % total concordance (both known normal and abnormal samples) in cfDNA testing in SE-HiSeq-WC and PE-NextSeq-Illumina respectively.

Although SE-HiSeq-WC cannot detect monosomy X and PE-NextSeq-Illumina can, it was decided to include the results of monosomy X in the interpretation for both methods. This is because it gives a more accurate comparison of concordance between the two methods.

Discrepant results between the POC genetic testing and the cfDNA testing could be caused by confined placental mosaicism. cfDNA analysis tests DNA derived from the whole placenta/ cytotrophoblasts only, whereas the POC testing may consist of fetal tissue and whole placental tissue or a small area of cytotrophoblasts. This could result in a discrepancy between the results.

2.5.4 Triploid Samples

Neither SE-HiSeq-WC nor PE-NextSeq-Illumina identified triploid samples, and therefore these were excluded from the primary interpretation. Triploid pregnancies have three copies of each chromosome and thus cannot be detected using a counting-based pipeline for identifying genetic aberrations (Park and Park, 2019). It is interesting to note that all triploid pregnancies were excluded from the data analysis in the (Yaron et al., 2020) cohort. It is

important that triploid samples can be identified rapidly due to the implications in management with regards to the choriocarcinoma risks. Triploid samples can be identified using histology and QF-PCR of POC.

2.5.5 Aims

At the beginning of this chapter, we proposed some questions, including:

- What are the levels of cfDNA are in early miscarriages and can it be used to reliably detect fetal aneuploidies and other chromosome abnormalities, including sub chromosomal abnormalities?
- To address whether the results obtained using cfDNA are comparable to the results achieved using routine testing of products of conception when a sample is available.
- Whether the levels of cfDNA correlate with the levels of beta human chorionic gonadotrophin (βhCG).Below I will discuss each of the aims.

The levels of cfDNA in early miscarriage were between < 1 % to 19 % in our study. While some aberrations were identified with a lower fetal fraction, the total concordance at a fetal fraction of <5 % was 45.7 % and 60.8 % for SE-HiSeq-WC and PE-NextSeq-Illumina respectively, compared to 80 % and 100 % respectively at ≥9 % fetal fraction. Where there is a low fetal fraction, it is likely that both SE-HiSeq-WC and PE-NextSeq-Illumina will call the sample as normal (i.e. a false negative result). This is because there is not enough fetal genetic material to detect the aberration. PE-NextSeq-Illumina appeared to be more accurate at detecting genetic aberrations than SE-HiSeq-WC at a low fetal fraction.

The recommended fetal fraction for NIPT is ≥4 % for trisomy 13, 18 and 21 (Oepkes et al., 2016, Skrzypek and Hui, 2017). From this study it would suggest that some of the less common aberrations are not detectable between 5-8 % fetal fraction, but many aberrations can be detected from >5 % fetal fraction. It is difficult to conclude an exact cut off as to when genetic abnormalities can be detected due to the low sample numbers and biological variation but it could be suggested the RPL cytogenetic follow up be completed using POC sample if the cfDNA sample fetal fraction is <5%.

This study and others (Yaron et al., 2020, Clark-Ganheart et al., 2015) have demonstrated that genetic aberrations can be identified using cfDNA in pregnancy loss. The concordance and inclusion/ exclusion criteria are shown in Table 18 .

Table 18- Concordance and inclusion/ exclusion criteria of SE- HiSeq-WC, PE-NextSeq-Illumina, Clark-Ganheart et al. (2015) and Yaron et al. (2020)

	Concordance (%)	Included	Excluded
SE-HiSeq-WC	57.5	mosaic and subchromosomal anomalies	triploid samples
PE-NextSeq-Illumina	72.4	mosaic and subchromosomal anomalies	triploid samples
Clark-Ganheart et al. (2015)	76	All gestational ages	triploid, mosaic, or subchromosomal
Yaron et al. (2020)	82	triploid samples and using pregnancy loss specific LLR thresholds	mosaic and subchromosomal anomalies

This study shows that some fetal aneuploidies in pregnancy loss are detectable by cfDNA however not all genetic aberrations are detectable. Two known translocations, a terminal

deletion of ~5.84Mb of chromosome 7 at 7q36.2q36.3 and a large terminal duplication of ~21.14 Mb at chromosome 19 at 19q13.12q13.43 (sample ID 133) and a terminal duplication of ~9.23 Mb of chromosome 4 at 4q34.3q35.2 and terminal deletion at ~29.78 Mb of chromosome 5 at 5q33.1q35.3 (sample ID 202). The terminal deletion of chromosome 19 (sample ID 133) was the only aberration to be detected by both SE-HiSeq-WC and PE-NextSeq-Illumina. None of the other aberrations were detected. The preliminary results using samples from the PAGE study suggested that SE-HiSeq-WC cannot detect genetic aberrations below 10 Mb. However, the ~29.78 Mb of chromosome 5 at 5q33.1q35.3 (sample ID 202) was also not detected. This could have been due to the low gestation and/ or low fetal fraction.

Results from cfDNA testing were comparable in some cases using SE-HiSeq-WC and PE-NextSeq-Illumina. Differences could be introduced at several stages in the process, for example the introduction of bias due to PCR within the TruSeq process.

Mosaic genetic aberrations (Sample IDs 51a, 57a and 586-UHC), were not detected by either SE-HiSeq-WC or PE-NextSeq-Illumina. The two explanations for this are the affected fetal fraction was too low to be detectable or that that the mosaicism was confined to the fetus and not present in the placenta/ cytotrophoblast where cfDNA is derived from. Mosaic samples cause the affected fetal fraction to be lower and therefore an affected fetal fraction which is lower than the total fetal fraction is suggestive of mosaicism. Whether mosaicism can be detected or not depends on the fetal fraction, the sequencing depth and the degree of mosaicism.

Statistical analysis showed that the cfDNA levels did correlate with the βhCG levels. The analysis showed that for every 10,000 mlU/mL increase of βhCG the fetal fraction increased by 0.3 % and was statistically significant.

2.5.6 Clinical Application

This study and others (Yaron et al., 2020, Clark-Ganheart et al., 2015) have shown that cfDNA can be utilised in miscarriage. However there are still some genetic aberrations that can be missed (e.g. triploid and mosaic samples). Therefore, cfDNA cannot completely replace current cytogenetic testing. A RPL algorithm has been proposed by (Yaron et al., 2020) which would utilise cfDNA testing in recurrent pregnancy loss (Figure 22).

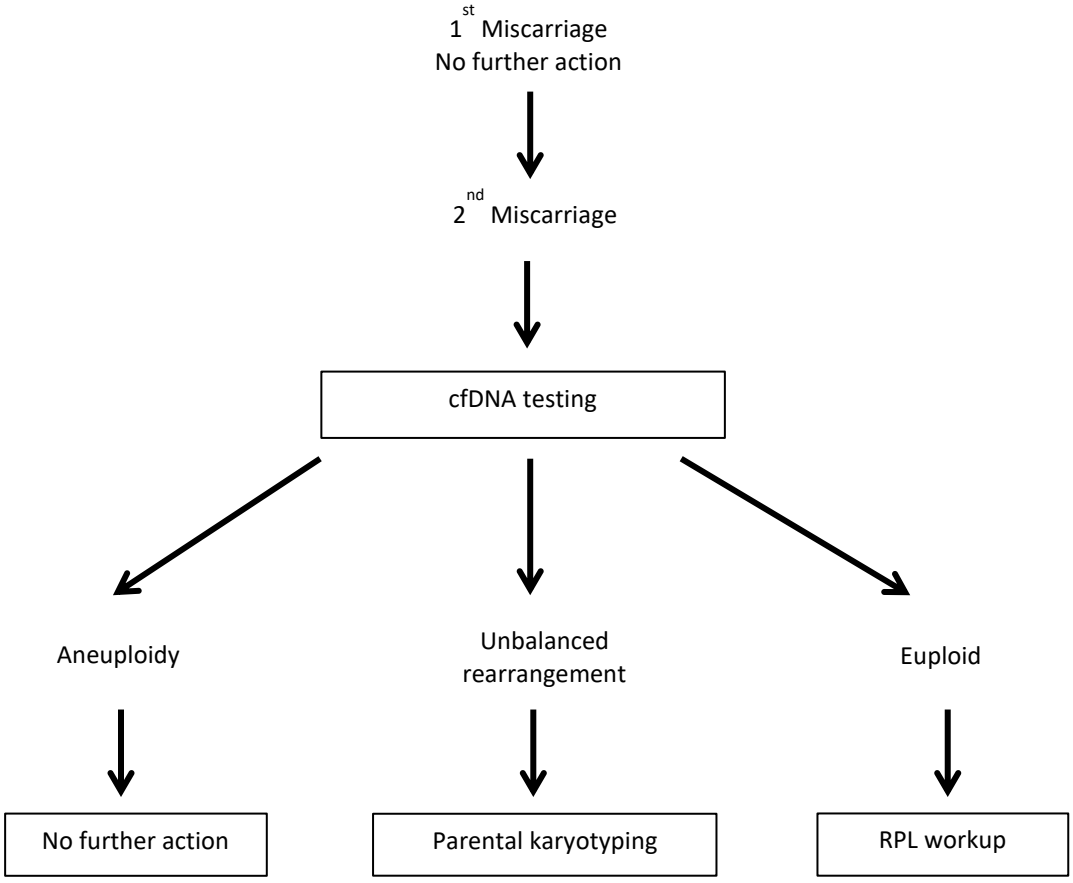


Figure 22- Proposed RPL algorithm by (Yaron et al., 2020)

When a third or subsequent pregnancy loss has been diagnosed, current guidelines are to test the pregnancy tissue for fetal aneuploidies. Providing there is still pregnancy tissue *in situ*, maternal blood samples can be collected to complete cfDNA testing and the collection of POC is also to be advised. If an aneuploidy is detected using cfDNA, no further work is required as numerical errors usually occur sporadically and the likelihood of a successful subsequent pregnancy is not negatively affected. If no genetic aberration is identified then cytogenetic analysis on POC should be initiated to see if there is a genetic aberration that is not detectable by cfDNA testing (e.g. triploid and mosaic samples). This would reduce the need to complete cytogenetic analysis on all POCs, however some genetic aberrations may be missed if POC is not available.

A full economic assessment has not been carried out to assess the costs benefit. However as the price of NGS decrease the cost benefit will greatly improve. When completing a cost assessment the price of alternative management and testing should also be included.

We have shown in this study that translocations can be identified providing the genetic aberration is large enough to be detected. In cases where an unbalanced rearrangement is identified, parental karyotyping should be recommended to assess whether one of the parents is a carrier of this translocation.

In our cohort of known cytogenetic aberrations, using SE-HiSeq-WC analysis, 24 cases were considered to have an aberration and 10/ 24 (41 %) of these cases were confirmed with POC testing. This would mean that 24 cases of POC testing would not need to be completed with the proposed algorithm. POC cytogenetic then has the potential to identify the genetic aberrations not identifiable using cfDNA testing.

2.5.7 Limitations of this Study

The limitations of this study are the low sample size, limited genetic variation and limited number of POC samples for comparable analysis. The results would have been further improved if we had been able to recruit more confirmed miscarriage samples with paired POC results. This could be overcome by extending the recruitment time and locations to increase the sample numbers and genetic variation and opening the study up to include all pregnancy losses regardless as to whether they are recurrent or not. It is difficult to increase the comparable results because pregnancy tissue is not always available from every pregnancy loss. (Yaron et al., 2020) ensured a cytogenetic result by taking a chorionic villus sample for every pregnancy loss. However, this is an invasive procedure and not routine care for this cohort of patients.

The limitations of SE-HiSeq-WC are that triploid pregnancies and sex chromosomes are not detectable (although aberrations in the sex chromosome can be identified from the raw data), the pipeline has not been optimised to detect autosomal monosomies and SE-HiSeq-WC cannot detect true fetal mosaicism. This is also true of PE-NextSeq-Illumina except this pipeline can detect sex chromosome abnormalities.

The samples run using PE-NextSeq-Illumina were run using the VeriSeq NIPT solution V2. The current recommendation for clinical prenatal application is that this testing should only be offered for viable pregnancies at >10 weeks gestation. The cfDNA samples that were analysed in this study were non-viable pregnancies and in most cases < 10 weeks gestation. (Yaron et al., 2020)used 50 of the samples as a training set to optimise the log to log ratio (LLR) scores. The optimised score was then used for the whole cohort of samples. It is likely

that optimising the LLR for SE-NextSeq-Illumina would also give a higher detection rate although this would compromise on the sensitivity and specificity.

Another limitation of this study is that a huge amount of bioinformatician involvement is needed, especially if a trained algorithm is used. This study also didn't complete a full economic assessment for costing.

2.6 Conclusion

Knowing the genetic result of a pregnancy loss can be applied during counselling patients for the prognosis of future pregnancies. It also helps to provide psychological support and relief from the guilt that can be associated with pregnancy loss.

Using cfDNA to diagnose whether a miscarriage was caused by a genetic aberration or not would have a huge clinical impact to patients where conventional cytogenetic testing is not available, either due to the availability of pregnancy tissue for testing or patients preferences. However this is only feasible where the pregnancy is in situ.

In conclusion, cfDNA can be used to detect a genetic aberration in miscarriages providing the maternal plasma sample is collected when the pregnancy tissue is still *in situ* and in cases where there is enough fetal fraction. Where the miscarriage has occurred at a later gestation and there is sufficient cfDNA, genetic aberrations including subchromosomal anomalies > 10 Mb can be identified and these have been comparable to the results obtained by routine testing of products of conception. However, in many cases of early pregnancy loss the fetal fraction is too low to reliably detect a genetic aberration. As it is often earlier gestation/ fetal

losses where there is no fetal tissue available, it is unlikely cfDNA testing can currently be used as an alternative as the fetal fraction is too low to give an accurate and reliable result.

The levels of cfDNA do increase with increasing βhCG. However, as βhCG is comparable within a pregnancy and not between pregnancies it is unlikely this would be able to be used as a predictor of cfDNA fetal fraction. Likewise, it is unlikely to be able to use cfDNA as an early predictor of miscarriage.

Chapter 3: Potential Genetic Causes of miscarriage in euploid pregnancies: a systematic review

Chapter 3- Potential Genetic Causes of miscarriage in euploid pregnancies: a systematic review

This chapter is the systematic review 'Potential genetic causes of miscarriage in euploid pregnancies: a systematic review' and has been published (Appendix 2) in Human Reproduction Update under the following reference:

> Emily Colley, Susan Hamilton, Paul Smith, Neil V Morgan, Arri Coomarasamy, Stephanie Allen, Potential genetic causes of miscarriage in euploid pregnancies: a systematic review, *Human Reproduction Update*, Volume 25, Issue 4, July-August 2019, Pages 452–472, https://doi.org/10.1093/humupd/dmz015

Authors roles: E.C.- study search, study selection, data extraction, quality assessment, and writing. S.H.- data extraction, quality assessment, and editing. P.S.- study design and critical appraisal of manuscript. N.M.- critical appraisal of manuscript and editing. A.C.- study design and critical appraisal of manuscript. S.A.- supervision, study selection, writing, and editing.

3.1 Abstract

Background: Approximately 50% of pregnancy losses are caused by chromosomal abnormalities, such as aneuploidy. The remainder have an apparent euploid karyotype, but it is plausible that there are cases of pregnancy loss with other genetic aberrations that are not currently routinely detected. Studies investigating the use of exome sequencing and chromosomal microarrays in structurally abnormal pregnancies and developmental disorders have demonstrated clinical application and/ or potential utility in these groups of patients. Similarly, there have been several studies that have sought to identify genes that

are potentially causative of, or associated with, spontaneous pregnancy loss, but the evidence has not yet been synthesized.

Objective and rationale: The objective was to identify studies that recorded monogenic genetic contributions to pregnancy loss in euploid pregnancies, establish evidence for genetic causes of pregnancy loss, identify the limitations of current evidence and make recommendations for future studies. This evidence is important in considering additional research into Mendelian causes of pregnancy loss and appropriate genetic investigations for couples experiencing recurrent pregnancy loss.

Search methods: A systematic review was conducted in MEDLINE (1946 to May 2018) and Embase (1974 to May 2018). The search terms “spontaneous abortion”, “miscarriage”, “pregnancy loss” or “lethal” were used to identify pregnancy loss terms. These were combined with search terms to identify the genetic contribution including “exome”, “human genome”, “sequencing analysis”, “sequencing”, “copy number variation”, “single nucleotide polymorphism”, “microarray analysis” and “comparative genomic hybridization”. Studies were limited to pregnancy loss up to 20 weeks in humans and excluded if the genetic content included genes which are not lethal *in utero*, PGD studies, infertility studies, expression studies, aneuploidy with no recurrence risk, methodologies where there is no clinical relevance and complex genetic studies. The quality of the studies was assessed using a modified version of the Newcastle-Ottawa scale.

Outcomes: A total of 50 studies were identified and categorised into three themes; whole-exome sequencing studies; copy number variation studies; and other pregnancy loss related studies including recurrent molar pregnancies, epigenetics and mitochondrial DNA

aberrations. Putatively causative variants were found in a range of genes, including *CHRNA1, DYNC2H1* and *RYR1,* which were identified in multiple studies. Copy number variants were also identified to have a causal or associated link with recurrent miscarriage.

Wider implications: Identification of genes causative of or predisposing to pregnancy loss will be of significant individual patient impact with respect to counselling and treatment. In addition, knowledge of specific genes that contribute to pregnancy loss could also be of importance in designing a diagnostic sequencing panel for patients with recurrent pregnancy loss and also in understanding the biological pathways that can cause pregnancy loss.

Key words: genetic causes, pregnancy loss, euploid miscarriage, exome sequencing, chromosomal array, single nucleotide variation, copy number variant.

3.2 Introduction

Pregnancy loss is the most common complication of pregnancy and multiple studies have sought to identify potentially causative genes involved in pregnancy loss. In this systematic review, we have synthesized the evidence to establish a genetic causality of pregnancy loss.

3.2.1 Miscarriage and recurrent pregnancy loss

Approximately 15 % of clinically recognised pregnancies end in pregnancy loss, with the majority occurring during the first trimester. Of these, 50 % are caused by chromosomal abnormalities such as aneuploidy (Hassold et al., 1980) and can be detected by conventional cytogenetic analysis. It is suggested that 86 % of these abnormalities are numerical, 6 % are

structural abnormalities and 8 % are due to other genetic mechanisms, such as chromosomal mosaicism and molar pregnancies (Goddijn and Leschot, 2000).

Recurrent Miscarriage (RM) is defined by the Royal College of Obstetricians and Gynaecologists (RCOG) as at least three consecutive miscarriages before 24 weeks gestation (Bakker et al.) and recurrent pregnancy loss (RPL) by the European Society of Human Reproduction and Embryology (ESHRE) November 2017 guidelines as the loss of two or more pregnancies (Weiss). In addition to genetic aetiology, a spectrum of non-genetic causes of RPL have also been identified, including thrombophilic factors, endocrinological causes, immunological and immunogenetic causes, sperm DNA fragmentation, uterine malformations and lifestyle factors such as smoking, reviewed by (Larsen et al., 2013a).

3.2.2 Cytogenetic and chromosomal microarray analysis

Traditionally, cytogenetic analysis of pregnancy tissue has been performed to identify genetic causes of RPL and to indicate the need for further analysis of parental samples where there is the possibility of a balanced chromosome rearrangement (e.g. translocation) in one of the parents. It is important to identify any numeric chromosome errors, such as trisomy, monosomy or polyploidy, since these are causes of pregnancy loss which usually occur sporadically, and the likelihood of a successful pregnancy outcome is not negatively affected in subsequent pregnancies. Where there is a balanced translocation in one of the parents, genetic counselling is important as there is likely to be a recurrence risk in future pregnancies and pre-implantation genetic testing (PGT), chorionic villus sampling or amniocentesis can be used to detect an abnormality in the conceptus. However, for couples with a translocation, medical management (e.g. natural conception and observation) has

been reviewed to have a higher live birth rate than IVF/PGD (Franssen et al., 2011, Hirshfeld-Cytron et al., 2011).

The most recent ESHRE guidelines for genetic analysis of products of conception (POC) give a conditional recommendation for genetic analysis but recommend that testing is carried out by array- comparative genomic hybridisation (CGH) instead of traditional karyotyping (Bender Atik et al., 2018). Conventional karyotype analysis identifies balanced and unbalanced chromosomal rearrangements and copy number variants (CNVs) to ~5Mb resolution. Chromosomal microarray analysis can now identify unbalanced CNVs below 1 Mb, with a resolution at the level of individual exons of genes in targeted regions of the genome (Miller et al., 2010). Microarray analysis is also less labour intensive as it is based on DNA analysis rather than cultured cells and has a higher success rate in poor quality tissue samples; however, the quality of tissue will impact the success and failure rate of both conventional karyotyping and array-CGH. Array-CGH has become the gold standard for genetic CNV analysis. It should, however, be noted that array CGH may miss some balanced chromosomal rearrangements and may also fail to identify maternal cell contamination.

3.2.3 Other genetic causes

In the case of pregnancy loss, with an apparently euploid karyotype, there may be genetic aberrations causative of pregnancy loss that are not currently known or routinely assessed. These could include single-nucleotide variants (SNVs) that affect individual genes and are detectable by sequencing or small sub-microscopic aberrations that affect a cluster of genes and are detectable by microarray analysis. In the case of SNVs, this is particularly important as many may follow a recessive or X-linked pattern of inheritance and therefore have a high

recurrence risk. CNVs detected in cases of pregnancy loss may unmask a recessive mutation in a relevant gene or involve dosage-sensitive genes, where loss or gain of copies affects the gene function. These regions may also represent benign CNVs seen frequently with no recorded effect on phenotype, although it remains possible that some may be involved in RPL. Evidence in humans and other species (Wilson et al., 2016b) suggests that many genes are important in early development, and can lead to embryonic lethality when functionally 'knocked out', resulting in pregnancy loss. More widespread genetic analysis of embryonic pregnancy loss may provide an opportunity to identify genes that are essential in early human development or where a lack of function leads to pregnancy loss.

3.2.4 Molar pregnancies

A molar pregnancy or hydatidiform mole (HM) is an abnormal pregnancy, which has cystic degeneration of the chorionic villi, abnormal proliferation of the trophoblast and abnormal development of the fetus. These can either be complete HM or partial HM, distinguishable by the extent of trophoblast proliferation and presence of embryonic tissue.

Complete HMs are usually diploid with all chromosomes of paternal origin. The majority arise from an anuclear ovum being fertilised by a haploid sperm and replicating its own chromosomes (uniparental paternal isodisomy) or rarely from an anuclear ovum fertilised by two sperm (uniparental paternal heterodisomy). Partial HMs are mostly triploid with 23 chromosomes of maternal origin and 46 of paternal origin.

Whilst HMs are usually triploid and sporadic and therefore outside the scope of this review, a minority of molar pregnancies are diploid and biparental, usually being recurrent and

familial. These may be caused by maternal autosomal recessive mutations in genes, such as *NLRP7* and *KHDC3L,* resulting in an abnormal epigenotype of imprinted loci. This results in abnormal gene expression, which causes abnormal placental trophoblast development and manifests as HM (Carey et al., 2015).

3.2.5 Whole exome sequencing

Advances in sequencing technology, including whole-exome sequencing (WES) and whole-genome sequencing (WGS), are increasingly providing the opportunity to detect genetic sequence variation and to characterise genetic mutations causing disease. WGS is the most extensive sequencing method and targets the entire genome, whereas WES targets the exome, which is the protein-coding region of the DNA. The exome makes up ~1 % of the human genome, and it is estimated to contain 85 % of the genetic mutations associated with disease (Choi et al., 2009). Generally, WES is the preferred method of sequencing because it is cheaper than WGS and has a smaller, more manageable data set while still comprehensively covering the coding regions of DNA. WGS has the advantage of analysing and giving a comprehensive view of the whole genome and has the potential to detect large structural variants, insertions/ deletions, SNVs and copy number changes. However, still relatively little is understood about the non-coding regions of the genome.

Studies investigating the use of WES in structurally abnormal pregnancies, late pregnancy losses and developmental disorders (Wright et al., 2015, Shamseldin et al., 2018, Carss et al., 2014) have demonstrated the clinical application in these patients. However, very few WES studies have reported analysis in pregnancy loss or lethal genes that could contribute to RPL. The few studies using WES to look for genetic aberrations in RPL have also tended to

represent only small patient cohorts. The ability to recognise and detect genetic mutations may have implications for routine genetic testing and clinical practice, especially when a pathogenic aberration is identified that can be reliably detected in future pregnancies.

3.2.6 Aims

There are several studies that have sought to identify genes causative of or associated with pregnancy loss, but the evidence has not yet been synthesised. We propose to review these studies and establish evidence of genetic causality of RPL, including reviewing appropriate methodologies. We will evaluate studies investigating Mendelian inheritance patterns, including autosomal recessive and dominant X-linked inheritance, and also *de novo* genetic causes, but we have excluded studies investigating more complex genetic associations, which have recently been systematically reviewed (Yaron et al., 2020).

3.3 Methods

3.3.1 Registration

This systematic review has been registered with PROSPERO (CRD42017073910).

3.3.2 Search

A systematic literature review to assess the studies investigating the genetic contribution to RPL was conducted in MEDLINE (1946 to May 2018) and Embase (1974 to May 2018) using Ovid (https://ovidsp.tx.ovid.com). The search terms used to identify pregnancy loss were ‘spontaneous abortion’, ‘miscarriage’, ‘pregnancy loss’ or ‘lethal’, and the search terms to identify the genetic contributions were ‘exome’, ‘human genome’, ‘sequencing analysis’,

'sequencing, copy number variation', 'single nucleotide polymorphism', 'microarray analysis' and 'comparative genomic hybridisation'. The search terms and corresponding Mesh terms are shown in Supplementary Table 1. Additional studies were also identified from references of selected studies.

3.3.3 Study selection

Studies were selected by two independent reviewers (E.C. and S.A.). Studies were first screened for eligibility using article titles and then by screening the study abstracts. Studies were included if they had pregnancy loss up to 20 weeks, but were not restricted if they also included some later losses, providing the genetic aberrations were defined. Studies were excluded if the genetic content included genes which were not lethal *in utero*, PGD studies, infertility studies, expression studies, aneuploidy with no recurrence risk, methodologies where there is no clinical relevance, and complex genetics. Both recurrent and sporadic pregnancy losses were included. The full inclusion and exclusion criteria are presented in Supplementary Table 2.

3.3.4 Data extraction process

Data on publication date, country, study objective, sample, phenotype and gestation, methods and analysis, study outcome and quality scores were extracted (E.C. and S.H.). Data extraction was checked by a second reviewer (S.A.). Each of the identified genes were found in Online Mendelian Inheritance in Man (OMIM) and the Mendelian Inheritance in Man (MIM) number, gene name, gene function, associated disease/phenotype and cytogenetic location were ascertained.

3.3.5 Quality assessment

The quality of each study was assessed (E.C. and S.H.) using a modified Newcastle-Ottawa scale (Supplementary Table 3). Each study was scored out of 12 and was judged on the sample size, inclusion/exclusion criteria, the genetic analysis method, statistical analysis, case definition, controls and comparability. The breakdown of each study's score is included in Supplementary Table 4.

3.4 Results

A total of 50 studies were included in the review. The initial search of the Medline and Embase databases identified 3404 potentially relevant articles. After screening the titles and abstracts, 74 full texts were obtained for detailed review. A total of 30 full articles were excluded because they were either not related to pregnancy loss, were more than 20 weeks gestation, or contained no genetic content. Examination of the bibliographies and journal indices generated six additional studies for the review. Figure 23 illustrates the study selection. The papers identified were categorised into three themes; WES studies, CNV studies and other studies related to pregnancy loss related studies including recurrent molar pregnancies.

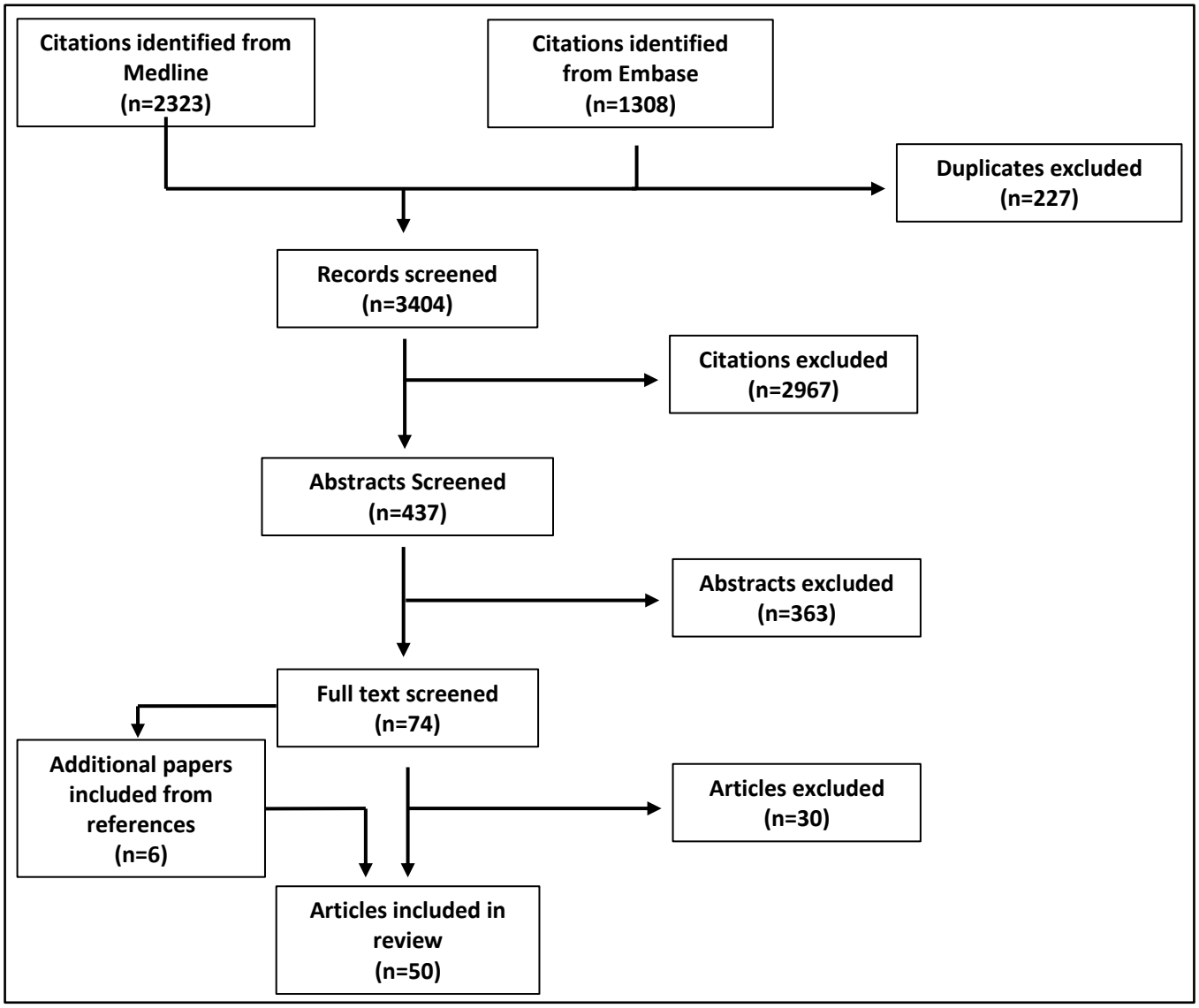


Figure 23- PRISMA flow diagram for the systematic review of the Potential genetic causes of miscarriage in euploid pregnancies.

The 50 studies that met the inclusion and exclusion criteria were all published in English between 2009 and 2018. Out of the studies identified, 21 were from Europe, 14 were from North America, 13 were from Asia and there was one study each from South America and Africa.

3.4.1 Whole exome sequencing

Thirteen studies were identified (Table 19) that used whole exome sequencing (WES) to identify single nucleotide variants (SNVs) in families with multiple pregnancy losses, a combination of pregnancy losses and terminations and termination of pregnancies for fetal scan findings. Eight of these studies focused on a single couple only (Bondeson et al., 2017, Cristofoli et al., 2017, Dohrn et al., 2015, Filges et al., 2014, Rae et al., 2015, Shamseldin et al., 2013, Tsurusaki et al., 2014, Wilbe et al., 2015). Six studies used WES analysis of trios (Filges et al., 2014, Dohrn et al., 2015, Wilbe et al., 2015, Cristofoli et al., 2017, Bondeson et al., 2017, Qiao et al., 2016).

Studies using WES identified variants in genes from both fetal and parental samples, thus allowing for the inheritance to be identified. One study identified compound heterozygous mutations in *KIF14* in a family with unexplained euploid miscarriages (Filges et al., 2014). The other studies included pregnancies terminated for a fetal abnormality including a homozygous missense mutation in *ECEL1* from a consanguineous couple with pregnancies terminated due to arthrogryposis multiplex congenita (Dohrn et al., 2015); a novel homozygous mutation in the *MUSK* gene in a non-consanguineous couple with a history of fetal akinesia deformation sequence (FADS) (Wilbe et al., 2015), compound heterozygous mutations in *STIL* from a non-consanguineous couple with fetal microcephaly (Cristofoli et al., 2017), a homozygous nonsense mutation in *CEP55* in a non-consanguineous family with two fetuses with Meckel-like syndrome (Bondeson et al., 2017) and compound heterozygous mutations in *IFT122* in a couple experiencing both RPL and later losses with scan abnormalities (Tsurusaki et al., 2014).

Two studies (Rae et al., 2015, Shamseldin et al., 2013) identified pathogenic variants by WES of fetuses affected with hydrops fetalis. The first identified pathogenic variant in the gene *FOXP3* was from a non-consanguineous couple whom had multiple male pregnancy terminations. *FOXP3* is an X-linked gene that is known to cause fetal akinesia syndrome (Rae et al., 2015). The second identified a novel mutation in the gene *CHRNA1* in a consanguineous couple (Shamseldin et al., 2013). Autosomal recessive mutations in this gene are also known to cause fetal akinesia.

A single study identified a homozygous missense variant in *NOP14* in pregnancy loss material from two consanguineous Iranian couples experiencing RPL. WES was completed on fetal tissue samples and the heterozygous copies of the variant were confirmed in the parents using Sanger sequencing (Suzuki et al., 2018).

Studies also used WES in larger cohorts. One study (Shamseldin et al., 2015) looked at consanguineous couples with two or more pregnancies diagnosed with non-immune hydrops fetalis (NIHF). Seven pathogenic variants previously known to cause NIHF (Shamseldin et al., 2015) were identified from 24 consanguineous couples with lethal NIHF.

Two Studies (Ellard et al., 2015, Qiao et al., 2016), analysed non-consanguineous couples with RPL. Variants in *GLE1*, *RYR1* and *DYNC2H1* were identified using WES of parental samples only (Ellard et al., 2015). Compound heterozygous variants were also identified in *DYNC2HI* and *ALOX15* in seven euploid pregnancy losses from four families (Qiao et al., 2016).

The final study used a slightly different approach and analysed a panel of 234 pre-selected RPL candidate genes from women affected by RPL. Using WES and bioinformatic filtering of

non-synonymous sequence variants, 27 variants were identified from the previously selected genes (Quintero-Ronderos et al., 2017). The genes in which variants were identified in the described sequencing studies are detailed in Table 20. However, genes from Quintero-Ronderos et al. (2017) have been excluded because they were from a pre-selected gene panel and therefore would introduce bias.

Table 19- Study characteristics from Single nucleotide variants by whole exome sequencing studies

Author, date and country	Study objective	Sample, phenotype and gestation	Methods and analysis	Study outcome	Quality scores
Rae et al., 2015 *Southampton, UK*	Assessment of genes causing pregnancy loss of male pregnancies.	One woman with multiple male pregnancy loss (hydrops fetalis). Two fetal tissue samples at 18 and 20 weeks gestation.	Fetal DNA screening using NGS panel (Sure Select XT kit – no information on gene number given).	Identified pathogenic variant in X-linked *FOXP3* gene. Known to cause IPEX syndrome, but gene not previously linked to NIHF.	3
Shamseldin et al., 2013 *Riyadh, Saudi Arabia*	Identify causative mutation in a family with RPL due to NIHF.	Consanguineous couple with two previous fetal losses and a history of fetal hydrops. One fetal tissue at 19 weeks gestation.	WES of affected fetus and filtering of variants by autozygome.	Identified novel mutation in *CHRNA1*, known to cause multiple pterygium and fetal akinesia syndrome.	6
Filges et al., 2014 *Vancouver, Canada*	Identify causal variants for a recurrent pattern of an undescribed lethal fetal congenital anomaly syndrome.	Two pregnancy tissue samples with normal female karyotypes at 21+4 and 18+5 weeks gestation from one family (non-consanguineous).	WES of DNA from parents and chorionic villus samples. Variants from trio filtered for autosomal recessive inheritance (homozygous or compound heterozygous).	Compound heterozygosity for two non-synonymous truncating mutations in *KIF14* in the one fetus.	4
Suzuki et al., 2018 *Yokohama, Japan*	Identify causative variant(s) of two consanguineous families with RPL.	Two Iranian consanguineous families with RPL (nine and two pregnancy loss in each family).	WES on fetal tissue samples. Filtering for compound heterozygosity and homozygous variants.	Homozygous variant in *NOP14* gene in both pregnancy losses with parents having heterozygous variants confirmed by Sanger sequencing.	3
Dohrn et al., 2015 *Aalborg, Denmark*	Clinical report of consanguineous couple with a history of pregnancies with Arthrogryposis multiplex congenita (AMC).	Two pregnancy tissue samples terminated at 14+2 and 13+4 weeks gestation from one consanguineous family.	WES of parents and fetuses. Variants filtered for autosomal recessive inheritance (homozygous in fetus).	Homozygous missense variant in *ECEL1*.	3
Wilbe et al., 2015 *Uppsala, Sweden*	WES on family trio with a history of recurrent fetal loss and fetal akinesia deformation syndrome.	One family trio (mother, father and 17 week fetal tissue) with five affected fetuses with severe/ lethal form of FADS.	WES on parental and one affected fetus. Variant filtering to identify potentially damaging variants following a recessive inheritance pattern.	Novel homozygous mutation in *MUSK* gene leading to a frameshift mutation and predicting a premature stop codon.	7
Cristofoli et al., 2017 *Leuven, Belgium*	Identify causative mutation in a family with profound fetal microcephaly associated with delayed gyrification of the corpus callosum.	One couple with 5 first trimester pregnancy losses. Two fetal tissue terminations sampled at 20 and 25 weeks gestation.	WES on fetus and parents using OneSeq Constitutional Research Panel for exome enrichment. Trio filtering of variants for recessive, *de novo* and hemizygous inheritance.	Identification of novel compound heterozygous mutation in *STIL* in both fetuses.	6
Bondeson et al., 2017 *Uppsala, Sweden*	Uncover the genetics of suspected autosomal recessive lethal fetal ciliopathy Meckel-like syndrome in a family with two affected fetuses with Meckel-like syndrome and 1 healthy sibling.	Family with four pregnancies (one pregnancy loss, one TOP, one IUFD, one healthy girl).	WES of trio (20 week TOP) with variant filtering for potentially damaging variants following recessive inheritance pattern. SNPs identified investigated for shared ancestral haplotype analysis.	Homozygous nonsense mutation in *CEP55* causing a premature stop codon in the affected fetus. The IUFD fetus was homozygous and the mother, father and healthy sibling were all heterozygous for the mutation.	5
Tsurusaki et al., 2014	Identify causative mutation in a	39-year-old women with two	WES for maternal sample, healthy	Compound heterozygous mutation in	33

Yokohama, Japan	family with RPL and scan abnormalities.	terminations (13 and 21 weeks), one intrauterine fetal death (13 weeks), four RPL (6-8 weeks) and one healthy son.	son and 21 week TOP due to skeletal anomalies.	*IFT122* identified in the TOP. Same heterozygous mutation was found in pregnancy loss tissue at 7 weeks by PCR analysis. One mutation identified in maternal sample. Mutation known to cause Cranioectodermal dysplasia-1.	
Shamseldin et al., 2015 *Riyadh, Saudi Arabia*	Analysis of consanguineous couples in which embryonic lethality appears to follow a Mendelian recessive pattern.	Twenty-four consanguineous couples with two or more pregnancies diagnosed with lethal NIHF.	Autozygosity mapping using genome-wide genotyping array from affected fetus. WES on affected fetus filtered for the autozygome and novel homozygous variants.	In eight fetuses, variants known to cause NIHF were detected in four genes (same gene affected in five families). Mutations in seven novel candidate genes were identified in 10 fetuses (same gene affected in four families).	5
Ellard et al., 2015 *Exeter, UK*	Investigate a strategy for diagnosis of autosomal recessive lethal disorders.	Three couples with multiple pregnancy terminations due to presumed autosomal recessive disorder.	WES of parental DNA samples only, filtering variants for autosomal recessive inheritance (homozygous or compound heterozygous). Pregnancy tissue subsequently tested for variants using sanger sequencing.	Proof of principle study detected variants in *GLE1*, *RYR1* and *DYNC2H1* (fetal akinesia syndrome). Demonstrates that WES of parental samples can be effective in diagnosing lethal recessive disorders.	8
Qiao et al., 2016 *British Columbia, Canada*	WES in recurrent pregnancy loss.	Seven fetal tissues from four families with RPL at <20 weeks gestation.	WES on couples and pregnancy tissue. Variant filtering for autosomal recessive, compound heterozygous and X-linked inheritance; pathway and disease association enrichment analysis.	Compound heterozygous variants identified in *DYNC2HI* and *ALOX15*.	5
Quintero-Ronderos et al., 2017 *Bogotá, Colombia*	Analyse non-synonymous sequence variants in RPL candidate genes for dissection of complex disease aetiology.	49 women with RPL at <20 weeks gestation.	WES of women (not partners), bioinformatic filtering of non-synonymous sequence variants in subset of 234 RPL candidate genes selected from review of literature and databases.	27 variants, in 22 genes, affecting 20 patients were considered strong candidates for relating to RPL molecular aetiology.	7

Quality scores were assessed using a modified Newcastle–Ottawa scale. FADS, fetal akinesia deformation sequence; IPEX, immunodysregulation polyendocrinopathy enteropathy X-linked; IUFD, intrauterine fetal death; NGS, next generation sequencing; NIHF, non-immune hydrops fetalis; RPL, recurrent pregnancy loss; SNP, single-nucleotide polymorphism; STIL, SCL/TAL1-interrupting locus; TOP, termination of pregnancy. ALOX15, 15-lipoxygenase, reticulocyte arachidonate; CEP55, centrosomal protein, 55-KD; CHRNA1, cholinergic receptor, nicotinic, alpha polypeptide 1; DYNC2H1, dynein, cytoplasmic 2, heavy chain 1; ECEL1, endothelin-converting enzyme-like 1; FOXP3, forkhead box P3; GLE1, RNA exportmediator; IFT122, intraflagellar transport 122; KIF14, kinesin family member 14; MUSK,muscle, skeletal, receptor tyrosine kinase, NOP14, nucleolar protein 14; RYR1, ryanodine receptor 1.

Table 20- Genes identified by WES studies

Gene	Paper Author	Gene/Locus name	Gene/Locus MIM number	Gene Function	Associated Disease/phenotype	Cytogenetic location
ALOX15	Qiao et al., 2016	15-LIPOXYGENASE, RETICULOCYTE ARACHIDONATE	152392	Implicated in anti-inflammation, membrane remodelling, and cancer development/metastasis.	None	17p13.2
CEP55	Bondeson et al., 2017	CENTROSOMAL PROTEIN, 55-KD	610000	Regulator of the final stages of mitosis.	Multinucleated neurons, anhydramnios, renal dysplasia, cerebellar hypoplasia, and hydranencephaly	10q23.33
CHRNA1	Shamseldin et al., 2013, Shamseldin et al., 2015	CHOLINERGIC RECEPTOR, NICOTINIC, ALPHA POLYPEPTIDE 1	100690	Nicotinic acetylcholine receptor, electrical signalling between nerve and muscle cells.	Myasthenic syndrome, congenital, Multiple pterygium syndrome	2q31.1
CTSA	Shamseldin et al., 2015	CATHEPSIN A	613111	Multifunctional enzyme, lysosomal multienzyme complex.	Galactosialidosis	20q13.12
DNAH14	Shamseldin et al., 2015	DYNEIN, AXONEMAL, HEAVY CHAIN 14	603341	Microtubule-associated motor protein complexes.	None	1q42.12
DYNC2H1	Ellard et al., 2015, Qiao et al., 2016	DYNEIN, CYTOPLASMIC 2, HEAVY CHAIN 1	603297	Ciliary intraflagellar transport (IFT).	Short-rib thoracic dysplasia 3 with or without polydactyly	11q22.3
ECEL1	Dohrn et al., 2015	ENDOTHELIN-CONVERTING ENZYME-LIKE 1	615065	Zinc metalloproteases.	Neuromuscular disease / Arthrogryposis Multiplex Congenita, Distal Arthrogryposis	2q37.1
FOXP3	Rae et al., 2015	FORKHEAD BOX P3	300292	Development and function of naturally occurring CD4 positive T regulatory cells.	Immunodysregulation, polyendocrinopathy, and enteropathy, X-linked	Xp11.23

FZD6	Shamseldin et al., 2015	FRIZZLED, DROSOPHILA, HOMOLOG OF, 6	603409	Wnt family member, directional cues to align either individual cells within an epithelial sheet or multicellular clusters.	Nail disorder, nonsyndromic congenital, 10, (claw-shaped nails)	8q22.3
GALNT14	Shamseldin et al., 2015	UDP-N-ACETYL-ALPHA-D-GALACTOSAMINE:POLYPEPTIDE N-ACETYLGALACTOSAMINYLTRANSFERASE 14	608225	Glycosyltransferases, enzymes catalyse the first step in the O-glycosylation.	None	2p23.1
GLE1	Ellard et al., 2015	RNA EXPORT MEDIATOR	603371	Regulates gene expression.	Arthrogryposis, Lethal congenital contracture syndrome 1	*9q34.11*
GUSB	Shamseldin et al., 2015	BETA-GLUCURONIDASE	611499	Lysosomal hydrolase (involved in the stepwise degradation of glucuronic acid-containing glycosaminoglycans).	Mucopolysaccharidosis VII	7q11.21
KIF14	Filges et al., 2014	KINESIN FAMILY MEMBER 14	611279	Intracellular transport and cell division.	Meckel syndrome 12, Microcephaly	1q32.1
IFT122	Tsurusaki et al., 2014	INTRAFLAGELLAR TRANSPORT 122	606045	Cell cycle progression, signal transduction, apoptosis, and gene regulation.	Cranioectodermal dysplasia 1 (DEC-1)	3q21.3-q22.1
MUSK	Wilbe et al., 2015	MUSCLE, SKELETAL, RECEPTOR TYROSINE KINASE	601296	Muscle-specific kinase for neuromuscular junction formation.	Fetal akinesia deformation sequence, Myasthenic syndrome, congenital, 9, associated with acetylcholine receptor deficiency	9q31.3
MYOM1	Shamseldin et al., 2015	MYOMESIN 1	603508	Striated muscle.	none	18p11.31
NEB	Shamseldin et al., 2015	NEBULIN	161650	Protein component of the cytoskeletal matrix within skeletal muscle.	Nemaline myopathy 2, autosomal recessive	2q23.3

NOP14	Suzuki et al., 2018	NUCLEOLAR PROTEIN 14	611526	18S rRNA processing and 40S ribosome biogenesis.	None	4p16.3
PIGC	Shamseldin et al., 2015	PHOSPHATIDYLINOSITOL GLYCAN ANCHOR BIOSYNTHESIS CLASS C PROTEIN	601730	Glycosylphosphatidylinositol biosynthesis.	None	1q24.3
RYR1	Ellard et al., 2015, McKie et al., 2014	RYANODINE RECEPTOR 1	180901	Calcium release channel of the sarcoplasmic reticulum.	Central core disease, King-Denborough syndrome, Minicore myopathy with external ophthalmoplegia, Neuromuscular disease, congenital, with uniform type 1 fiber	19q13.2
STIL	Cristofoli et al., 2017	SCL/TAL1-INTERRUPTING LOCUS	181590	Centriole duplication during the cell cycle.	Microcephaly 7, primary, autosomal recessive	1p33
THSD1	Shamseldin et al., 2015	THROMBOSPONDIN TYPE 1 DOMAIN-CONTAINING PROTEIN 1	616821	Adhesive glycoprotein	None	13q14.3
UBN1	Shamseldin et al., 2015	UBINUCLEIN 1	609771	Proliferation.	None	16p13.3

3.4.2 CNVs

Thirteen studies and one meta-analysis (Bagheri et al., 2015) (Table 21), were identified which looked for CNVs in either fetal tissue, parental samples, or both by chromosomal microarray analysis. Three different microarray platforms were used for analysis, either single-nucleotide polymorphism (SNP) array, oligonucleotide (oligo) array or bacterial artificial chromosome (BAC) array.

Six studies reported CNVs in pregnancy loss (Zhang et al., 2009, Viaggi et al., 2013, Levy et al., 2014, Zhang et al., 2016, Donaghue et al., 2017, Zhou et al., 2016), four studies in RPL (Raman et al., 2019, Nagirnaja et al., 2014, Karim et al., 2017, Park and Park, 2019) and three studies with a mixture of both pregnancy loss and RPL (Wang et al., 2017, Warren et al., 2009, Borrell, 2019). Seven of the studies included parental samples and therefore the inheritance of reported CNVs was determined. Six of the studies did not include parental samples, and therefore the inheritance pattern of the CNVs reported in these studies could not be determined.

The pregnancy losses reported were pregnancies of varying gestational age, with the majority of pregnancy losses at less than 20 weeks. In three studies (Raman et al., 2019, Park and Park, 2019, Viaggi et al., 2013), all pregnancy losses tested were less than 12 weeks gestation. Two papers (Borrell, 2019, Park and Park, 2019) also identified pregnancies with developmental abnormalities and used hystero-embryoscopy to allow morphological examination of the fetus *in utero* prior to genetic analysis.

Of the studies which determined the inheritance of the CNVs, there were 30 *de novo*, and 43 inherited CNVs (Levy et al., 2014, Raman et al., 2019, Borrell, 2019, Park and Park, 2019, Wang et al., 2017, Warren et al., 2009). In general, the studies showed a 2.2-13 % detection rate (DR) of pathogenic CNVs (Donaghue et al., 2017, Levy et al., 2014, Wang et al., 2017, Warren et al., 2009, Zhang et al., 2016, Zhang et al., 2009) plus a 0.9 % to 3.3 % DR of variants of unknown significance (VOUS) (Donaghue et al., 2017, Wang et al., 2017, Zhang et al., 2016, Qiao et al., 2016). An additional meta-analysis study (Bagheri et al., 2015) compared the characteristics and contributions of rare and common CNVs from four of the other studies by reclassifying CNVs according to the prevalence of healthy controls using Database of Genomic Variants (Bagheri et al., 2015, Raman et al., 2019, Borrell, 2019, Park and Park, 2019, Viaggi et al., 2013). They concluded that common CNVs were specifically enriched in immunological pathways and rare CNVs were not, although the small number of rare CNVs may have hampered this conclusion. However, both rare and common CNVs could have a role in pregnancy loss, as rare CNVs have a two times higher gene density and contain more genes studied in mouse knockouts and common CNVs contain more genes in biological pathways relevant to pregnancy. The studies that identified VOUS were in accordance with each other and suggested the rate of between 2-3 %.

Of particular interest is to find recurrent CNVs that are associated with pregnancy loss. Maisenbacher et al. (2017) determined the frequency of the 22q11.2 deletion in a large cohort of pregnancy loss samples using a SNP microarray. The 22q11.2 deletion was detected in 15 (0.07%) of 22451 POCs, with an overall incidence of 1/ 1497. They concluded that this was higher than the reported general population prevalence (1/4000- 1/6000). Likewise, Nagirnaja et al. (2014) identified CNV regions on chromosome 5 (5p13.3),

disrupting the *PDZD2* and *GOLPH3* genes. There was significant association with an increased risk of RPL. *PDZD2* and *GOLPH3* are predominately expressed in the placenta, suggesting a functional relevance; however neither, of these genes haves previously been linked to placental function or pregnancy complications (Nagirnaja et al., 2014).

Table 21- Study characteristics from studies looking at other genetic findings in pregnancy loss including recurrent molar pregnancies

Reference	Sample Size	Clinical Information		Array	Results from POC analysis			Quality Score
		Sample	Gestation		CNVs	Size	Inherited vs. De novo	
Warren et al., 2009 *Utah, USA*	30	POC- miscarriage	10-20 weeks	BAC and Oligo	4 CNVs (3 gains, 1 loss)	All <300 kb	4 *de novo*	4
Zhang et al., 2009 *Beijing, China*	58	POC- miscarriage	First trimester	Oligo	5 CNVs (4 gains, 1 loss)	108 - 1460 kb	No parental samples	5
Rajcan-Separovic et al., 2010b *Vancouver, Canada*	17	POC- miscarriage with abnormal morphology	Not given (CRL 4-30 mm)	BAC and Oligo	6 CNVs (in 5 miscarriages) (4 gains, 2 losses)*	<250 Kb	2 maternal, 1 paternal, 2 *de novo,* 1 unknown	4
Viaggi et al., 2013 *Genova, Italy*	40	POC- miscarriage	<12 weeks	Oligo	14 CNVs (9 gains, 5 losses)*	<5 Mb	No parental samples	5
Levy et al., 2014 *New York, USA*	2389	POC- miscarriage	<20 weeks (mean 7.7)	SNP	33 CNVs- 12 clinically significant, 21 VOUS	All <10 Mb	Clinically significant- 5 maternal, 6 paternal, 1 unknown. VOUS- 17 maternal, 3 paternal, 1 unknown	6
Zhang et al., 2016 *Guangzhou, China*	60	POC- Miscarriage	<20 weeks (mean 17.2)	SNP and Oligo	4 pCNVs, 2 VOUS	3 >10 Mb, 3 <10 Mb	No parental samples	6
Zhou et al., 2016 *New Haven, USA*	234	POC- 128 culture failure(CF-POC) and 106 normal karyotype (NK-POC)	6-37 weeks	Oligo	5 pCNVs from CF-POC, 6 pCNVs from NK-POC	10 <10 Mb, 1 >10 Mb (from CF-POC)	No parental samples	9
Donaghue et al., 2017 *London, UK*	1911	POC- miscarriage	8-40 weeks (mean 16)	Oligo	260 CNVs	197 (7.2%) >10 Mb, 25 (0.9%) <10 Mb, 24 (0.9) VOUS, 14 (0.5%) IF	No parental samples	5
Rajcan-Separovic et al., 2010a *Vancouver, Canada*	26	POC- RPL	<20 weeks	Oligo	13 CNVs (6 gains,7 losses)*	<1600 Kb	6 maternal, 7 paternal	7
Robberecht et al., 2012 *Leuven, Belgium*	32	POC- miscarriage with abnormal morphology (1-6 previous losses)	<11 weeks (mean 7.5 weeks)	BAC and SNP	11 CNVs (8 gains, 3 losses)*	<1 Mb	2 de novo, 9 inherited	5

Wang et al., 2017 *Nanjing, China*	535	POC- SA and RM	<13 weeks (mean 9.8)	SNP	31 pCNVs	12 <10Mb, 19 >10Mb	1 maternal, 3 paternal, 22 *de novo,* 5 lost at follow up	9
Nagirnaja et al., 2014 *Tartu, Estonia*	309 female, 249 male	Parental blood samples- RPL (≤3 miscarriages)	<22 weeks	SNP	N/A	N/A	N/A	8
Karim et al., 2017 *Jeddah, Saudi Arabia*	44	Parental samples with RPL (29.5% >5 miscarriages)	79.5% first trimester, 16% second trimester, 4.5% third trimester	Oligo	36 CNVs (25 gains, 11 loses)	all <10 Mb	N/A	7
Bagheri et al., 2015 *Vancouver, Canada*	101	Meta- analysis of studies			24 rare CNVs		14 familial, 1 *de novo,* 9 unknown origin	7

These CNVs were subsequently reviewed and reclassified by Bagheri *et al*. (2015). Quality scores were assessed using a modified Newcastle–Ottawa scale. Array, chromosomal microarray analysis; BAC, bacterial artificial chromosome microarray; CNV, copy number variation; CRL, crown to rump length; IF, incidental finding; Oligo, oligonucleotide microarray; pCNV, pathogenic copy number variant; POC, products of conception; VOUS, variants of unknown significance.

3.4.3 Recurrent molar pregnancies

Eleven studies (Table 22) were identified which evaluated the genetics of diploid and biparental recurrent HM (RHM) pregnancies. One study (Parry et al., 2011) identified biallelic mutations in *C6orf221* in three consanguineous families with familial biparental HMs. Three studies (Abdalla et al., 2012, Brown et al., 2013, Ulker et al., 2013) reported case studies of an individual consanguineous family, two non-consanguineous families and two consanguineous families with RHMs. Autosomal recessive mutations were identified in the *NLRP7* gene and were considered to be responsible for the occurrence of HMs. Deveault et al. (2009) investigated 13 women experiencing RHM, some with a family history of molar pregnancies and 11 *NLRP7* variants were identified. Mutation analysis of the *NLRP7* gene in 35 women experiencing RPL with at least one HM revealed 17 different mutations (Qian et al., 2011). Qian et al. (2011) also suggested that one defective allele in *NLRP7* causes diploid androgenic moles and two defective alleles causes diploid biparental moles.

Two studies (Huang et al., 2013, Kang et al.) investigated cohorts of women to see whether mutations in the *NLRP7* gene could also be responsible for RPL without history of molar pregnancy. Messaed et al. (2011) investigated 135 women with either RPL or at least one HM and sequencing of *NLRP7* exons identified two patients with RPL to have *NLRP7* mutations. Huang et al. (2013) also showed significant association between RPL and *NLRP7* polymorphisms. In contrast, two further studies (Andreasen et al., 2013, Manokhina et al., 2013) identified no disease-causing mutations in *NLRP7* in women with RPL and similarly Aghajanova et al. (2015) found no mutations in *NLRP7*, *NLRP2* or *KHDC3L* (*C6orf221*).

3.4.4 Other genetic causes

Two studies (Seyedhassani et al., 2010a, Seyedhassani et al., 2010b) analysed and sequenced mitochondrial tDNA in 96 women with RPL. Four variants in threonine tRNA and one variant in proline tRNA were observed, but in some cases these were also observed in controls (Seyedhassani et al., 2010a), which calls into question the significance of these findings. Analysis of mitochondrial D-loop sequences showed a higher rate of point mutations in RPL patients than in controls. In total 89 out of 153 variants were only identified in women with RPL and 22 of these mutations were considered to be significant (Seyedhassani et al., 2010b).

X-chromosome inactivation occurs during early embryogenesis and has also been proposed to have an aetiological role in RPL. Skewed X-chromosome inactivation (XCI) status was compared between women with RPL and healthy controls. Extremely skewed XCI (defined as >90 %) was identified in 17.7% of RPL women compared to 1.6 % of extremely skewed XCI in controls (Bagislar et al., 2006).

Six further papers were identified that discussed specific genes and their contribution to pregnancy loss. Each paper (Bendroth-Asmussen et al., 2016, Bhuiyan et al., 2008, Lopez-Carrasco et al., 2013, McKie et al., 2014, Stouffs et al., 2011, Zhang et al., 2016) investigated an individual gene or genes. In a case study of a 30-year-old women with pregnancy loss from glycogen storage disease Type IV (GSD-IV), DNA extracted from placental tissue identified compound heterozygous mutations in *GBE1* (Bendroth-Asmussen et al., 2016).

Another case study, a consanguineous Arabian family with pregnancy losses, stillborn, fetal demise and two live children, had homozygosity mapping. This led to the screening of the

HERG gene in the live children, parents and stillborn. Homozygous nonsense mutations in *HERG* were identified in the child with polymorphic ventricular tachycardia and the same heterozygous mutation in the parents and unaffected child. Amniotic fluid cells from the stillborn child were also homozygous for the same *HERG* mutation (Bhuiyan et al., 2008).

Three rare homozygous *RYR1* variants were identified using genome-wide linkage studies and sequencing of *RYR1* coding exons. Initially, a *RYR1* homozygous nonsense mutation was detected in two fetuses with FADS/ lethal multiple pterygium syndrome (LMPS). The parents were both homozygous for the same mutation. When 66 further probands with FADS/ LMPS phenotype were screened for germline *RYR1* mutations two further potential homozygous mutations were detected (McKie et al., 2014).

In a larger study, 100 couples with at least three unexplained pregnancy losses had *WNT6* mutation analysis performed. *WNT6* has previously been shown to have an important role for stromal cell proliferation during decidualisation in mice. Four novel mutations were identified in the women with RPL but not in the male partners or healthy controls (Zhang et al., 2015), although there was no conclusive evidence for pathogenicity.

Ten aberrations were identified in *MSH4*, *DNMT3L* and *SYCP3* in 23 couples with RPL. Six of these aberrations were predicted to alter the amino acid sequence. All but one of these aberrations was considered a likely SNV. The mutation in the *SYCP3* gene was shown to have a 78 % likelihood of causing a deleterious effect on protein function due to an alteration in the amino acid sequence changing a non-polar isoleucine into a polar threonine (Stouffs et al., 2011). Another study (Lopez-Carrasco et al., 2013) targeted the two spindle checkpoint genes *AURKB* and *SYCP3* in 102 patients with either RPL or spermiogram alterations. One

heterozygous intronic deletion was identified in *SYCP3* with no *in silico* causative indication. Six aberrations were identified in *AURKB,* however a deletion and two nucleotide changes were considered to have no functional alteration or be frequent variants respectively. Three rare missense variants were identified in *AURKB*, with two of these variants found in a couple with pregnancy loss.

Table 22-Characteristics of studies looking at other genetic findings in pregnancy loss including recurrent molar pregnancies.

	Author, Date and Country	Study Objective	Sample, phenotype and gestation	Methods and analysis	Study Outcome	Quality Scores
Recurrent Molar	Deveault et al., 2009 *Montreal, Canada*	Investigate the role of *NLRP7* in recurrent HMs.	Thirteen patients, each with RHM or recurrent reproductive wastage. Five with family history of molar pregnancy, eight with no fetal heart.	Sequencing of *NLRP7* exons.	Eleven *NLRP7* variants in parents with diploid biparental, diploid androgenic, triploid and tetraploid moles.	6
Recurrent Molar	Parry et al., 2011 *Leeds, UK*	Report biallelic mutations of *C6orf221* in three families with FBHM.	Three consanguineous families with FBHM, and 14 further probands.	Homozygosity sampling using genome-wide SNP array, NGS enriching for homozygous regions.	Biallelic mutations in *C6orf221* in three families.	4
Recurrent Molar	Abdalla et al., 2012 *Alexandria, Egypt*	Mutation analysis of *NLRP7* and *C6orf221* genes in women with RHM.	Three women from two families with first trimester losses with repeat HMs.	Cytogenetic analysis; Sequence analysis of *NLRP7* and *C6orf221* genes.	Two autosomal recessive mutations were identified in *NLRP7*.	4
Recurrent Molar	Brown et al., 2013 *Vermont, USA*	Case study of one patient.	One woman with successive molar pregnancies in the first trimester. One complete HM and two partial moles.	Chromosomal microarray analysis of DNA from peripheral blood.	Homozygosity for clinically relevant *NLRP7* mutation. SNP array ruled out deletion.	4
Recurrent Molar	Ulker et al., 2013 *Istanbul, Turkey*	Investigate genetic disposition of families with RHM pregnancies.	Two consanguineous Turkish families with recurrent HM.	Sequencing of *NLRP7* axons and SNP array.	1 family with homozygous *NLRP7* nonsense mutation. Another family with heterozygous deletion affecting the *NLRP2* and *NLRP7* genes.	3
RPL/ Recurrent Molar	Qian et al., 2011 *Zhejiang, China*	*NLRP7* mutation analysis in Chinese patients with recurrent reproductive wastage, including at least one HM.	Thirty-five women with RPL with at least one HM.	Mutation analysis of NLRP7 exons.	Founder mutations in *NLRP7* - diploid androgenic moles in patients with one defective allele, diploid biparental moles in patients with two defective alleles.	9
RPL/ Recurrent Molar	Messaed et al., 2011 *Montreal, Canada*	Investigate the role of *NLRP7* in sporadic moles and other forms of reproductive wastage.	One hundred thirty-five women with three spontaneous abortions or at least one HM.	Sequencing of *NLRP7* exons.	*NLRP7* is responsible for some cases of recurrent spontaneous abortion.	9
RPL/ Recurrent Molar	Andreasen et al., 2013 *Aarhus, Denmark*	Investigate the association between molar pregnancy and RPL regarding changes in the *NLRP7* and *C6orf221/ KHDC3L* genes.	Nineteen women with RPL and one HM, five women with RHM, seven women with HM and with family members with HM.	DNA from maternal blood - NLRP7 and *KHDC3L* were sequenced using PCR.	No unequivocal pathogenic mutations in *NLRP7* or *KHDC3L*; although eight rare non-synonymous variants (NSVs) in *NLRP7* were observed and 3 NSVs in KHDC3L.	7
RPL	Huang et al., 2013 *Tainan, Taiwan*	Do genetic variants of *NLRP2* and *NLRP7* confer susceptibility to idiopathic RPL?	One hundred forty-three women with a least two consecutive pregnancy losses at <12 weeks.	Genomic DNA extracted from lymphocytes and SNPs identified using end-point TaqMan assays from genomic DNA.	A SNP in *NLRP7* showed significant association with idiopathic RPL. A SNP in *NLRP2* showed a marginally significant association.	10
RPL	Manokhina et al., 2013 *Vancouver, Canada*	Elucidate which subpopulations of women with adverse reproductive outcomes should be screened for *NLRP7/ C6orf221*.	Seventeen women with RPL or complete HM.	Peripheral blood DNA sequencing of *NLRP7* and *C6orf221 exons*. Six non-synonymous *NLRP7* variants genotyped in larger clinical groups.	No association between *NLRP7* variants and RPL or partial HM.	8

RPL	Aghajanova et al., 2015 *Houston, USA*	Are mutations in *NLRP2*, *NLRP7* or *KHDC3L* associated with recurrent pregnancy loss or infertility?	Ninety-four women with unexplained infertility and 24 women with RPL.	Sequencing of coding exons of *NLRP7*, *NLRP2* and *KHDC3L* in genomic DNA.	No disease-causing mutations were identified in *NLRP2*, *NLRP7* and *KHDC3L* in either unexplained infertility or RPL.	6
Imprinting/ molar	Docherty et al., 2015 *Southampton, UK*	Report genetic causes in multi-locus imprinting disturbances (MLID) patients, to seek genetic causes of the epigenetic errors in MLID.	Thirty-nine MLID patients and 33 mothers of patients with MLID.	WES and sanger sequencing to confirm NLRP5 variants, target methylation analysis using methylation-specific PCR.	Identification of *NLRP5* variants in five mothers of patients with MLID and hypomethylation of imprinted loci observed in MLID patients exposed to maternal *NLRP5* variants.	6
Mito chon drial DNA	Seyedhassani et al., 2010a *Tehran, Iran*	Analysis of mitochondrial tRNA (Thr) and tRNA (Pro) in women with RPL.	Blood samples from 96 women with at least 3 pregnancy losses at <20 weeks gestation.	Mitochondrial proline and threonine tRNAs were sequenced.	Four mutations in tRNA (Thr) and one mutation in tRNA (Pro) were identified in some women.	5
Mitochond rial DNA	Seyedhassani et al., 2010b *Tehran, Iran*	Examine mitochondrial D-loop deletions and nucleotide alterations in samples taken from RPL women.	Ninert-six blood samples from women with at least 3 pregnancy losses at <20 weeks gestation.	Mitochondrial D-loop DNA sequence analysis.	One hundred fifty-three different mutations in D-loop region. Higher rate of mutations in D-loop in maternal blood of women with RPL in comparison to control group.	10
X-Chromo some inactiva tion	Bagislar et al., 2006 *Ankara, Turkey*	Is there a difference in X-chromosome inactivation patterns in patients with RPL?	Eighty patients with RPL and 160 controls.	Methylation sensitive restriction enzyme digest and PCR. Radioactive labelling and densitometry analyses.	Skewed X chromosome inactivation in 20.9% RPL vs. 6.4% controls, with extreme skewing in 17.7 % RPL vs. 1.6 % controls.	8
GBE1	Bendroth-Asmussen et al., 2016 *Copenhagen, Denmark*	Case report of 30 year old woman with pregnancy loss in the first trimester. Presenting Glycogen storage disease Type IV (GSD-IV) as a cause of early pregnancy loss.	Pregnancy tissue at 10+1 weeks gestation.	Placental histopathology with Periodic acid-Schiff staining and immune-histochemical stains and sequencing of *GBE1*.	Compound heterozygosity for mutations in *GBE1*. GSD-IV can cause early pregnancy loss, and diagnosis can be made on histopathological examination.	2
HERG	Bhuiyan et al., 2008 *Amsterdam, Netherlands (Multi centre)*	Identify causes of RPL and late fetal loss due to arrhythmias in a consanguineous Arabian couple.	Male child with polymorphic ventricular tachycardia	Homozygosity mapping of cardiac ion channel genes.	Homozygous nonsense mutation in *HERG* (consanguineous parents both had the same heterozygous mutation).	3
RYR1	McKie et al., 2014 *Cambridge, UK*	Identify any novel genetic causes of the FADS and multiple pterygium syndrome (MPS) phenotypes.	Two pregnancy tissues affected by FADS/ MPS at 12+6 and 14+0 weeks gestation. Thirty-six families with FADS/LMPS and 30 families with Escobar variant of MPS.	Genome-wide linkage using DNA from fetal material and sequencing of *RYR1* coding exons in probands.	Linkage study identified three rare homozygous *RYR1* variants which were also identified in families with FADS/LMPS phenotypes.	5
WNT6	Zhang et al., 2015 *Jinan, China*	Investigate whether mutations in WNT6 play a role in unexplained RPL.	One hundred couples with at least three pregnancy loss; 100 matched controls.	Mutation analysis of *WNT6* exons.	Four novel mutations in women with unexplained pregnancy loss, none in males or controls. No conclusive evidence for pathogenicity.	8
22q11.2	Maisenbacher et al., 2012 *San Carlos, USA*	Determine the incidence of the 22q11.2 deletion in the pregnancy loss population.	Twenty-two thousand five hundred forty-one POC specimens received for pregnancy loss testing.	Genotypes using SNP arrays across the genome.	Fifteen samples with 22q11.2 deletion (0.07% of the sampled population), 12 from normal karyotypes and 2 from abnormal karyotypes.	6

Maturation Arrest	Stouffs et al., 2011 *Brussels, Belgium*	Gain insight into maturation arrest and the relationship with mutations in genes essential for meiosis and the relevance to RPL.	Forty azoospermic patients and 23 couples with at least two consecutive pregnancy losses.	Genomic DNA sequencing of *DNMT3L*, *SYCP3* and *MSH4* genes.	Five and six aberrations affecting amino acid sequence in azoospermic men and RPL patients respectively. Some aberrations were also identified in controls.	7
Spindle Checkpoint genes	Lopez-Carrasco et al., 2013 *Valencia, Spain*	Investigate involvement of spindle checkpoint genes in patients with chromosomal instability and reproductive problems.	One hundred two patients (43 of which have RPL and 46 with spermiogram alterations).	Genomic DNA sequencing of *AURKB* and *SYCP3* genes.	Three rare heterozygous missense variants in *AURKB*.	10
Epigenetics	Zheng et al., 2013 *Guangzhou, China*	Do assisted reproduction techniques affect DNA methylation of imprinted genes and can aberrant methylation of imprinted genes account for pregnancy loss?	320 pregnancy tissue samples from pregnancy loss after ART or natural conception between 6-26 weeks.	Paternally methylated gene, *H19*, and maternally methylated genes, *LITI* and *SNRPN*, were analysed using bisulfite pyrosequencing of genomic DNA and methylation analysis.	Two regions of hyper-methylation found in *H19*. Mean percentage of methylation higher in pregnancy loss samples. Positive correlation between percentages of methylation of all three genes and pregnancy loss.	11

Quality scores were assessed using a modified Newcastle–Ottawa scale. FBHM, familial biparental hydatidiform mole; HM, hydatidiform mole; RHM, recurrent hydatidiform mole. *C6orf221,* chromosome 6 open reading frame 221; *DNMT3L,* DNA methyltransferase 3-like protein; *GBE1,* glycogen branching enzyme; *H19,* H19, imprinted maternally expressed noncoding transcript; *HERG,* human ether-a-go-go-related gene; *KHDC3L,* KHDC3-like protein, subcortical maternal complex member; *LIT1,* long qt intronic transcript 1; *MSH4,* MutS, E. coli, homolog of, 4; *NLRP2,* NLR family, pyrin domain-containing 2; *NLRP5,* NLR family, pyrin domain-containing 5; *NLRP7,* NLR family, pyrin domain-containing 7; *SNRPN,* small nuclear ribonucleoprotein polypeptide N; *SYCP3,* synaptonemal complex protein 3; *WNT6,* wingless-type MMTV integration site family, member 6.

3.5 Discussion

In this systematic review we have identified 50 papers which investigated genetic contributions other than aneuploidy to pregnancy loss. The studies identify some key areas, including identification of SNVs by WES, identification of CNVs by microarray analysis, and investigation of a group of genes associated with diploid and biparental recurrent molar pregnancies that are linked to pregnancy loss. Other genetic contributions, such as epigenetics and mitochondrial DNA (mtDNA), were also investigated in individual papers. There were also studies reporting sequencing of candidate genes already known to be associated with pregnancy loss with or without structural abnormalities.

We have summarised the current evidence below for each of these categories, and then discuss the implications of these findings both for future studies and also for genetic investigation of couples experiencing RPL.

3.5.1 Whole exome sequencing

Advances in next generation sequencing are vastly improving and enabling a molecular diagnosis for a range of disorders and clinical pathways. As the cost of WES decreases, the technology is becoming more widely used and clinically applicable. This review identified a number of studies (Table 19) over the last 4 years which have used WES to look for as yet unidentified genetic causes of pregnancy loss. The majority of these studies looked at individual patients or couples with RPL, some of which showed ultrasound scan abnormalities during the pregnancy (Bondeson et al., 2017, Cristofoli et al., 2017, Wilbe et al., 2015, Tsurusaki et al., 2014). More recently a small number of studies have been

published studying larger cohorts of patients and exploring possible strategies for genetic investigation of these patients (Ellard et al., 2015, Qiao et al., 2016, Shamseldin et al., 2015). This review included studies where patients suffered multiple pregnancy losses with phenotypic findings in all or some of their pregnancy losses. This included ultrasound scan abnormalities and post-mortem findings, and in some cases, where patients opted for termination of pregnancy. These were thought to be important to include because there could be a range of phenotypic effects caused by a genetic abnormality in a lethal gene, which could include abnormalities and late fetal death in some pregnancies, but earlier pregnancy loss in others.

Bioinformatic filtering is required when studying the whole exome in order to provide a more manageable approach to interpretation of the data. In most of these studies 'trios' of patients were sequenced, and bioinformatic modelling of inheritance patterns was used to limit the number of variants identified. In most cases, patterns of autosomal recessive inheritance (or X-linked recessive in male fetal losses) were modelled to look for variants. As might be expected, very often the couples investigated were consanguineous or possibly from populations isolated geographically. An alternative autozygosity mapping approach was used by Shamseldin et al. to restrict the genes that were analysed by WES (Shamseldin et al., 2013, Shamseldin et al., 2015) and a 'proof of principle' study (Ellard et al., 2015) developed a technique to identify autosomal recessive lethal disorders using WES in couples with RPL.

It is important to note that where autosomal recessive mutations are identified as a cause of pregnancy loss, this will guide counselling and treatment options for the couple as there is a

1:4 recurrence risk in future pregnancies, and prenatal diagnosis or PGD would be available to the couple.

Interestingly, genes that were identified from these WES studies are associated with processes that have an early role in developmental biology and are essential in embryogenesis. Some key processes include centrosome integrity, anti-inflammatory/ immune responses, proliferation and maintenance of epithelial cells, maintenance and development of collagen and muscle tissues and blood coagulation. The majority of WES studies focused on individual families. Therefore the genes detected are limited to preselected cases and it is not possible to group them together for a meta-analysis to ascertain the DRs.

Immune cells present early during pregnancy, especially during implantation where the maternal immune system has to tolerate the implanting embryo. The immune response during implantation is not currently well understood. However, the maternal immunity shifts from cell-mediated immunity to humoral (antibody mediated) immunity to protect the embryo from rejection. Aberrations in several genes, *ALOX15 (Qiao et al., 2016)*, *CR1* (Quintero-Ronderos et al., 2017), *FOXP3* (Rae et al., 2015) and *TLR3* (Filges et al., 2014) were identified and are known to be involved in inflammatory and immune defences. Mutations in these genes could be causing defects resulting in early pregnancy loss because the immune response is rejecting the embryo.

During embryogenesis, cells differentiate and proliferate. Potentially causative mutations were identified in *FLT1* (Quintero-Ronderos et al., 2017), *LIFR* (Quintero-Ronderos et al., 2017) and *UBN1* (Shamseldin et al., 2015) genes involved in cell differentiation and

proliferation. Mutations in the two genes *TRO* and *CHD11* were both identified (Quintero-Ronderos et al., 2017) and are involved in cell adhesion. As cell differentiation, cell proliferation and cell adhesion are an important part of fetal growth during pregnancy, disruption in these genes could cause the pregnancy to fail.

Mutations in genes involved in tissue formation were also identified. In particular, *CDH1* (Quintero-Ronderos et al., 2017) and *FZD6* (Shamseldin et al., 2015) are specifically involved in cell adhesion, *MMP10* and *MMP9* (Quintero-Ronderos et al., 2017) for extracellular remodelling and *MUSK* (Wilbe et al., 2015) and *MYOM1* (Shamseldin et al., 2015) for formation of neuromuscular junctions and striated muscle.

During pregnancy, blood passes through the placenta for the exchange of gases, nutrients, electrolytes and waste products between the mother and fetus. Mutations in three genes, *F5*, *FGA* and *THBD* (Quintero-Ronderos et al., 2017), were identified. These are involved in the coagulation pathway. The flow of blood is necessary for the fetus to grow and any disruption causing the blood to clot could result in loss of the pregnancy.

In summary, WES of POC or fetal DNA and parental DNA is a promising method to identify variants in genes which might be responsible for RPL and/ or fetal abnormalities. Where aberrations are inherited from the parents, a genetic diagnosis may provide invaluable information for preimplantation screening or prenatal diagnosis in future pregnancies. However, studies with larger unbiased cohorts are needed to conclusively determine detection rates and clinical utility of WES in this group of patients.

3.5.2 Chromosomal microarray analysis

In some cases, CNVs either as gains or losses may be responsible for pregnancy loss of a fetus with an apparently normal karyotype. CNVs, both rare and common, may be impacting pregnancy-related genes or pathways, resulting in pregnancy loss. These may involve single genes or clusters of genes which are deleted, duplicated or disrupted.

Studies identified by our systematic review are summarised in Table 21. Due to the diverse approaches taken, the studies are difficult to compare collectively. Cohorts reported sporadic pregnancy loss and RPL, different gestations and different methods of analysis. Some studies (Bagheri et al., 2015, Raman et al., 2019, Borrell, 2019, Warren et al., 2009, Levy et al., 2014, Park and Park, 2019, Wang et al., 2017) analysed both fetal tissue and parental DNA concurrently (i.e. a trio) to identify whether CNVs were *de novo* or inherited. This is important in assessing both the likely pathogenicity of the finding and the associated recurrence risk. Where the CNV is also detected in a parent, it is less likely to be causative of a pregnancy loss in isolation. It is possible that inherited CNVs could still cause RPL where the CNV co-occurs with an autosomal recessive gene mutation (SNV) on the other allele or where genes present within the CNV are relevant to genomic imprinting or embryonic/ placental growth (Raman et al., 2019, Borrell, 2019).

Relatively little is known about the genes and pathways involved in pregnancy loss, and therefore many CNVs identified will be classed as uncertain clinical significance. One study analysed CNVs in parents experiencing idiopathic RPL using functional enrichment analysis, identifying biological pathways that were significantly over-represented, such as antigen binding and immune signalling (Karim et al., 2017, Nagirnaja et al., 2014). Enrichment was

identified in genes associated with immunoregulatory interactions at the feto-maternal interface and impaired immune signalling (Nagirnaja et al., 2014).

Identification of pregnancies with developmental abnormalities using hystero-embryoscopy enables genetic abnormalities to be compared with developmental abnormalities and growth disorganisation of the embryo. CNVs identified where there is a developmental abnormality present are more likely to indicate genes important in early development. In addition to evaluating a genetic cause for pregnancy loss, such studies can provide an opportunity to identify and evaluate the function of the genes. Where variants are identified in genes, through analysis of an enriched cohort, such as this, it is easier to interpret their clinical significance.

Several studies explored the possibility of uniparental disomy (UPD) and looked for regions of loss of heterozygosity in euploid embryos (Levy et al., 2014, Park and Park, 2019, Wang et al., 2017). The pathological relevance of UPD is difficult to evaluate as not all platforms are capable of detecting UPD (e.g. Oligo BAC array) and therefore are difficult to compare. Pregnancy loss could be due to UPD resulting in unmasking of an underlying lethal recessive disease gene(s) or imprinted genes.

CNVs were identified in the highly imprinted region 11p15.5. This region is abundant with imprinted genes and has an important role in the maternal-fetal exchange. Aberrant methylation or duplication of imprinted genes in this region could cause pregnancy loss (Zhang et al., 2016).

3.5.3 Recurrent molar pregnancies

Although the majority of HMs are sporadic, a small minority are recurrent and/ or familial. A number of studies looked at the role of genes including *NLRP7*, C*6orf221* (*KHDC3L*) and *NLRP2* in pregnancy loss manifesting as recurrent molar pregnancy. In the cases reviewed, the HMs are euploid, and are instead caused by autosomal recessive mutations in genes that code for the cell machinery that labels the parental origin of the two sets of chromosomes.

It is thought that *NLRP7* and *C6orf221* are components of an oocyte complex that forms during oogenesis and determines the epigenetic status of the oocyte genome by inactivating genes. It is likely that variants in *NLRP7* cause Hydatidiform moles by impairing the normal imprinting process causing maternal genes to be expressed when they should not be.

Studies have explored the role of *NLRP2*, *NLRP5, NLRP7* and *C6orf221* in other forms of pregnancy loss such as partial moles, RPL, stillbirth, infertility and multi-locus imprinting disturbance (MLID) (Aghajanova et al., 2015, Andreasen et al., 2013, Huang et al., 2013, Manokhina et al., 2013, Kang et al., Docherty et al., 2015). These have shown conflicting results, many showing no evidence of *NLRP7*, *NLRP2* and *C6orf221* mutations in women with RPL (Aghajanova et al., 2015, Andreasen et al., 2013, Manokhina et al., 2013).

Evidence from several papers suggests that genes involved in oocyte development, maturation and epigenetic reprogramming are likely to be important in a subset of pregnancy losses. One of the most studied epigenetic modifications is DNA methylation. DNA methylation is implicated in the regulation of imprinting and expression of imprinted genes are thought to be important for the development and physiology of the placenta (Togneri et al., 2019). Aberrant DNA methylation of several imprinted loci (*H19*, *LIT1* and

SNRPN) was demonstrated in pregnancy losses, with increasing methylation of these genes showing a positive correlation with pregnancy loss. It is possible that inappropriate DNA methylation may either be a contributing factor or consequence of the defect that led to pregnancy loss (Zheng et al., 2013). It also remains to be investigated as to whether there are wider epigenetic defects at other loci. Zheng et al. (2013) propose a multifactorial threshold model for pregnancy loss where additional genetic and environmental factors may also play a role.

3.5.4 Other genetic causes

Mitochondria have been hypothesised to have an important role in development. They predominantly regulate the production of ATP, used to regulate cellular metabolism. Processes such as cell proliferation and development require high energy giving the mitochondria an important role during pregnancy. Seyedhssani et al. (Seyedhassani et al., 2010a, Seyedhassani et al., 2010b) have identified mutations in mtDNA in women with RPL. Furthermore a significant number of mutations were identified in the D-loop of mtDNA. The D-loop contains essential elements for mtDNA transcription and disruption could affect the transcription or translation of mtDNA, in turn compromising embryonic development or causing pregnancy loss.

It is hypothesised that skewed XCI could be involved in the pathogenesis of RPL. Bagislar et al. (2006) and colleagues demonstrated extremely skewed XCI in 17.7 % of patients with RPL. It is suggested that skewed XCI could expose X-linked variants that are lethal in the hemizygous state. In addition, a more recent review (Sui et al., 2015) including 12 case-control studies on skewed XCI with or without RPL. In patients with RPL, skewed XCI was

significantly higher, although the significance drops with fewer losses and for less extreme skewing. Although the association between RPL and skewed XCI is unclear, two mechanisms have been proposed. Firstly, if a female carrier with a recessive lethal X-linked genetic mutation and skewed XCI has a male fetus who inherits the X-linked genetic mutation, it could lead to pregnancy loss. Secondly, an X-linked genetic mutation could cause follicular atresia and an increase in aneuploid embryos resulting in pregnancy loss (Sui et al., 2015).

Six papers (Bendroth-Asmussen et al., 2016, McKie et al., 2014, Stouffs et al., 2011, Zhang et al., 2016, Bhuiyan et al., 2008, Lopez-Carrasco et al., 2013) describe targeted sequence analysis of specific candidate genes (*GBE1*, *RYR1*, *WNT6*, *DNMT3L*, SYCP3, *MSH4, HERG* and AURKB) in either an individual case of pregnancy loss (Bendroth-Asmussen et al., 2016, Bhuiyan et al., 2008) or in patient cohorts (McKie et al., 2014, Stouffs et al., 2011, Zhang et al., 2016, Lopez-Carrasco et al., 2013). This targeting was informed by factors including histopathological examination of placental tissue observed in fetal arrhythmia, scan findings and functional prediction of gene pathways.

3.5.5 Limitations of current evidence

This review was completed in a systematic manner by two independent reviewers making it reproducible. The limitation of this study however is the quality of the studies published to date. Each study was scored according to our modified Newcastle-Ottawa scale (Supplementary Table 4) with a few of the studies being of poor quality and scoring as little as three or four in our scale.

The most common limitations in these studies related to the small size of the studied cohorts, with several focusing on a single family, and many of the studies lacking information on control populations or statistical analysis. Work on small groups, and in particular a single family, may detect genetic abnormalities that have occurred in isolation or are very rare. In many cases this results in identification of variants in unique candidate genes with no definitive causal effect. Therefore larger cohorts are needed to replicate these findings and to determine how relevant these findings are to other couples with RPL.

There was also limited availability of functional data in many of the studies. A few studies supplemented their cases with information on scan abnormalities or post-mortem abnormalities detected in cases of losses and hystero-embryoscopy to correlate genetic findings with findings in the embryo. The studies were also difficult to compare and collate as there was multiple variations in the cohorts studied and the methods of analysis.

3.6 Conclusion

It is evident that there are many genetic and environmental factors that result in a successful pregnancy and a disruption in any of these could contribute to pregnancy loss.

From the genetic perspective this includes both clearly pathogenic genetic causes such as sporadic aneuploidy and translocations and other potential genetic causes such as smaller CNVs and mutations in genes important in early fetal development. In addition, there are likely to be complex genetic contributions such as multi-factorial inheritance, changes in methylation (epigenetics), and mitochondrial function, which could be contributing to pregnancy loss. These more complex genetic mechanisms may be influenced by

environmental factors such as diet, medication, pollutants and lifestyle which could provide a cumulative effect resulting in pregnancy loss.

The papers we have identified have demonstrated that monogenic aetiologies could contribute to a proportion of pregnancy losses. However as most studies have been carried out in highly selected families or small cohorts, additional studies are required to further assess if this technology is generalisable to more couples experiencing RPL.

It is plausible that cases of pregnancy loss (particularly in RPL) may have causative mutations not detectable with routine cytogenetic analysis or fetal scans, but are detectable by WES. Although WES is not currently recommended for routine diagnostic use for pregnancy losses, identification of genes associated with pregnancy loss will be of significant individual patient impact with respect to treatment and availability of PGD. If monogenetic etiologies of RPL and the overall prevalence of monogenetic causes of pregnancy loss are better elucidated through larger, well-designed studies, the identification of non-aneuploid causes of RPL could be of significant patient impact.

Knowledge of specific genes that contribute to pregnancy loss could also be of importance in understanding the biological pathways that can cause pregnancy loss. However, much larger and more comparable cohort studies are required in all of these areas to determine causality of candidate genes and to dissect out these effects, as at present many of these findings are of uncertain clinical significance. Functional analysis such as embryoscopy studies and *in vivo* animal modelling may assist in further assessment of the mutation effect on early embryonic development.

RPL is a complex problem influenced by many different aetiologies. Currently, with the exception of aneuploidy and other chromosomal abnormalities, routine investigation for the genetic contributions, causing pregnancy loss is limited. With increased knowledge of additional non-aneuploid contributions to RPL, additional genetic testing recommendations may be made in the future to couples experiencing RPL. These would have implications for diagnosis and recurrence risks.

Chapter 4: Whole Exome Sequencing

Chapter 4- Whole Exome Sequencing

Whole exome sequencing (WES) is a method to target the specific 'protein-coding' regions of DNA known as the exome. The exome contributes to approximately 1% of the human genome, but contains approximately 85 % of the genetic mutations associated with disease (Choi et al., 2009). With advances in sequencing technology, the opportunity to detect genetic sequence variation and to characterise genetic mutations causing monogenic disease is increasing. WES has been used to evaluate developmental disorders. For example, the Deciphering Developmental Disorders (DDD) Study (Wright et al., 2015) and the Prenatal Assessment of Genomes and Exomes (PAGE) study (Lord et al., 2019), have shown that exome sequencing increases the diagnostic yield (Drury et al., 2015). However, very few WES studies have been reported for miscarriage, and it is possible that genes which are lethal during early development could be contributing to recurrent miscarriage.

A few published studies have used WES in patient samples with recurrent pregnancy loss. These studies tend to have very small cohorts and often focus on fetal abnormalities. Several studies, identified in the systematic review in Chapter Three, identified families where WES in trios had been used to identify variants in both the fetus and parents to identify the cause of the fetal loss and determine the inheritance.

A large study from the Wellcome Trust and Francis Crick Institute has looked at genes essential for mouse embryonic development (Wilson et al., 2016a). Using 31 lethal and 11 sub-viable, novel gene knock-out embryos, 42 genes found were considered to be lethal or sub-viable and 398 different mammalian phenotypes were recorded. This study could have a

huge impact in deciphering essential genes for human development and human embryonic lethal genes.

WES can potentially detect some large structural variants, and small insertions/deletions, single nucleotide variants and copy number changes within the exome. However, very few WES studies have reported analysis in early pregnancy loss or lethal genes which could contribute to recurrent miscarriage. The few studies using WES to look for genetic aberrations in recurrent pregnancy loss also tend to only represent small patient cohorts. The ability to recognise and detect genetic mutations may have implications for routine genetic testing and clinical practice, especially when a pathogenic aberration is identified that can be reliably detected in future pregnancies. However this group of patients are potentially challenging in comparison to structurally abnormal pregnancies. In most cases it will be difficult to enrich the cohort for pregnancies that are more likely to have a genetic abnormality and it is difficult to functionally assess the contribution of any variants that are identified.

More widespread genetic analysis of euploid miscarriages may provide an opportunity to identify genes, essential in early human development and where a lack of function leads to miscarriage. This could identify in some couples why apparently normal conceptuses miscarry and would provide important information to that couple for future pregnancy planning. Our systematic review reflects that larger studies are needed to conclusively determine the clinical utility of WES in this group of patients. We aim to identify the best WES method for us and to run whole exome sequencing on selected samples from our

recurrent miscarriage patient cohort, focusing on trios of maternal, paternal and fetal DNA samples.

4.1 Pilot Study

In order to establish the best method of whole exome sequencing in our patient cohort an initial pilot study was conducted to assess three different sequencing methods including cost, ease, sequence output, coverage and variant detection.

4.1.1 Methods

As per ethical approval, maternal and paternal whole blood samples were collected as part of the Tommy's miscarriage recruitment from patients with recurrent pregnancy loss. Where available, POC samples were collected from the family. DNA was extracted from both whole blood and tissue using the QIAsymphony DSP DNA Midi kit (QIAGEN) according to manufactures guidelines on a QIAsymphony robot.

Three trios (maternal, paternal and POC) from families 82, 141 and 154 were sequenced using three different WES methods from external providers, UoB-TruSeq, TE-Agilent-SS and BGI-AgilentV6. The coverage of each WES method assessed using the FETALCES-001: Exome Comparison gene panel in Congenica consisting of 1542 genes (Appendix 3).

Table 23- Three different Whole Exome Sequencing Providers.

	Sequencing Provider	NGS Platform	Library Prep	Sequencing Coverage	Cost Per sample
UoB-TruSeq	University of Birmingham Sequencing Service	Illumina NextSeq500	TruSeq exome kit (Illumina)	30x	£664
TE-Agilent-SS	Theragen Etex	Illumina NovaSeq 6000	Agilent Sureselect V6 exome capture kit	100x	£250
BGI-AgilentV6	BGI	BGISEQ-500	Agilent V6 kit	100x	£261

FASTQ files were downloaded from the respective sources and the Next generation sequencing (NGS) data and family pedigree from each WES method was uploaded via a secure FTP site to Congenica (https://www.congenica.com/about-us/), genetic variant interpretation software to be processed (previously Sapientia). Variant calling bioinformatics analysis was completed in Congenica using family information for inheritance filtering.

Variants were filtered using the FETALCES-001: Exome Comparison gene panel which consists of genes selected from Fetal Clinical exome derived from the PAGE study (Lord et al., 2019). This panel of genes was selected to assess the coverage of each WES method.

4.1.2 Results

Three WES family trios from Families 82, 141 and 154 were assessed. The data from all three trios was successfully downloaded however only the data from UoB-TruSeq and TE-Agilent-SS was successfully uploaded into Congenica. The data from BGI-AgilentV6 was not able to be processed due to the data being generated using BGI's sequencer. This has a slightly different FASTQ format and is not compatible with Congenica.

4.1.2.1 Coverage

The WES coverage for each trio was assessed at >20x coverage for two WES methods, UoB-TruSeq and TE-Agilent-SS using the FETALCES-001: Exome Comparison gene panel in Congenica. BGI-AgilentV6 coverage was not assessed as the data could not be uploaded to Congenica. The coverage was assessed using the coverage of each individual gene per patient on the gene panel and taking an average of these. The mean WES coverage for UoB-TruSeq at >20x coverage was 97.17%, 96.94% and 98.43% for families 82, 141 and 154 respectively. The mean WES coverage for TE-Agilent-SS at >20x coverage was 98.95%, 99.24% and 99.25% for families 82, 141 and 154 respectively.

Figure 24 below shows a summary of the coverage. The coverage of individual genes can be found in Appendix 4-6, where at least one sequencing method was fewer than 90% coverage at ≥20x coverage.

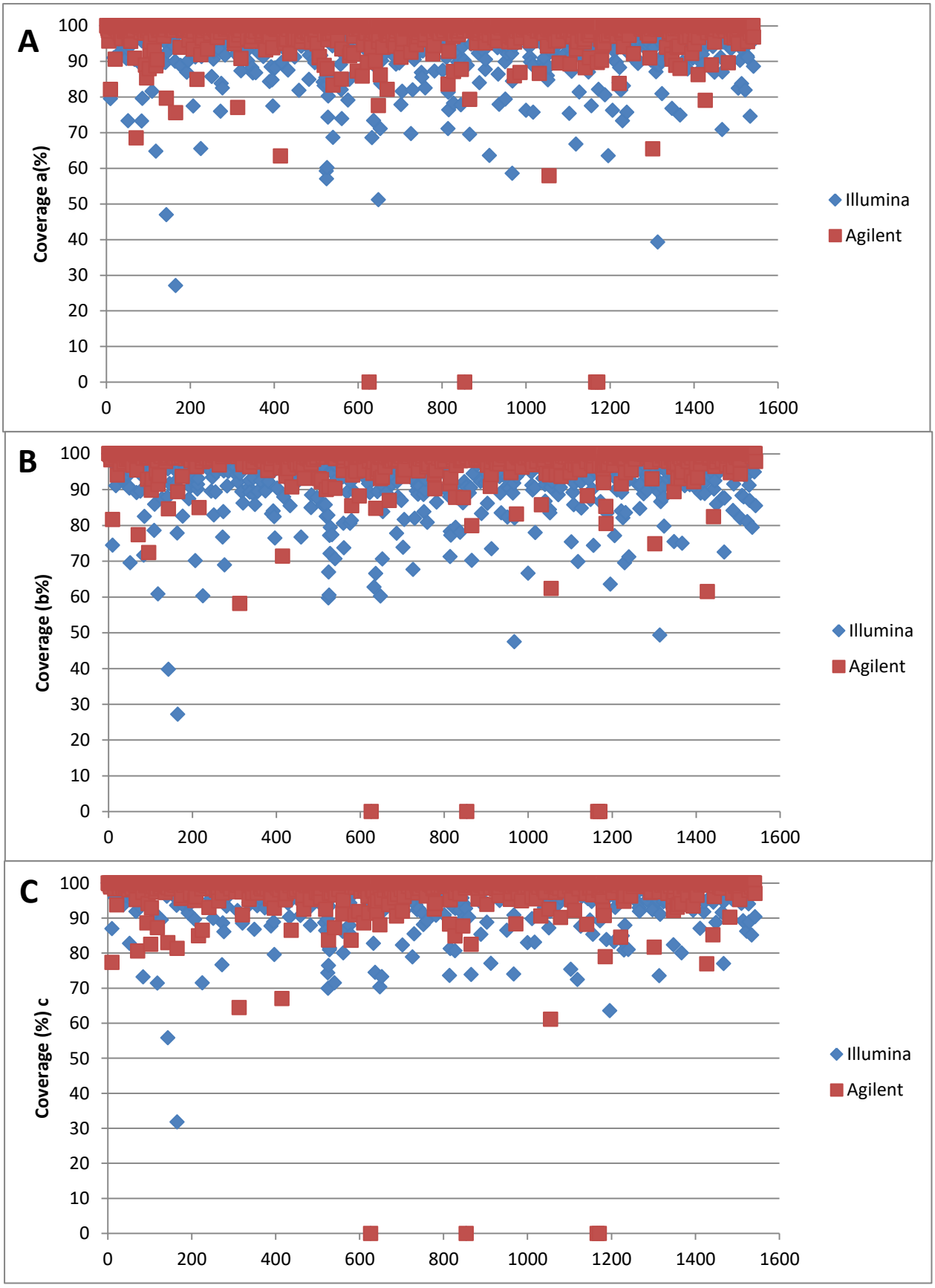


Figure 24-Gene Coverage comparisons for WES

Comparison of UoB-TruSeq and TE-Agilent-SS for A) Family 82, B) Family 141 and C) Family 154. Coverage has been calculated using the FETALCES-001: Exome Comparison gene panel in Congenica.

4.1.2.2 Variants

Variants were identified in UoB-TruSeq and TE-Agilent-SS sequencing data using Congenica from the FETALCES-001: Exome Comparison gene panel (Table 24). Unfortunately the data from BGI-AgilentV6 was not able to be processed by Congenica so this cannot be included in the comparison between the variants. The variants described below were identified following inheritance filtering within each trio. Variants have not been confirmed by Sanger sequencing.

Variants that were considered artefactual were excluded. Artefacts can be introduced due to poor quality DNA preparations, structural variation, sequence context or can be sequencing method specific or sequencing dependent, e.g. repetitive regions. A variant is usually felt to be artefactual if it is seen multiple times in the patient cohort or sequencing run, if it is at low frequency in population control databases (GnomAD), often in repetitive or GC rich regions, and often present as *de novo fp* flagged, skewed read splits and messy sequence alignments. Analysis filters can be applied in Congenica to flag artefactual variants.

For Family 82: Of 24 variants filtered from UoB-TruSeq sequencing data, 23 were excluded as artefacts, and of four variants filtered from TE-Agilent-SS sequencing three were excluded as artefacts. Both Illumina and Agilent sequencing data identified the same variant.

For Family 141: Of 25 variants filtered from UoB-TruSeq sequencing data, seven were excluded as artefacts, and of seven variants filtered from TE-Agilent-SS sequencing none were excluded as artefacts. Illumina and Agilent sequencing data identified six of the same

variants. A variant in *CDT1* was only identified in Agilent sequencing data and a variant in *PRG4* was only identified in Illumina sequencing data.

For Family 154: Of 22 variants filtered from UoB-TruSeq sequencing data, seven were excluded as artefacts, and of 11 variants filtered from TE-Agilent-SS sequencing six were excluded as artefacts. Illumina and Agilent sequencing identified five of the same variants. A variant in *CUX2* was only identified in Agilent sequencing data and two variants were identified in *HYDIN* in Agilent sequencing data were in Illumina sequencing data only one variant was identified. Two variants were identified in *PKD1* in Illumina sequencing data where in Agilent sequencing data only one variant was identified.

Table 24- Variants and Artefacts

Number of artefacts and variants in Illumina and Agilent sequencing. Filtering according to gene panel and inheritance in trios.

	UoB-TruSeq		**TE-Agilent-SS**	
	Artefacts	**Variants**	**Artefacts**	**Variants**
Family 82	23	1	3	1
Family 141	18	7	0	7
Family 154	15	7	5	6

4.1.3 Discussion

An initial pilot study to identify the most appropriate sequencing provider of whole exome sequencing for this study was conducted using three family trios. The sequencing methods included were UoB-TruSeq, TE-Agilent-SS and BGI-AgilentV6.

The pilot study was used to assess the three different methods for delivery, quality control, data receipt, upload of data to Congenica, sequencing depth and number of artefacts (Table 25).

Table 25- Assessments of Pilot study.

	UoB-TruSeq	TE-Agilent-SS	BGI-AgilentV6
Delivery	Delivered in person	Pre-paid FedEx account	Courier service
Quality control	Passed	Passed	Slightly or degraded samples
Sequencing	Successful	Successful	Successful using Kappa kits
Data receipt	BaseSpace https://emea.illumina.com/products/by-type/informatics-products/basespace-sequence-hub.html	Theragen Etex- File Transfer Protocol (FTP) server ftp://ftp.theragenetex.com/	BGI online https://www.bgionline.com/
Data upload to Congenica	Successful	Successful	Unsuccessful
Sequencing depth	97.51 % (average)	98.87 % (average)	Failed
Artefacts	18.6 (average)	2.6 (average)	Failed

The pilot study was used to identify the best method of WES for our patient cohort and we assessed three different sequencing methods based on cost, ease, coverage and variant detection. As the data from BGI-AgilentV6 was unable to be processed in Congenica this was excluded from the comparisons.

In terms of cost per sample, UoB-TruSeq was by far the most expensive. This is likely because it was completed by the University of Birmingham and not a competitive company. However, extra costs were not needed to transport the samples, which have not been considered on the cost per sample. Transporting the samples to the University of Birmingham was easier than arranging a courier for the other samples. Sending the biological samples outside the country was more difficult because they had to go through customs.

TE-Agilent-SS sequencing data had the better coverage than Illumina for all three probands. UoB-TruSeq also had more artefacts then TE-Agilent-SS, however the same variants were detected in both UoB-TruSeq and TE-Agilent-SS sequencing data. UoB-TruSeq actually detected one extra variant in family 154 and there were a few differences between TE-Agilent-SS and UoB-TruSeq variants in families 141 and 154.

In conclusion, although the UoB-TruSeq sequencing had a slight lower coverage for these trios, there were more artefacts and the sequencing was more expensive, the same variants were identified in both UoB-TruSeq and TE-Agilent-SS, and with the ease of transferring the samples to the University of Birmingham Sequencing Service, it was concluded to complete an additional 4 trios using UoB-TruSeq sequencing.

4.2 Whole Exome Sequencing of Trios

Having identified that UoB-TruSeq was the preferred sequencing service for our study; four additional trios were sequenced using the University of Birmingham Sequencing Service.

4.2.1 Methods

4.2.1.1 Trio sample selection

A total of seven trios, collected as part of the Tommy's miscarriage study from patients with recurrent pregnancy loss, were selected for WES. The selection criteria for these seven cases were, a known euploid pregnancy loss, DNA from the entire trio (Maternal, paternal and POC DNA) and consent from the family to send the DNA to an external laboratory. Therefore, the seven trios unintentionally were from seven male pregnancy losses. This was coincidental. The trios consisted of DNA from pregnancy tissue and maternal and paternal whole blood. DNA was extracted from both whole blood and tissue using the QIAsymphony DSP DNA Midi kit (QIAGEN) according to manufactures guidelines on a QIAsymphony robot.

4.2.1.2 Library Preparation and sequencing

At least 100 ng of DNA was submitted to the University of Birmingham Sequencing Service where the WES service was provided. Whole exome Sequencing was completed by the University of Birmingham Sequencing Service using TruSeq Exome Library Prep Kit (Illumina) as per manufacturer's instructions. In brief:

4.2.1.2.1 DNA Fragmentation

100 ng of genomic DNA was fragmented into 150 bp fragments using Covaris shearing. This produces dsDNA fragments of 150 bp with a 3’ or 5’ overhang.

4.2.1.2.2 Repair Ends and Select Library Size

The 3’ or 5’ overhang was removed from the sheared DNA using End Repair Mix containing an exonuclease activity. The fragment length was then optimised using sample purification beads.

4.2.1.2.3 Adenylate 3’ End and Ligate Adapters

To prevent the DNA fragments ligating together and forming a chimera, an ‘A’ nucleotide was added to the 3’ end of the blunt fragment. This also provided a complementary overhang for the ‘T’ nucleotide for ligating the adapter to the DNA fragment. The indexing ligating adapter allows for hybridisation onto the flow cell.

4.2.1.2.4 Enrich DNA Fragments

DNA with ligated adapters was selectively enriched using PCR to amplify the DNA in the library. The PCR Primer Cocktail anneals to the end of the ligating adaptor to minimise the number of PCR cycles which prevents skewing the library. After the DNA has been enriched the libraries are quantified using Qubit dsDNA HS Assay Kit and the library size distribution checked on an Agilent Technologies 2100 Bioanalyser.

4.2.1.2.5 Pool Libraries and hybridise probes

100 ng of each sample with a unique index was combined into a single pool with coding exome oligos to bind targeted DNA regions with capture probes. Streptavidin Magnetic beads select the capture probes which hybridise to the region of interest and the non-specific binding was washed away from the beads. This wash was completed twice to ensure a high specificity of the captured regions.

4.2.1.2.6 Clean up and amplification of enriched library

The capture library was cleaned using Purification beads and the enriched library amplified using an eight cycle PCR program. A clean up was completed after this PCR program to purify the enriched library. The enriched library was quantified using Qubit dsDNA HS Assay Kit and the library quality assessed using an Agilent Technologies 2100 Bioanalyser using a High Sensitivity DNA chip.

4.2.1.2.7 Denature Library

To denature the library, 0.2 N NaOH was added to the library. After incubation, an equal amount of 200 mM Tris-HC1 was added to the denatured libraries to ensure the NaOH was fully hydrolysed.

4.2.1.2.8 Library Dilution

The denatured library was diluted to a 4 nM loading concentration using HT1 reagent and a 1% PhiX spike in was added as a sequencing control.

4.2.1.2.9 Sequencing

The library was sequenced using a NextSeq500 using a v2.5 150 cycles high output (75 paired end) flow cell.

4.2.1.3 Variant interpretation

FASTQ files were downloaded from the respective sources and the Next generation sequencing (NGS) data and family pedigree from each WES method was uploaded via a secure FTP site to Congenica (https://www.congenica.com/about-us/), genetic variant interpretation software to be processed. Variant calling bioinformatics analysis was completed in Congenica using family information.

Variants were filtered using the FETALCES-001: Exome Comparison gene panel which consists of genes selected from Fetal Clinical exome derived from the PAGE study (Lord et al., 2019). This panel of genes was selected because they have been shown to be developmentally and prenatally relevant, and therefore the first tier of genes that we felt should be examined. Variant inheritance was filtered using autosomal dominant, homozygous autosomal recessive, compound heterozygous autosomal recessive, X-linked dominant or X-linked recessive inheritance filters. Allele population frequency databases were set to ≤ 0.01 and the read depth was set to ≥ 20. The quality was assessed by filtering out DRAGENHardINDEL and DRAGENHardSNP QC filters (Figure 25). Variants were annotated based on the ACMG guidelines (Richards et al., 2015) (Table 26).

Once analysis had been carried out using the gene panel, the data was also opened up to look at variants across the whole exome for all seven trios (including those that were

sequenced to identify the preferred sequencing method). Variant inheritance was filtered using *de novo* and autosomal dominant, autosomal recessive, autosomal compound heterozygous, X-linked dominant and X-linked recessive. The allele population frequency data bases were set to ≤0.0005 and the read depth was set to ≥20. Variants were excluded if they were artefactual, selected based on whether they had a potential phenotypic link, and then classified using the ACMG classification.

Copy number variants and complex rearrangements were not detected by the filtering and so will be missed by this analysis.

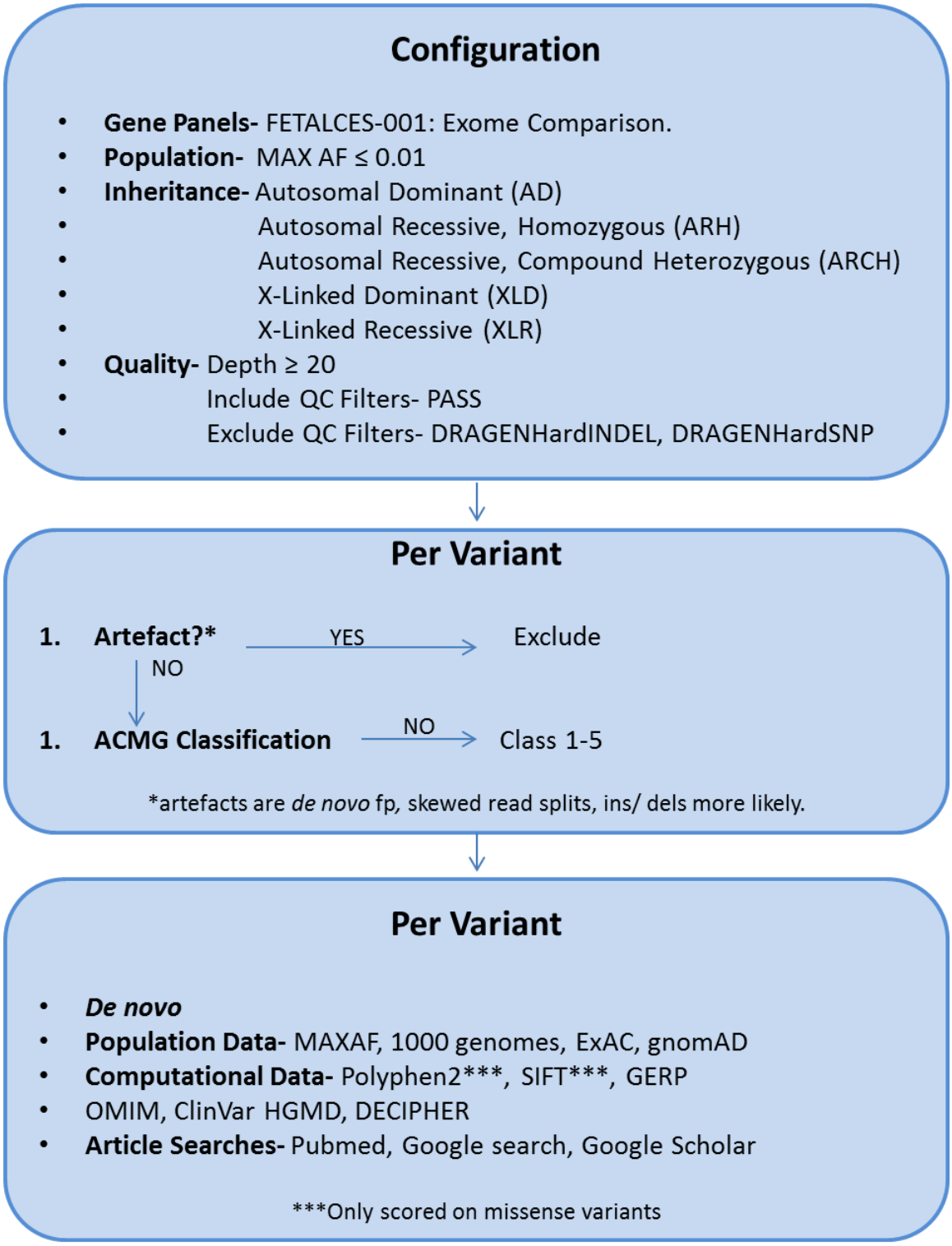


Figure 25- Flow chart summary for analysis of trios using FETALCES-001: Exome Comparison gene panel in Congenica.

Table 26- ACMG evidence framework

Table adapted from Richards et al. (2015). Full description of criteria can be found in Richards et al. (2015)

BA, benign alone; BS, benign strong; BP, benign supporting; PM, pathogenic moderate; PP, pathogenic supporting; PS, pathogenic strong; PVS, pathogenic very strong.

	Benign			Pathogenic			
	Stand Alone	**Strong**	**Supporting**	**Supporting**	**Moderate**	**Strong**	**Very Strong**
Population Data	**BA1**- Allele frequency is >5% in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium.	**BS1**- Allele frequency is greater than expected for disorder. BS2- Allele frequency is greater than expected for disorder.			**PM2**- Absent from controls (or at extremely low frequency if recessive) in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium.	**PS4**- The prevalence of the variant in affected individuals is significantly increased compared with the prevalence in controls	
Computational and predictive data			**BP1**-Missense variant in a gene for which primarily truncating variants are known to cause disease. **BP3**-In-frame deletions/insertions in a repetitive region without a known function. **BP4**-Multiple lines of computational evidence suggest no impact on gene or gene product (conservation, evolutionary, splicing impact, etc.). **BP7**-A synonymous (silent) variant for which splicing prediction algorithms predict no impact to the splice consensus sequence nor the creation of a new splice site AND the nucleotide is not highly conserved.	**PP3**-Multiple lines of computational evidence support a deleterious effect on the gene or gene product (conservation, evolutionary, splicing impact, etc.).	**PM4**- Protein length changes as a result of in-frame deletions/insertions in a non-repeat region or stop-loss variants. **PM5**- Novel missense change at an amino acid residue where a different missense change determined to be pathogenic has been seen before.	**PS1**- Same amino acid change as a previously established pathogenic variant regardless of nucleotide change.	**PVS1**- Null variant (nonsense, frameshift, canonical ±1 or 2 splice sites, initiation codon, single or multiexon deletion) in a gene where LOF is a known mechanism of disease.
Functional data		**BS3**-Well-established in vitro or in vivo functional studies show no damaging effect on protein function or splicing.		**PP2**-Missense variant in a gene that has a low rate of benign missense variation and in which missense variants are a common mechanism of disease.	**PM1**-Located in a mutational hot spot and/or critical and well-established functional domain (e.g., active site of an enzyme) without benign variation.	**PS3**-Well-established in vitro or in vivo functional studies supportive of a damaging effect on the gene or gene product.	

Segregation data		**BS4**-Lack of segregation in affected members of a family.		**PP1**-Cosegregation with disease in multiple affected family members in a gene definitively known to cause the disease.			
De novo data					**PM6**- Assumed de novo, but without confirmation of paternity and maternity.	**PS2**- De novo (both maternity and paternity confirmed) in a patient with the disease and no family history.	
Allelic data			**BP2**-Observed in trans with a pathogenic variant for a fully penetrant dominant gene/disorder or observed in cis with a pathogenic variant in any inheritance pattern.		**PM3-**For recessive disorders, detected in trans with a pathogenic variant.		
Other database data			**BP6**-Reputable source recently reports variant as benign, but the evidence is not available to the laboratory to perform an independent evaluation		**PP5**-Reputable source recently reports variant as pathogenic, but the evidence is not available to the laboratory to perform an independent evaluation.		
Other data			**BP5**-Variant found in a case with an alternate molecular basis for disease.		**PP4**- Patient's phenotype or family history is highly specific for a disease with a single genetic etiology.		

4.2.2 Results

Seven trios were sequenced by the University of Birmingham Sequencing Service using Illumina Sequencing chemistry. The family history of each trio is shown below in Table 27.

Table 27- Family history of trios

Tommy's ID	Gravida/ Para	Parental Karyotypes	Maternal Age	Paternal Age	Additional Info/ Comments
Family 26	G8P1	unknown	29	unknown	-
Family 57	G9P3	unknown	41	37	First pregnancy twins with only one surviving.
Family 68	G6P1	46XX/ 46XY	41	40	-
Family 75	G8P0	46XX/ 46XY	32	32	First pregnancy was terminated.
Family 82	G11P2	46XX/ 46XY	39	33	Two live births were with different partner.
Family 141	G8P0	46XX/ 46XY	33	36	Consanguineous Couple (first cousins)
Family 154	G5P1	unknown	32	33	-

4.2.2.1 Coverage

The coverage for each trio was assessed at >20x using the FETALCES-001: Exome Comparison gene panel in Congenica. The mean WES coverages for these trios using UoB-TruSeq were 94.92 %, 86.40 %, 94.58 % and 92.57 **%** for family's 26, 57, 68 and 75 respectively.

4.2.2.2 Variants

Variants were identified from the Illumina sequencing data with and without the FETALCES-001: Exome Comparison gene panel in Congenica. Three variants were identified in Family 26 with the gene panel and six variants were identified without the gene panel (Table 28). No variants were identified in Family 57 with the gene panel and three variants were identified without the gene panel (Table 29). Two variants were identified in Family 68 with the gene panel and six variants were identified without the gene panel (Table 30). Three variants were identified in Family 75 with the gene panel and five variants were identified without the gene panel). One variant was identified in Family 82 with the gene panel and six variants were identified without the gene panel (Table 32). Seven variants were identified in Family 141 with the gene panel and 16 variants were identified without the gene panel (Table 33). Seven variants were identified in Family 154 with the gene panel and seven variants were identified without the gene panel (Table 34).

Table 28- Family 26

Gene (Transcript)	Variant	Zygosity (Inheritance)	AF MAX	Computational Evidence PolyPhen Prediction SIFT Prediction GERP Score	OMIM number	Gene-Phenotype Relationship (Inheritance)	Read Split (Index case) N/mut	Pathogenicity (Classification) Significance
With FETALCES-001: Exome Comparison gene panel								
COL18A1 **(NM_130445.2)**	c.1852G>C p.(Gly618Arg)	Heterozygous (Paternal)	0	possibly damaging 0.98 4.54	120328	Knobloch syndrome, type 1 (AR)	46/40	Uncertain significance (PM2, PP3) Inherited, no 2nd variant
MAMLD1 **(NM_005491.3)**	c.*861A>C -	X-linked Hemizygous (Maternal)	0.00714	 -0.91	300120	Hypospadias 2, (XLR)	1/33	Uncertain Significance (BS2) Probable incidental finding
WDR73 **(NM_032856.2)**	c.1132delC p.(Arg378Alafs *25)	Heterozygous (*de novo*)	0.00405	 	616144	Galloway-Mowat syndrome 1 (AR)	5/15	Likely Pathogenic (PM2, PVS1_M, PM6) No 2nd variant
Without gene panel								
ARHGEF40 **(NM_018071.4)**	c.3427C>T p.(Arg1143Trp)	Heterozygous (paternal)	<0.0001	Possibly damaging 0.773 Deleterious 0 3.93	610018	-	15/14	Uncertain Significance (PM2, PP3) Inherited, no 2nd variant
HELZ2 **(NM_001037335.2)**	c.5528C>G p.(Ala1843Gly)	Heterozygous (Paternal)	0	Possibly damaging 0.844 tolerated 0.06 1.03	611265	-	15/10	Uncertain Significance (PM2) Inherited, no 2nd variant
PLEC **(NM_000445.3)**	c.10550A>C p.(Gln3517Pro)	Heterozygous (Paternal)	<0.0001	Probably damaging 0.947 2.67	601282	Epidermolysis bullosa simplex with muscular dystrophy (AR)	31/31	Uncertain significance (PM2, PP3) Inherited, no 2nd variant
SPOP **(NM_003563.3)**	c.782G>A p.(Gly261Glu)	Heterozygous (*de novo*)	0	Probably damaging 0.996 deleterious 0 5.43	602650	-	58/46	Uncertain significance (PM2, PP3, PS2_M) Uncertain
TMEMI31L **(NM_015196.3)**	c.1550-8A>C -	Heterozygous (Paternal)	0	 -1.22	616243	-	15/14	Uncertain Significance (PM2) Inherited, no 2nd variant
ZNF705E **(NM_001278713.1)**	c.63G>A p.(Trp21*)	Homozygous (Biparental)	0	1.12 	-	-	0/32	Uncertain Significance (PM2, PM4) Uncertain

AF MAX, maximum allele frequency; PolyPhen Prediction, Polymorphism Phenotyping- to predict possible impact of an amino acid substitution on the structure and function of a protein; SIFT, sorting intolerant from tolerant- to predict how amnio acid substitution affects the protein function; OMIM, online mendelian inheritance in man; GERP score, Genomic Evolutionary Rate Profiling.

Table 29- Family 57

Gene (Transcript)	Variant	Zygosity (Inheritance)	AF MAX	Computational Evidence PolyPhen Prediction SIFT Prediction GERP Score	OMIM number	Gene-Phenotype Relationship (Inheritance)	Read Split (Index case) N/mut	Pathogenicity (Classification) Significance
Without gene panel								
CDIP1 **(NM_013399.2)**	c.329G>A p.(Gly110Asp)	Heterozygous *(De novo)*	0	Probably damaging 0.931 deleterious 0.01 5.95	610503	-	22/22	Uncertain significance (PM2, PP3, PS2_M) Uncertain
CTTN **NM_138565.2)**	c.502C>T p.(Arg168*)	Heterozygous *(De novo)*	<0.0001	 3.49	164765	-	12/14	Uncertain significance (PM2, PS2_M) Uncertain
FAM222B **NM_018182.2**	c.899G>A p.(Arg300His)	Heterozygous *(De novo)*	0.000123	Possibly damaging 0.646 Tolerated 0.09 4.28	-	-	11/25	Uncertain significance (PM2, PP3, PS2_M) Uncertain

AF MAX, maximum allele frequency; PolyPhen Prediction, Polymorphism Phenotyping- to predict possible impact of an amino acid substitution on the structure and function of a protein; SIFT, sorting intolerant from tolerant- to predict how amnio acid substitution affects the protein function; OMIM, online mendelian inheritance in man; GERP score, Genomic Evolutionary Rate Profiling.

Table 30- Family 68

Gene (Transcript)	Variant	Zygosity (Inheritance)	AF MAX	Computational Evidence PolyPhen Prediction SIFT Prediction GERP Score	OMIM number	Gene-Phenotype Relationship (Inheritance)	Read Split (Index case) N/mut	Pathogenicity (Classification) Significance
With FETALCES-001: Exome Comparison gene panel								
COL18A1 **(NM_130445.2)**	c.3539G>A p.(Arg1180Gln)	Heterozygous (Maternal)	0.00423	possibly damaging 0.798 4.32	120328	Knobloch syndrome, type 1 (AR)	35/27	Uncertain Significance (PM2, PP3) Inherited, no 2nd variant
COX6B1 **(NM_001863.4)**	c.247C>T p.(Pro83Ser)	Heterozygous (*De novo)*	0.00027	possibly damaging 0.461 4.87	124089	Mitochondrial complex IV deficiency (AR)	48/48	Uncertain significance (PM2, PS2_M) Inherited, no 2nd variant
Without gene panel								
ADAMTS7 **(NM_014272.3)**	c.1510C>T p.(His504Tyr)	Heterozygous (Paternal)	0	Probably damaging 0.99 deleterious 0 3.85	605009	-	22/15	Uncertain Significance (PM2, PP3) Inherited, no 2nd variant
ATP11B **(NM_014616.2)**	c.3170A>G p.(Gln1057Arg)	Homozygous (Biparental)	0.000485	benign 0.439 tolerated 0.06 4.11	605869	-	0/44	Uncertain significance (PM2, BP4) Uncertain
CHRNA3 **(NM_001166694.1)**	c.907_908delCT p.(Leu303Aspfs*115)	Heterozygous (Maternal)	0.000385	 -0.75	118503	Susceptibility to Lung cancer	35/25	Uncertain significance (PM2, PM4) Inherited, no 2nd variant
CNTNAP4 **(NM_033401.3)**	c.3161G>A p.(Cys1054Tyr)	Heterozygous (Paternal)	0.000182	 -0.579	610518	-	79/68	Uncertain significance (PM2) Inherited, no 2nd variant
DIP2C **(NM_014974.2)**	c.3901G>A p.(Val1301Met)	Heterozygous (*De novo)*	0	Probably damaging 0.999 5.91	611380	-	12/12	Uncertain significance (PM2_M, PP3, PS2) Uncertain
ST5 **(NM_005418.3)**	c.1573T>C p.(Ser525Pro)	Homozygous (Biparental)	0	benign 0.004 tolerated 0.25 4.38	140750	-	0/31	Uncertain significance (PM2, BP4) Uncertain

AF MAX, maximum allele frequency; PolyPhen Prediction, Polymorphism Phenotyping- to predict possible impact of an amino acid substitution on the structure and function of a protein; SIFT, sorting intolerant from tolerant- to predict how amnio acid substitution affects the protein function; OMIM, online mendelian inheritance in man; GERP score, Genomic Evolutionary Rate Profiling.

Table 31- Family 75

Gene (Transcript)	Variant	Zygosity (Inheritance)	AF MAX	Computational Evidence PolyPhen Prediction SIFT Prediction GERP Score	OMIM number	Gene-Phenotype Relationship (Inheritance)	Read Split (Index case) N/mut	Pathogenicity (Classification) Significance
With FETALCES-001: Exome Comparison gene panel								
PCNT **(NM_006031.5)**	c.5612_5629delCCGAGAGAAATTTAGAAA p.(Ala1871_Ile1877delinsVal)	Heterozygous (Maternal)	0	 0.597	605925	Microcephalic osteodysplastic primordial dwarfism, type II (AR)	31/22	Uncertain significance (PM2, PM4) Inherited, inherited 2nd variant, uncertain
PCNT **(NM_006031.5)**	c.7802C>T P.(Ala2601Val)	Heterozygous (Paternal)	0.0002	Possibly damaging 0.75 -4.85	605925	Microcephalic osteodysplastic primordial dwarfism, type II (AR)	42/39	Uncertain significance (PM2) Inherited, inherited 2nd variant, uncertain
TRPV3 **(NM_145068.3)**	c.2236G>A p.(Val746Ile)	Heterozygous (Paternal)	0.000116	Probably damaging 0.994 tolerated 0.1 5.52	607066	Palmoplantar keratoderma, nonepidermolytic, focal 2 (AD)	11/13	Uncertain significance (PM2) Inherited, uncertain
Without gene panel								
DIP2B **(NM_173602.2)**	c.3344C>T p.(Ser1115Phe)	Heterozygous (Paternal)	0.000233	Probably damaging 0.995 deleterious 0 4.87	611379	Mental retardation, FRA12A type (AD)	25/23	Uncertain significance (PM2, PP3) Inherited, uncertain
KDMSA **(NM_001042603.1)**	c.4516C>T p.(Arg1506Trp)	Heterozygous (Paternal)	**< 0.0001**	Probably damaging 0.973 deleterious 0.02 4.87	180202	-	13/21	Uncertain significance (PM2, PP3) Inherited, no 2nd variant
RRP12 **(NM_015179.3)**	c.2239G>C p.(Asp747His)	Heterozygous (Maternal)	0	benign 0.004 tolerated 0.11 5.55	617723	-	14/13	Uncertain significance (PM2, BP4) Inherited, no 2nd variant
SHPRH **(NM_173082.3)**	c.3794C>T p.(Thr1265Ile)	Heterozygous (Maternal)	<0.0001	Possibly damaging 0.491 deleterious 0.01 4.89	608048	-	41/34	Uncertain significance (PM2, PP3) Inherited, no 2nd variant
SLC25A53 **(NM_001012755.3)**	c.453C>G p.(Phe151Leu)	Hemizygous (Maternal)	0	benign 0.34 tolerated 0.45 2.33	300941	-	0/26	Uncertain significance (PM2, BP4) Inherited, no 2nd variant

AF MAX, maximum allele frequency; PolyPhen Prediction, Polymorphism Phenotyping- to predict possible impact of an amino acid substitution on the structure and function of a protein; SIFT, sorting intolerant from tolerant- to predict how amnio acid substitution affects the protein function; OMIM, online mendelian inheritance in man; GERP score, Genomic Evolutionary Rate Profiling.

Table 32- Family 82

Gene (Transcript)	Variant	Zygosity (Inheritance)	AF MAX	Computational Evidence PolyPhen Prediction SIFT Prediction GERP Score	OMIM number	Gene-Phenotype Relationship (Inheritance)	Read Split (Index case) N/mut	Pathogenicity (Classification) Significance
With FETALCES-001: Exome Comparison gene panel								
ANOS1 **(NM_000216.2)**	c.1759G>T p.(Val587Leu)	Hemizygous (Maternal)	0.003867	benign 0.088 tolerated 0.15 4.43	300836	Hypogonadotropic hypogonadism 1 with or without anosmia (Kallmann syndrome 1) (XLR)	5/61	Likely Benign (BS2, BP4) None
Without gene panel								
ALPK1 **(NM_025144.3)**	c.1875G>T p.(Leu625Phe)	Heterozygous (Paternal)	0.0002	benign 0.008 tolerated 0.75 0.347	607347	-	41/49	Uncertain significance (PM2, BP4) Inherited, no 2nd variant
IQCK **(NM_153208.1)**	c.452C>T p.(Ala151Val)	Heterozygous (*de novo*)	**0**	Probably damaging 0.995 deleterious 0.04 5.22	-	-	46/50	Uncertain significance (PM2, PS2_M, PP3) Uncertain
STK36 **(NM_001243313.1)**	c.994A>G p.(Ser332Gly)	Heterozygous (Maternal)	< 0.0001	benign 0.005 tolerated 0.31 2.98	607652	-	26/26	Uncertain significance (PM2, BP4) Inherited, no 2nd variant
TAF9B **(NM_015975.4)**	c.686A>G p.(Gln229Arg)	Hemizygous (Maternal)	0.000117	benign 0.412 tolerated 0.44 4.2	300754	-	5/27	Uncertain significance (PM2, BP4) Inherited, no 2nd variant
TEX2 **(NM_018469.3)**	c.1859C>T p.(Pro620Leu)	Heterozygous (*De novo)*	0	Probably damaging 1 deleterious 0 5.78	-	-	17/14	Uncertain significance (PM2, PP3, PS2_M) Uncertain

AF MAX, maximum allele frequency; PolyPhen Prediction, Polymorphism Phenotyping- to predict possible impact of an amino acid substitution on the structure and function of a protein; SIFT, sorting intolerant from tolerant- to predict how amnio acid substitution affects the protein function; OMIM, online mendelian inheritance in man; GERP score, Genomic Evolutionary Rate Profiling.

Table 33- Family 141

Gene (Transcript)	Variant	Zygosity (Inheritance)	AF MAX	Computational Evidence PolyPhen Prediction SIFT Prediction GERP Score	OMIM number	Gene-Phenotype Relationship (Inheritance)	Read Split (Index case) N/mut	Pathogenicity (Classification) Significance
With FETALCES-001: Exome Comparison gene panel								
CDT1 **(NM_030928.3)**	c.553G>C p.(Val185Leu)	Heterozygous (Maternal)	0.00509	Benign 0.039 tolerated 0.11 2.67	605525	Meier-Gorlin syndrome (AR)	60/51	Uncertain Significance (BP4) Inherited, no 2nd variant
CENPJ **(NM_018451.4)**	:c.2896_2898del GAG p.(Glu966del)	Heterozygous (Paternal)	0. 00116	5.25	609279	Seckel syndrome (AR), Microcephaly 6, primary (AR)	107/97	Uncertain Significance (PM4) Inherited, no 2nd variant
CENPJ **(NM_018451.4)**	c.1564G>A p.(Gly522Ser)	Heterozygous (Maternal)	0	benign 0.011 tolerated 0.22 0.527	609279	Seckel syndrome (AR), Microcephaly 6, primary (AR)	42/48	Uncertain Significance (PM2, BP4) Inherited, no 2nd variant
NAA15 **(NM_057175.3)**	c.382C>T p.(Arg128*)	Heterozygous (*De novo)*	0	-0.187	608000	Mental retardation (AD)	25/43	Likely pathogenic (PM2, PVS1_S, PS2_M) Incidental finding
OCRL **(NM_000276.3)**	c.375G>T p.(Glu125Asp)	Hemizygous (Maternal)	**0**	benign 0.001 tolerated 0.38 0.803	300535	Dent disease 2 (XLR), Lowe syndrome (XLR)	1/97	Likely Benign (BS2, BP4) None
POLG **(NM_002693.2)**	c.2910C>A p.(Asn970Lys)	Heterozygous (Maternal)	0	Probably damaging 1 deleterious 0 5.24	174763	Mitochondrial DNA depletion syndrome 4A. Progressive external ophthalmoplegia (AD)	77/57	Uncertain Significance (PM2, PP3) Inherited, uncertain
RYR1 **(NM_001042723.1)**	c.13498G>C p.(Asp4500His)	Heterozygous (Maternal)	0.0097459	3.93	180901	Minicore myopathy with external ophthalmoplegia (AR) Malignant hyperthermia susceptibility 1,central core disease (AD)	41/47	Uncertain Significance (BS2) Inherited, uncertain
Without gene panel								
AKR7A3 **(NM_012067.2)**	c.331T>C p.(Tyr111His)	Homozygous (Biparental)	0.000303	Probably damaging 0.978 deleterious 0 3.21	608477	-	0/54	Uncertain Significance (PM2, PP3) Uncertain

BBX **(NM_020235.6)**	**c.1736C>T p.(Pro579Leu)**	**Homozygous (Biparental)**	**0.000424**	**Probably damaging 1** deleterious 0.01 6.07	-	-	**0/93**	**Uncertain Significance (PM2, PP3)** Uncertain
CACHD1 **(NM_020925.2)**	c.3593G>A p.(Arg1198Gln)	Homozygous (Biparental)	0.000182	Possibly damaging 0.885 Deleterious low confidence 0.03 4.75	-	-	0/112	Uncertain Significance (PM2, PP3) Uncertain
COL12A1 **(NM_004370.5)**	c.1903C>T p.(Pro635Ser)	Heterozygous (Maternal)	< 0.0001	Possibly damaging 0.734 5.68	120320	Bethlem myopathy 2 (AD)	29/20	Uncertain Significance (PP3) Inherited, uncertain
CPD **(NM_001304.4)**	c.1150G>C p.(al384Leu)	Homozygous (Biparental)	0.000303	benign 0.261 deleterious 0.01 3.9	603102	-	0/141	Uncertain Significance (PM2) Uncertain
EFCAB13 **(NM_152347.4)**	c.118A>G p.(Ile40Val)	Heterozygous (Paternal)	0.0002	benign 0.028 tolerated 0.26 -0.81	-	-	41/49	Uncertain Significance (PM2, BP4) Inherited, no 2nd variant
KIF21A **(NM_017641.3)**	c.4372G>A p.(Gly1458Arg)	Homozygous (Biparental)	0	Probably damaging 0.999 deleterious 0 5.72	608283	Congenital fibrosis of extraocular muscles (AD)	0/148	Uncertain Significance (PM2, PP3) Uncertain
NSRP1 **(NM_032141.3)**	c.1030C>G p.(His344Asp)	Homozygous (Biparental)	0	benign 0.393 deleterious 0.01 3.05	616173	-	0/67	Uncertain Significance (PM2) Uncertain
NUDT16 **(NM_152395.2)**	c.388G>C p.(Ala130Pro)	Homozygous (Biparental)	0	Probably damaging 0.991 deleterious 0 2.84	617381	-	0/48	Uncertain Significance (PM2, PP3) Uncertain
PKP2 **(NM_004572.3)**	c.404C>A p.(Ser135Tyr)	Homozygous (Biparental)	0	Possibly damaging 0.897 deleterious 0 4.76	602861	Arrhythmogenic right ventricular dysplasia (AD)	0/57	Uncertain Significance (PM2, PP3) Uncertain
SLC41A3 **(NM_017836.3)**	c.631C>T p.(Arg211*)	Homozygous (Biparental)	0.0001	4.99	610803	-	0/42	Uncertain Significance (PM2) Uncertain

SPATA20 (NM_022827.3)	**c.2135C>T p.(Ala712Val)**	**Homozygous (Biparental)**	**0**	**benign 0.005** tolerated 0.57 3.17	**613939**	**-**	**0/72**	**Uncertain Significance (PM2, BP4)** Uncertain
TEX13A (NM_031274.3)	c.677C>T p.(Ser226Phe)	Hemizygous (Maternal)	0.000105	-0.583	300312	-	0/57	Uncertain Significance (PM2) Inherited
TGFB1I1 (NM_001042454.2)	c.1132C>T p.(Pro378Ser)	Homozygous (Biparental)	< 0.0001	Probably damaging 1 tolerated 0.12 4.58	602353	-	0/25	Uncertain Significance (PM2) Uncertain
VWA2 (NM_001272046.1)	c.2092dupC p.(Gln698Profs*6)	Heterozygous (Paternal)	< 0.0001	5.25	-	-	34/36	Uncertain Significance (PM2) Inherited, no 2nd variant
XIRP2 (NM_152381.5)	c.1295G>A p.(Ser432Asn)	Heterozygous (Maternal)	0	benign 0.009 1.79	609778	-	53/53	Uncertain Significance (PM2, BP4) Inherited, no 2nd variant

AF MAX, maximum allele frequency; PolyPhen Prediction, Polymorphism Phenotyping- to predict possible impact of an amino acid substitution on the structure and function of a protein; SIFT, sorting intolerant from tolerant- to predict how amnio acid substitution affects the protein function; OMIM, online mendelian inheritance in man; GERP score, Genomic Evolutionary Rate Profiling.

Table 34- Family 154

Gene (Transcript)	Variant	Zygosity (Inheritance)	AF MAX	Computational Evidence PolyPhen Prediction SIFT Prediction GERP Score	OMIM number	Gene-Phenotype Relationship (Inheritance)	Read Split (Index case) N/mut	Pathogenicity (Classification) Significance
With FETALCES-001: Exome Comparison gene panel								
AP4M1 **(NM_004722.3)**	c.514C>T p.(Arg172Cys)	Heterozygous (Maternal)	0. 001	probably damaging 0.998 deleterious 0 4.72	602296	Spastic paraplegia 50 (AR)	105/ 111	Uncertain significance (PM2, PP3) Inherited, no 2nd variant
HYDIN **(NM_001270974.1)**	c.10652A>G p.(His3551Arg)	Heterozygous (Paternal)	0.003339	possibly damaging 0.702	610812	Ciliary dyskinesia, primary, 5 (AR)	15/9	Uncertain significance (BS2, PP3)
KMT2D **(NM_003482.3)**	c.13930C>G p.(Pro4644Ala)	Heterozygous (Paternal)	0.000962	unknown 0 0.121	602113	Kabuki syndrome (AD)	23/19	Uncertain significance (BS2)
KMT2D **(NM_003482.3)**	c.10241A>G p.(Lys3414Arg)	Heterozygous (Maternal)	0.000669	Probably damaging 0.952 3.96	602113	Kabuki syndrome (AD)	40/55	Uncertain significance (BS2, PP3) Inherited, uncertain
NHS **(NM_198270.2)**	c.3152C>T p.(Thr1051Ile)	Hemizygous (Maternal)	0.00644	benign 0.012 tolerated 0.12 4.1	300457	Nance-Horan syndrome (XLD)	0/18	Likely benign (BS2, BP4) None
PKD1 **(NM_001009944.2)**	c.7429C>T p.(Arg2477Cys)	Heterozygous (Maternal)	0.003395	Possibly damaging 0.901 4.77	601313	Polycystic kidney disease 1 (AD)	20/15	Uncertain significance (BS2, PP3) Inherited, uncertain
PKD1 **(NM_001009944.2)**	c.3854G>A p.(Arg1285Gln1)	Heterozygous (Maternal)	0.000313	benign 0.003 -0.762	601313	Polycystic kidney disease 2 (AD)	38/53	Likely Benign (BS2, BP4) None
Without gene panel								
ACSBG1 **(NM_015162.4)**	c.253C>T p.(Arg85Trp)	Heterozygous (Paternal)	0.000427	Possibly damaging 0.582 deleterious 0.04 2.92	614362	-	51/53	Uncertain significance (PM2, PP3) Inherited, no 2nd variant
ARL6IP6 **(NM_152522.5)**	c.449delT p.(Leu150Hisfs*2)	Heterozygous (*De novo*)	0	2.98	616495	-	57/58	Uncertain significance (PM2, PS2_M) Uncertain
CIC **(NM_015125.3)**	c.3214G>A p.(Val1072Met)	Heterozygous (Paternal)	<0.0001	Probably damaging 0.971 5.05	612082	Mental retardation (AD)	39/46	Uncertain significance (PM2, PP3) Inherited, uncertain
COL6A6 **(NM_001102608.1)**	c.1696C>T p.(Arg566*)	Heterozygous (Paternal)	0.000485	3.23	616613	-	68/84	Uncertain significance (PM2) Inherited, no 2nd variant

IL7R **(NM_002185.3)**	c.876+12delT -	Heterozygous (Paternal)	0.000122		146661	Severe combined immunodeficiency, T-cell negative, B-cell/natural killer cell-positive type (AR)	96/93	Uncertain significance (PM2, PM4)
				2.82				Inherited, no 2nd variant
VPS13C **(NM_017684.4)**	c.10697A>T p.(Tyr3566Phe)	Heterozygous (Paternal)	0	Possibly damaging 0.857	608879	Parkinson disease 23, early onset (AR)	142/122	Uncertain significance (PM2, PP3)
				deleterious 0.01				
				5.7				Inherited, uncertain
VPS13C **(NM_017684.4)**	c.4174G>A p.(Val1392Met)	Heterozygous (Maternal)	0	benign 0.055	608879	Parkinson disease 23, early onset (AR)	44/39	Uncertain significance (PM2, BS4)
				tolerated 0.13				
				3.14				Inherited, no 2nd variant

AF MAX, maximum allele frequency; PolyPhen Prediction, Polymorphism Phenotyping- to predict possible impact of an amino acid substitution on the structure and function of a protein; SIFT, sorting intolerant from tolerant- to predict how amnio acid substitution affects the protein function; OMIM, online mendelian inheritance in man; GERP score, Genomic Evolutionary Rate Profiling.

4.2.3 Discussion

In total seven family trios were sequenced using Illumina sequencing chemistry at the University of Birmingham Sequencing Service. Each family had suffered from recurrent miscarriage and cytogenetic testing had determined the pregnancy loss tested using WES was euploid. Unfortunately the data is not available for all the cytogenetic results of all the families' previous pregnancy losses. All the sequenced pregnancy tissue was coincidentally from male fetuses.

Variants identified from these trios were analysed firstly with the FETALCES-001: Exome Comparison gene panel at an allele frequency of <0.01. This gene panel contains genes which have been demonstrated to be causative of fetal anomalies and developmental disorders and has been derived through the PAGE study (Lord et al., 2019). The trios were also analysed without the gene panel at an allele frequency of <0.0005. Each variant was assessed using the VEP consequence and was filtered using autosomal dominant with *de novo*, autosomal recessive, compound autosomal recessive inheritance and X-linked inheritance. All variants were assessed using the ACGS/ ACMG criteria (Richards et al., 2015, Ellard, 2019)

The ACGS/ ACMG criteria were developed to standardise interpretation of sequence variants. The recommendation of the guidelines is that they are used according to the disease and inheritance pattern. Therefore ACGS criteria are not designed to be used if there is no reported disease association with a gene or if there is no Human Phenotype Ontology (HPO) term. In the case of this research, there are no HPO terms as these are early pregnancy losses with no ultrasound scan findings, and for many of the genes not in the

FETALCES-001: Exome Comparison gene panel there is currently no reported disease association. The ACMG criterion has still been used. However, in cases where the ACMG criterion relies on the mechanism of the disease (e.g. loss of function) and there is no disease associated with the disease, the ACMG criterion was not used. For example, PVS1 was not used for truncating variants if there was no disease associated with the gene which may have a loss of function mechanism. Alternatively where there was an associated disease, PVS1 was downgraded (PVS1_S) for variants which had an associated disease caused by loss of function variants as there was no HPO term and these could be incidental findings. This highlights the difficulty with interpretation of the clinical significance of variants identified in these cohorts by exome sequencing.

As the WES was completed in trios, the maternity and paternity were confirmed by the number of maternally/ paternally inherited variants. In cases where there was a *de novo* variant in a gene which has no disease associated with it, the *de novo* criteria PS2 was downgraded to moderate (PS2_M) evidence because there wasn't a disease association and/or associated HPO terms. If there was a disease association, and the variant was seen often in variant databases (GnomAD- https://gnomad.broadinstitute.org/), the ACMG criterion BS2 was applied.

4.2.3.1 Family 26

Three variants were identified in Family 26 (Table 28) using the gene panel FETALCES-001: Exome Comparison gene panel. *COL18A1* and *MAMLD1* were classified as variants of unknown significance (VUS) using the ACMG standards and guidelines and *WDR73* identified as likely pathogenic. As the variant in WDR73 was a frame shift variant resulting in a premature stop coding the ACMG criterion PVS1_S was used and downgraded as there were

no HPO terms. PVS1 was then downgraded further to PVS1_M as the frameshift variant was towards the end of the protein. WDR73 regulates cell cycle progression, proliferation and survival by its interaction with microtubules (Jinks et al., 2015) and homozygous mutations in *WDR73* are known to cause Galloway-Mowat syndrome 1, caused by either frameshift variants (Colin et al., 2014, Vodopiutz et al., 2015) or point mutations (Ben-Omran et al., 2015) that result in a premature stop codon. The likely pathogenic variant in family 26 is a *de novo* heterozygous, frameshift variant in the gene *WDR73* causing a premature termination of translation. However, as there was no second variant identified, this variant would not be considered causative of the miscarriage. The read split of *WDR73* is skewed (5/15) which suggests that this variant could be an artefact or from maternal cell contamination (MCC). As the fetal DNA was obtained from products of conception (POC) some maternal tissue could have been sampled when the fetal DNA was extracted which resulted in MCC. If this was MCC then the fetus could have been homozygous for this variant. Alternatively this skewed read split could be caused by a deletion on the other allele or from uniparental disomy.

Six variants were identified in family 26 without the gene panel. All the variants were VUS, and four were heterozygously inherited from a parent, with no second variant inherited and therefore unlikely to be relevant to the pregnancy loss. For example, mutations in *PLEC* are not lethal *in utero* and there is no second variant in the *PLEC* gene. A *de novo* missense variant in *SPOP* was identified in a conserved domain. *SPOP* is likely to be involved in the cell cycle and somatic mutations have been associated with prostate cancer (Barbieri et al., 2012, Zuhlke et al., 2014) and endometrial tumours (Le Gallo et al., 2012). The PS2 criteria was downgraded to PS2_M as there was no disease associated with the gene, and therefore

the significance of this finding is uncertain. A homozygous truncating variant was identified in the *ZNF705E* gene which could potentially be of interest, however the ZNF705E gene is not well characterised, is not on OMIM, and not reported in the literature.

4.2.3.2 Family 57

DNA from two pregnancies was collected as part of the Tommy's trial from family 57. One pregnancy loss was a mosaic trisomy 16 and the other was euploid. Only the euploid pregnancy loss was sequenced in the WES trio. No variants were identified in Family 57 using the FETALCES-001: Exome Comparison gene panel and three VUS were identified without the gene panel. None of the VUS had a gene-phenotype relationship and *FAM222B* was also not an OMIM gene. Two *de novo* variants in the genes *CTTN* and *CDIP1* were identified with a heterozygous zygosity. Multiple paternal SNPS were identified without any filters applied and so non-paternity can be excluded. *CTTN* is the gene for Cortactin which is a regulator of actin cytoskeleton including cell migration and invasion (Yamaguchi and Condeelis, 2007) and is also a known oncogene (Hui et al., 1997, Luo et al., 2006). *CDIP1* is a cell death-inducing p53 target acting as a signal transducer for endoplasmic reticulum stress-mediated apoptosis (Namba et al., 2013), suggesting these genes may be of some importance. The variants identified in *CTTN and CDIP* are seen zero times and once heterozygously in allele frequency databases (GnomAD) respectively and are in highly conserved orthologs.

4.2.3.3 Family 68

Two variants were identified in Family 68 using the FETALCES-001: Exome Comparison gene panel and both variants were classed as VUS. A maternally inherited heterozygous mutation

in *COL18A1* with no variant on the other allele and a *de novo* mutation in *COX6B1*, with an AF Max of 0.00027, which argues against lethality as this is an autosomal recessive disorder and no other variants were identified on the other allele.

With the gene panel removed, 6 variants were identified in family 68. Three variants were identified which were heterozygously inherited. It is therefore unlikely that these three variants are pathogenic as they have been parentally inherited with no second variant identified in the gene. A heterozygous *de novo* variant was identified in *DIP2C.* However, as there is low coverage in the parents it is difficult to be sure that this is a true *de novo* variant. *DIP2C* has little known function and is not an OMIM morbid gene.

Two VUS missense variants were identified in the homozygous state in the proband, with biparental inheritance. *ST5* is a tumour suppressor gene and has been shown to be expressed in fetal tissue (Gohring et al., 2010) and *ATP11B* is an ATPase which transports ions across membranes. As there is little in the literature about the function of these genes and any potential disease associations then the significance of this finding is uncertain.

Many biparentally inherited variants were identified in Family 68, although most of them were filtered out due to their maximum allele frequency being above 0.0005. This suggests that this may be a consanguineous coupling, however this has not been confirmed clinically.

4.2.3.4 Family 75

Three variants in two genes were identified in the pregnancy loss of Family 75 using the FETALCES-001: Exome Comparison gene panel. *TRPV3* was classified as VUS, was inherited from a parent and is unlikely to be of any significance to the pregnancy loss. Two variants were identified in the gene *PCNT,* which were both classified of VUS. The *PCNT* gene on

chromosome 21 encodes the PCNT protein which localises to centrosomes in the cell cycle and is involved with regulating spindle microtubule formation during mitosis (Flory et al., 2000). In the absence of PCNT, the mitotic spindles are disorganised and premature centriole separation causes mis-segregation of chromosomes (Kim and Kim, 2019, Rauch et al., 2008). The *PCNT* gene is implicated in Microcephalic osteodysplastic primordial dwarfism, type II (MOPDII) and has been shown to be caused by both homozygous and compound heterozygous mutations in *PCNT* (Rauch et al., 2008, Kantaputra et al., 2011) and Willems et al have concluded that MOPDII is caused by a loss of function of PCNT (Willems et al., 2010). While both of the variants in *PCNT were* classified as VUS they could be of interest from the gene function and the disorder can cause IUGR which could be relevant to the pregnancy loss. Both variants had a 0 or very low AF MAX score and were not recorded in the ClinVar database. Of particular interest is the maternally inherited inframe deletion (Figure 26) where the deletion of 18 base pairs is reflected in the coverage. Where the deletion has occurred in one allele it has evidently reduced the coverage at that region.

Figure 26-Maternally inherited Inframe deletion of *PCNT.*

Coverage of both Mother and patient shows an 18 base pair deletion on one allele.

Five variants were identified without the gene panel. The *DIP2B* gene is associated with mental retardation but as this variant was classified as VUS and inherited paternally it is very unlikely that this variant would have caused an incidental finding of mental retardation. Three heterozygous variants in *KDMSA, RRP12,* and *SHPRH* were identified and all classified as VUS. These variants are all missense variants which were parentally inherited and have no disease associated with the gene. A hemizygous variant in *SLC25A53* was also identified maternally. Again, this was VUS with no disease association.

4.2.3.5 Family 82

One variant was identified in Family 82 using the FETALCES-001: Exome Comparison gene panel. A maternally inherited variant in *ANOS1* was identified which encodes for the glycoprotein, anosmin, which is involved with the migration of Gonadotrophin releasing hormone (GNRH) and development of olfactory nerves (Cariboni et al., 2004). Either deletions (Bick et al., 1992, Hardelin et al., 1993) or mutations (Oliveira et al., 2001) in the *ANOS1* gene have been identified to cause the X-linked condition, Hypogonadotropic hypogonadism with or without anosmia (Kallmann syndrome) in male fetuses. However, the phenotype is unlikely to be severe enough to be significant for this pregnancy loss and is a likely benign variant.

Without the gene panel, five variants were identified and classified as VUS. Two *de novo* variants, *IQCK* and *TEX2,* were heterozygous in non-OMIM genes and three variants, *ALPK1, STK36* and *TAF9B,* were parentally inherited and VUS with no phenotype association so are unlikely to be of any significance.

4.2.3.6 Family 141

Seven variants in six genes were identified in the consanguineous Family 141 using the FETACLCES-001: Exome Comparison gene panel. A *de novo* stop gain variant was identified in *NAA15* and was classified as likely pathogenic. As there were no HPO terms the ACMG classifications PVS1 and PS2 were both downgraded. It is likely that this variant in *NAA15* is an incidental finding associated with mental retardation but it is not associated with and unlikely to have caused the pregnancy loss.

Two autosomal recessive compound heterozygous, variants were identified in *CENPJ* (one missense variant and one deletion of a single amino acid). One variant was inherited maternally and one variant was inherited paternally and both variants were classified as VUS. *CENPJ* is a centrosomal protein which has a role in microtubule nucleation (Hung et al., 2000) and has been associated with Seckel syndrome (Al-Dosari et al., 2010) and microcephaly (Bond et al., 2005, Leal et al., 2003). The disease mechanism is reported to be loss of function, therefore it is less likely that this finding is clinically significant. However Seckle syndrome is associated with IUGR and there are small numbers of pathogenic missense variants reported on ClinVar, therefore it can't be completely excluded.

A maternally inherited X-linked variant was identified in *OCRL*. This was classified as likely benign and is therefore unlikely to be an incidental finding or significant to the pregnancy loss. Maternal inherited variants in *CDT1, POLG* and *RYR1* were also identified in the conceptus of family 141. These variants were all classified as VUS and as they were also identified heterozygously in the maternal sample, they are unlikely to have any significance to the pregnancy loss.

Without the gene panel, 16 variants were identified in family 141 and all 16 were classified as VUS. Of particular interest was a biparental homozygous mutation in *PKP2. PKP2* is the gene for Plakophilin which are structural proteins found within the desmosome structure in the cytoskeleton (Bonne et al., 2000). Variants in the *PKP2* gene have been associated with Arrythmogenic right ventricular dysplasia (Gerull et al., 2004) inherited in a dominant manner. However *PKP2* null mice are embryonically lethal (Grossmann et al., 2004), and there are some reports in the literature of homozygous variants in the PKP2 gene causing more severe cardiac anomalies such as hypoplastic left heart (Verhagen et al., 2018,

Ramond et al., 2017). Therefore it is possible that this finding is relevant in the context of early fetal loss.

In addition, there were 10 other variants which were homozygous and biparentally inherited in this consanguineous family, four parentally inherited variants and one maternally derived hemizygous variant. However, none of these variants were thought to be of any significance. These results show the potential difficulties with interpretation of large-scale sequencing data from consanguineous or related couples that may share lots of rare variants.

4.2.3.7 Family 154

Seven variants in five genes were identified in family 154. Variants in *AP4M1* and *HYDI* were identified heterozygously and inherited maternally and paternally respectively. Variants in *AP4M1* have been shown to cause Spastic paraplegia (Verkerk et al. 2009, Tuysuz et al. 2014) and variant in *HYDIN* have been identified to cause Ciliary dyskinesia (Olbrich et al. 2012). However as these variants are VUS and in autosomal recessive diseases it is unlikely these variants are causative of the pregnancy loss in the absence of a second variant.

Five variants were in autosomal dominant genes. Two missense variants were identified in the *KMT2D* gene which causes Kabuki syndrome, characteristic of congenital mental retardation and other phenotypes (Niikawa et al., 1981). However as Kabuki syndrome is caused by loss of function variants and these missense variants were parentally inherited, it is unlikely these variants were causative of the pregnancy loss or incidental findings. Two missense variants were identified in the *PKD1* gene. Again, these were inherited parentally and unlikely to be causative of the pregnancy loss or incidental findings. An X-linked

dominant variant was identified in the *NHS* gene, this was hemizygous and maternally inherited but was classified as likely benign.

Seven additional variants were identified without the gene panel and all classified as VUS. As the variant in the *ACSBG1* gene was a frameshift variant, causing a premature stop codon, with no disease associated the ACMG classification PVS1 was not used and the PS2 classification for *de novo* variants was downgraded as there were no HPO terms. The stop gain variant in the *COL6A6* gene also had no disease associated, so the ACMG classification PVS1 was also not used. The five other variants identified without the gene panel were all classified as VUS and parentally inherited so unlikely to be of any significance.

4.3 Conclusion

The systematic review in chapter three identified potential genetic causes in euploid miscarriages and concluded that while there are many genetic disruptions that could be causative of miscarriage they are not well known or understood. In this chapter we aimed to identify potentially genetic causes of euploid miscarriages in our recurrent miscarriage patient cohort. It is important to note that miscarriage has multifactorial and therefore environmental factors, biological factors and multiple genetic factors can combine to increase the risk of miscarriage. It is likely that in some cases small genetic contributions to miscarriage are of complex inheritance and do not follow simple Mendelian inheritance. Therefore, variants identified by WES could be associated with miscarriage but not a definitive cause.

Initially a pilot study was used to identify the best method of sequencing for our patient cohort. Three sequencing methods, Illumina Agilent and BGI were selected and the results compared. While Agilent was cheaper, had more coverage and fewer artefacts, it was concluded that Illumina was the preferred method of sequencing due to ease and identifying slightly more variants upon analysis.

Seven trios were sequenced in total by the University of Birmingham sequencing service using Illumina sequencing chemistry. The data was uploaded to Congenica and the variants were analysed with and without the FETALCES-001: Exome Comparison gene panel. While there were several variants in interesting genes, none of these were considered to be definitively causative of the pregnancy loss. Seven trios is a very small cohort and it was unlikely that genetic causes of miscarriage will be identified with such a small cohort. This study also identifies that routine WES within this patient group is difficult to assess due to the lack of phenotype information. Ideally, a much larger study cohort with ultrasound scan findings would be needed to identify and prove any significant genetic association to miscarriage. Unfortunately this would be costly, time consuming, and difficult to be conclusive in many cases.

One of the biggest weaknesses of this study was that while it was known that the trios were from euploid pregnancy losses and the families had recurrent miscarriage, the cytogenetics of the previous pregnancy losses was not known and DNA was not available. If DNA was available from multiple pregnancy losses from the same families the variants identified could be compared between pregnancies. While in most cases variants were seen to be inherited both maternally and paternally, Family 57 had all *de novo* variants from the filtering settings. This is unusual and brings in to question the family relationship or a

sample mix up. However, when all filters were removed from the analysis of family 57 both maternally and paternally inherited variants were observed.

While no definitive genetic cause for the pregnancy loss was identified in this cohort, it is still likely plausible that some pregnancy losses are associated with genetic aberrations detected by WES. The systematic review in chapter three identifies monogenic contributions to pregnancy loss in euploid pregnancies. Knowledge of these specific genes which contribute to pregnancy loss could be of important clinical significance. Much larger cohort studies and functional follow-on studies are required to identify candidate genes which are significant to pregnancy loss.

Chapter 5: Conclusion

Chapter 5- Conclusion

Pregnancy loss is the most common adverse outcome of pregnancy and occurs in 1 in 4 pregnancies. There are many risk factors associated with recurrent miscarriage with genetic factors being the most common cause of pregnancy loss. Approximately 50 % of miscarriages are caused by chromosomal abnormalities.

This thesis explored the use of cfDNA in the evaluation of miscarriage and genetic causes of miscarriage. A clinical trial was implemented to explore whether cfDNA can be used to give a genetic diagnosis for the cause of miscarriage, a systematic review was published to identify potential genetic causes in apparently euploid pregnancies and whole exome sequencing was used to identify variants in genes which could be causative of miscarriage within our trial cohort.

5.1 The use of cfDNA in the investigation of miscarriage

Current analysis of recurrent miscarriage uses cytogenomics to detect abnormalities in the fetal genome. This relies on the collection of pregnancy tissue, which faces many difficulties. Two previous studies have been published which used cfDNA to detect aneuploidies in nonviable pregnancies.

A prospective cohort study (Clark-Ganheart et al., 2020) analysed 50 cfDNA samples of non-viable pregnancies. The gestations by ultrasound scan ranged from 6.1-38.4 weeks. Thirty-eight of the 50 samples had a reportable result including eight samples which demonstrated trisomies. This study demonstrated that cfDNA can be used from non-viable pregnancies to

identify aneuploidies. However, this study had a larger number of later gestations and recommended cfDNA in non-viable pregnancies after a gestation of eight weeks only.

The study by (Yaron et al., 2020) used cfDNA to analyse pregnancy loss at less than 14 weeks gestation. 86 pregnancies had cfDNA results with comparable POC results (which was achieved by CVS sampling) and the median FF was 5 %. Out of the 86 samples, 64 % had a chromosomal abnormality and a detection rate of 55 % was achieved using the log-likelihood ratio (LLR) standardly used for NIPT. To increase the sensitivity, a pregnancy-loss specific threshold was developed using a 50 sample 'training set'. This increased the detection rate to 82 %. Trisomies not detectable using NIPT LLR thresholds but detectable using pregnancy- loss specific LLR thresholds were Trisomy 2, 4 and 17. Monosomy X and triploidy were not identified using either threshold.

Our cfDNA study cohort were recruited through Tommy's miscarriage Centre at Birmingham Women's Hospital and included pregnancy losses in the first trimester and therefore the cfDNA sample were <13 weeks gestation.

Using the SE-HiSeq-WC method, 59.5 % of samples were identified concordantly to POC testing. WISECONDOR was initially used as it was a published freely available pipeline that could be implemented in to West Midland's Regional Genetics Laboratory.

The PE-NextSeq-Illumina method of cfDNA was completed externally by an Illumina laboratory in Cambridge and processed in a 24-sample batch through a modified Illumina VeriSeq NIPT solution v2 workflow. In our study, 75.9 % of samples were identified concordantly to POC testing (not including triploid samples). It was initially thought that the reduced performance in our study was due to the lower gestations; however the difference

in pipelines may have also contributed to this difference. Using the pregnancy-loss specific LLR improved the detection rate in the (Yaron et al., 2020) publication and it is feasible that having a pregnancy-loss specific LLR for the cfDNA samples in the study sequenced using PE-NextSeq-WC would improve the detection rate.

The next step in this study may be to conduct a trial using an algorithm similar to the one proposed by (Yaron et al., 2020), trialling fetal fraction cut offs and using a pipeline with modified LLRs to optimise the detection rate for this cohort of patients.

Yaron et al. (2020) have proposed an algorithm for RPL using cfDNA rather than CMS analysis of POC. However, currently cfDNA testing is not reliable enough to offer solely for RPL. Therefore we propose the below RPL algorithm, aimed at the recommended guidelines for the UK (Figure 27).

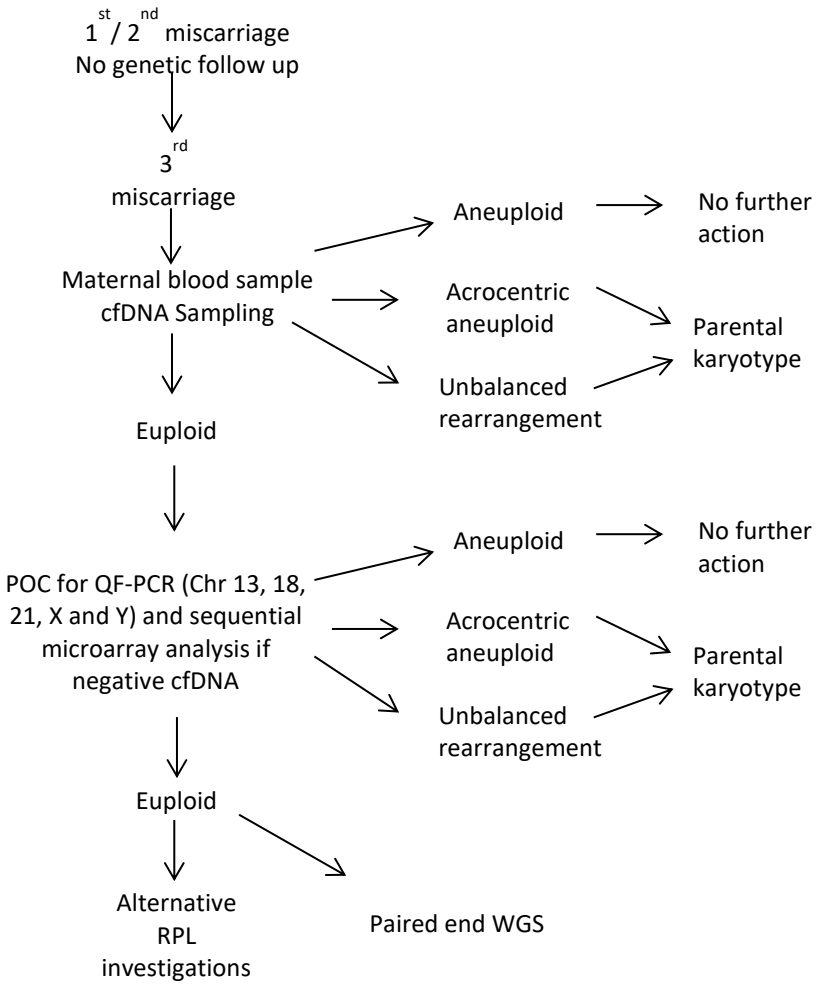


Figure 27- Recurrent pregnancy loss algorithm

Expanded from (Yaron et al., 2020).

RPL workup, in the UK, is only recommended for third and subsequent miscarriages; therefore, there is no genetic follow up for the first or second pregnancy loss. We propose that a maternal blood sample, cfDNA sample, and paternal blood sample (where available) is collected at the point of miscarriage diagnosis and the POC collected if and when available. Initially cfDNA sampling will be conducted. Cases of unbalanced rearrangements or acrocentric aneuploidies should be followed up by parental karyotyping to establish whether one or more of the parents carry a balanced rearrangement, which would confer

increased risk in subsequent pregnancies. Currently cfDNA testing is not accurate enough where an apparently normal result is obtained, as genetic aberrations with low fetal fractions and triploid samples are not identified. Therefore, in the case of an apparently euploid result on cfDNA, POC testing by QF-PCR (Chr 13, 18, 21, X and Y) and sequential chromosomal microarray analysis (if the QF-PCR is normal) can be activated. Routinely QF-PCR is completed for chromosomes 13, 18, 21, X and Y as large chromosomal abnormalities in these chromosomes are the only ones viable with life. However, in a miscarriage setting including all the acrocentric chromosomes would be of benefit. If the POC testing shows an apparent euploid result then alternative RPL investigations are recommended.

A recent study (Dong et al., 2019) has also looked at the potential of using low-pass, paired-end, genomic sequencing on maternal and paternal peripheral blood samples. Recurrent miscarriage (≥2) couples received low-pass, paired-end, genomic sequencing on maternal and paternal peripheral blood to identify chromosomal rearrangements and CNVs. In total, 2167/2180 (99.4 %) yielded results. Low-pass, paired-end, genomic sequencing on maternal and paternal identified 127 patients (11.7 %) with balanced structural chromosomal rearrangements, including 78 (7.2 %) balanced translocations and 49 inversions (4.5 %). All results were confirmed by PCR and Sanger sequencing. CNVs were also identified using low-pass, paired-end, genomic sequencing and 2124 deletions and 4623 duplications were identified, averaging approximately two deletions and 4.2 duplications per couple. Using ACMG guidelines, six CNVs were identified as pathogenic and 12 CNVs were identified as uncertain significance and were validated using CMA. Detecting parental chromosomal aberrations is important in the management of RM couples and the study by (Dong et al., 2019) has demonstrated that low-pass genomic sequencing can be used to decipher this in

RM patients with increased resolution and detection rates. This could be used as an adjunct to parental karyotyping.

5.2 Potential Genetic Causes of miscarriage in euploid pregnancies

This thesis also explored other potential genetic causes of recurrent miscarriage in euploid pregnancies. While it is known that 50 % of miscarriages are caused by chromosomal abnormalities it is feasible that apparently euploid pregnancy losses could have an underlying genetic cause not currently detected by routine testing. We published a systematic review that explored this and completed WES in seven trios with apparently euploid pregnancy losses.

The systematic review identified 50 studies which found putatively causative variants in a range of genes. Key aspects of the literature included genetic aberrations identified through WES and microarray analysis, genes associated with diploid and biparental recurrent molar pregnancies and other genetic contributions such as epigenetics and mitochondrial DNA.

WES studies identified genes that have an early role in developmental biology and are essential in embryogenesis and CNV studies identified CNVs were significantly represented in genes involving antigen binding and immune signalling genes. Evidence also suggested that genes involved in oocyte development, maturation, and epigenetic reprogramming are likely to be important in a subset of pregnancy losses.

Although genes were identified to have causal or associated links with recurrent miscarriage it is difficult to state this conclusively. However, there are likely to be complex genetic

contributions to pregnancy loss and knowledge of specific genes could contribute to the understanding of biological pathways that contribute to pregnancy loss. Routine investigations of pregnancy loss focus on aneuploidy and other chromosomal abnormalities. However, with increased knowledge of additional non-aneuploid contributions to RPL, additional genetic testing recommendations may be made in the future to couples experiencing RPL.

5.3 Whole Exome Sequencing

The WES sequencing completed in chapter 4, while not identifying any variant in a gene that could conclusively be considered causative of the pregnancy loss, it did identify some interesting variants. The only variant that was *de novo* likely pathogenic was a variant in *WDR73* which regulates cell cycle regression and proliferation. This is important during embryogenesis so it could be postulated that this could have had an effect on the pregnancy loss. However, variants in the *WDR73* can be causative of Galloway-Mowat syndrome and while it is not known to be prenatally morbid, death in early childhood may occur. Eight other *de novo* variants were identified in four of the trios. All seven of these *de novo* variants were classified as uncertain significance using the ACMG criteria and they are unlikely to be causative of recurrent pregnancy loss due to being sporadic. Many of these genes are currently uncharacterised and others are involved in the cell cycle regulation (*SPOP, CDIP1, CTTN*). Three homozygous variants (*ATP11B. ST5, PKP2*) and three compound heterozygous variants (*PCNT, KMT2D, CENPJ*) were identified that were biparentally inherited. While these variants were all classified as uncertain significance, it was interesting that these genes are involved in cell organisation and regulation. It is difficult to say whether

any of these findings are significant, and this small study highlights that routine WES would be difficult in the cohort of patients due to lack of patient phenotype. It is therefore difficult to draw any definitive conclusions without further studies.

Although WES is not currently recommended for routine diagnostic testing of pregnancy loss, if in the future monogenic causes of pregnancy loss are identified, through large, well-designed studies, the identification of non-aneuploid causes of RPL could be of significant impact. Studies could include knock-out mouse models to see if genes where variants were identified are embryonic lethal. Functional studies could also be used to identify where a gene may have an impact on the embryonic development.

5.4 Closing remarks

As more is known about the causes of pregnancy loss, and genetic analysis becomes higher resolution and more cost effective, more genetic testing can be offered in the case of pregnancy loss.

This thesis has highlighted that there are multiple ways to manage the genetics of recurrent miscarriage and as genetic technologies and understanding improves so does the management. This would have implications for diagnosis and recurrence risks.

References

(NICE), N. I. F. H. A. C. E. 2012. Ectopic pregnancy and miscarriage: diagnosis and initial management London.

(NICE), N. I. F. H. A. C. E. 2016. High-throughput non-invasive prenatal testing for fetal RHD genotype. London.

ABDALLA, E. M., HAYWARD, B. E., SHAMSEDDIN, A. & NAWAR, M. M. 2012. Recurrent hydatidiform mole: detection of two novel mutations in the NLRP7 gene in two Egyptian families. *Eur J Obstet Gynecol Reprod Biol,* 164**,** 211-5.

ABU-HEIJA, A. 2014. Thrombophilia and Recurrent Pregnancy Loss: Is heparin still the drug of choice? *Sultan Qaboos Univ Med J,* 14**,** e26-36.

AGHAJANOVA, L., MAHADEVAN, S., ALTMAE, S., STAVREUS-EVERS, A., REGAN, L., SEBIRE, N., DIXON, P., FISHER, R. A. & VAN DEN VEYVER, I. B. 2015. No evidence for mutations in NLRP7, NLRP2 or KHDC3L in women with unexplained recurrent pregnancy loss or infertility. *Human Reproduction,* 30**,** 232-8.

AHMADI, M. H., HANTUOSHZADEH, S., OKHOVAT, M. A., NASIRI, N., AZARKEIVAN, A. & AMIRIZADEH, N. 2016. Fetal RHD Genotyping from Circulating Cell-Free Fetal DNA in Plasma of Rh Negative Pregnant Women in Iran. *Indian J Hematol Blood Transfus,* 32**,** 447-453.

AL-DOSARI, M. S., SHAHEEN, R., COLAK, D. & ALKURAYA, F. S. 2010. Novel CENPJ mutation causes Seckel syndrome. *J Med Genet,* 47**,** 411-4.

ALBERRY, M., MADDOCKS, D., JONES, M., ABDEL HADI, M., ABDEL-FATTAH, S., AVENT, N. & SOOTHILL, P. W. 2007. Free fetal DNA in maternal plasma in anembryonic pregnancies: confirmation that the origin is the trophoblast. *Prenat Diagn,* 27**,** 415-8.

ANDERSEN, A. M., ANDERSEN, P. K., OLSEN, J., GRONBAEK, M. & STRANDBERG-LARSEN, K. 2012. Moderate alcohol intake during pregnancy and risk of fetal death. *Int J Epidemiol,* 41**,** 405-13.

ANDREASEN, L., CHRISTIANSEN, O. B., NIEMANN, I., BOLUND, L. & SUNDE, L. 2013. NLRP7 or KHDC3L genes and the etiology of molar pregnancies and recurrent miscarriage. *Molecular Human Reproduction,* 19**,** 773-81.

ANDREWS, D.; Padhukasahasram, B.; Kim, S. P1-66. Validation of VeriSeq NIPT Solution v2, a paired-end PCR-free WGS-based methodology for noninvasive prenatal screening of fetal aneuploidies and CNVs on a genome-wide scale. Poster Abstracts of the ISPD 23rd International Conference on Prenatal Diagnosis and Therapy, Singapore, 7–11 September 2019. Prenat Diagn. 2020, 40 (Suppl. 1), 62–63.

ARIEL, I., LERER, I., YAGEL, S., COHEN, R., BEN-NERIAH, Z. & ABELIOVICH, D. 1997. Trisomy 2: confined placental mosaicism in a fetus with intrauterine growth retardation. *Prenat Diagn,* 17**,** 180-3.

ASHOOR, G., SYNGELAKI, A., POON, L. C., REZENDE, J. C. & NICOLAIDES, K. H. 2013. Fetal fraction in maternal plasma cell-free DNA at 11-13 weeks' gestation: relation to maternal and fetal characteristics. *Ultrasound Obstet Gynecol,* 41**,** 26-32.

AYKUT, A., ONAY, H., SAGOL, S., GUNDUZ, C., OZKINAY, F. & COGULU, O. 2013. Determination of fetal rhesus d status by maternal plasma DNA analysis. *Balkan J Med Genet,* 16**,** 33-8.

BAGHERI, H., MERCIER, E., QIAO, Y., STEPHENSON, M. D. & RAJCAN-SEPAROVIC, E. 2015. Genomic characteristics of miscarriage copy number variants. *Molecular Human Reproduction,* 21**,** 655-61.

BAGISLAR, S., USTUNER, I., CENGIZ, B., SOYLEMEZ, F., AKYERLI, C. B., CEYLANER, S., CEYLANER, G., ACAR, A. & OZCELIK, T. 2006. Extremely skewed X-chromosome inactivation patterns in women with recurrent spontaneous abortion. *Aust N Z J Obstet Gynaecol,* 46**,** 384-7.

BAKKER, I., BOON, E. M. J. & SISTERMANS, E. A.
BARBIERI, C. E., BACA, S. C., LAWRENCE, M. S., DEMICHELIS, F., BLATTNER, M., THEURILLAT, J. P., WHITE, T. A., STOJANOV, P., VAN ALLEN, E., STRANSKY, N., NICKERSON, E., CHAE, S. S., BOYSEN, G., AUCLAIR, D., ONOFRIO, R. C., PARK, K., KITABAYASHI, N., MACDONALD, T. Y., SHEIKH, K., VUONG, T., GUIDUCCI, C., CIBULSKIS, K., SIVACHENKO, A., CARTER, S. L., SAKSENA, G., VOET, D., HUSSAIN, W. M., RAMOS, A. H., WINCKLER, W., REDMAN, M. C., ARDLIE, K., TEWARI, A. K., MOSQUERA, J. M., RUPP, N., WILD, P. J., MOCH, H., MORRISSEY, C., NELSON, P. S., KANTOFF, P. W., GABRIEL, S. B., GOLUB, T. R., MEYERSON, M., LANDER, E. S., GETZ, G., RUBIN, M. A. & GARRAWAY, L. A. 2012. Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat Genet,* 44**,** 685-9.
BAUD, D., GOY, G., JATON, K., OSTERHELD, M. C., BLUMER, S., BOREL, N., VIAL, Y., HOHLFELD, P., POSPISCHIL, A. & GREUB, G. 2011. Role of Chlamydia trachomatis in miscarriage. *Emerg Infect Dis,* 17**,** 1630-5.
BEN-OMRAN, T., FAHIMINIYA, S., SORFAZLIAN, N., ALMURIEKHI, M., NAWAZ, Z., NADAF, J., KHADIJA, K. A., ZAINEDDIN, S., KAMEL, H., MAJEWSKI, J. & TROPEPE, V. 2015. Nonsense mutation in the WDR73 gene is associated with Galloway-Mowat syndrome. *J Med Genet,* 52**,** 381-90.
BENDER ATIK, R., CHRISTIANSEN, O. B., ELSON, J., KOLTE, A. M., LEWIS, S., MIDDELDORP, S., NELEN, W., PERAMO, B., QUENBY, S. & VERMEULEN, N. 2018. ESHRE guideline: recurrent pregnancy loss. 2018**,** hoy004.
BENDROTH-ASMUSSEN, L., AKSGLAEDE, L., GERNOW, A. B. & LUND, A. M. 2016. Glycogen Storage Disease Type IV: A Case With Histopathologic Findings in First-Trimester Placental Tissue. *International Journal of Gynecological Pathology,* 35**,** 38-40.
BHUIYAN, Z. A., MOMENAH, T. S., GONG, Q., AMIN, A. S., GHAMDI, S. A., CARVALHO, J. S., HOMFRAY, T., MANNENS, M. M., ZHOU, Z. & WILDE, A. A. 2008. Recurrent intrauterine fetal loss due to near absence of HERG: clinical and functional characterization of a homozygous nonsense HERG Q1070X mutation. *Heart Rhythm,* 5**,** 553-61.
BICK, D., FRANCO, B., SHERINS, R. J., HEYE, B., PIKE, L., CRAWFORD, J., MADDALENA, A., INCERTI, B., PRAGLIOLA, A., MEITINGER, T. & BALLABIO, A. 1992. Brief report: intragenic deletion of the KALIG-1 gene in Kallmann's syndrome. *N Engl J Med,* 326**,** 1752-5.
BLOOM-FESHBACH, K., SIMONSEN, L., VIBOUD, C., MOLBAK, K., MILLER, M. A., GOTTFREDSSON, M. & ANDREASEN, V. 2011. Natality decline and miscarriages associated with the 1918 influenza pandemic: the Scandinavian and United States experiences. *J Infect Dis,* 204**,** 1157-64.
BOND, J., ROBERTS, E., SPRINGELL, K., LIZARRAGA, S. B., SCOTT, S., HIGGINS, J., HAMPSHIRE, D. J., MORRISON, E. E., LEAL, G. F., SILVA, E. O., COSTA, S. M., BARALLE, D., RAPONI, M., KARBANI, G., RASHID, Y., JAFRI, H., BENNETT, C., CORRY, P., WALSH, C. A. & WOODS, C. G. 2005. A centrosomal mechanism involving CDK5RAP2 and CENPJ controls brain size. *Nat Genet,* 37**,** 353-5.
BONDESON, M. L., ERICSON, K., GUDMUNDSSON, S., AMEUR, A., PONTEN, F., WESSTROM, J., FRYKHOLM, C. & WILBE, M. 2017. A nonsense mutation in CEP55 defines a new locus for a Meckel-like syndrome, an autosomal recessive lethal fetal ciliopathy. *Clinical Genetics.*
BONNE, S., VAN HENGEL, J. & VAN ROY, F. 2000. Assignment of the plakophilin-2 gene (PKP2) and a plakophilin-2 pseudogene (PKP2P1) to human chromosome bands 12p11 and 12p13, respectively, by in situ hybridization. *Cytogenet Cell Genet,* 88**,** 286-7.
BOOTS, C. & STEPHENSON, M. D. 2011. Does obesity increase the risk of miscarriage in spontaneous conception: a systematic review. *Semin Reprod Med,* 29**,** 507-13.
BORRELL, A. 2019. *RE: Personal communication: Illumina NIPT Lab Partner Day.*
BROWN, L., MOUNT, S., REDDY, R., SLIM, R., WONG, C., JOBANPUTRA, V., CLIFFORD, P., MERRILL, L. & BROWN, S. 2013. Recurrent pregnancy loss in a woman with NLRP7 mutation: not all molar pregnancies can be easily classified as either "partial" or "complete" hydatidiform moles. *International Journal of Gynecological Pathology,* 32**,** 399-405.

BRUNO, D. L., BURGESS, T., REN, H., NOURI, S., PERTILE, M. D., FRANCIS, D. I., NORRIS, F., KENNEY, B. K., SCHOUTEN, J., ANDY CHOO, K. H. & SLATER, H. R. 2006. High-throughput analysis of chromosome abnormality in spontaneous miscarriage using an MLPA subtelomere assay with an ancillary FISH test for polyploidy. *Am J Med Genet A,* 140**,** 2786-93.

BUSSEN, S., SUTTERLIN, M. & STECK, T. 1999. Endocrine abnormalities during the follicular phase in women with recurrent spontaneous abortion. *Hum Reprod,* 14**,** 18-20.

CAREY, L., NASH, B. M. & WRIGHT, D. C. 2015. Molecular genetic studies of complete hydatidiform moles. *Transl Pediatr,* 4**,** 181-8.

CARIBONI, A., PIMPINELLI, F., COLAMARINO, S., ZANINETTI, R., PICCOLELLA, M., RUMIO, C., PIVA, F., RUGARLI, E. I. & MAGGI, R. 2004. The product of X-linked Kallmann's syndrome gene (KAL1) affects the migratory activity of gonadotropin-releasing hormone (GnRH)-producing neurons. *Hum Mol Genet,* 13**,** 2781-91.

CARP, H., GUETTA, E., DORF, H., SORIANO, D., BARKAI, G. & SCHIFF, E. 2006. Embryonic karyotype in recurrent miscarriage with parental karyotypic aberrations. *Fertil Steril,* 85**,** 446-50.

CARSS, K. J., HILLMAN, S. C., PARTHIBAN, V., MCMULLAN, D. J., MAHER, E. R., KILBY, M. D. & HURLES, M. E. 2014. Exome sequencing improves genetic diagnosis of structural fetal abnormalities revealed by ultrasound. *Hum Mol Genet,* 23**,** 3269-77.

CARVALHO, B., DORIA, S., RAMALHO, C., BRANDAO, O., SOUSA, M., MATIAS, A., BARROS, A. & CARVALHO, F. 2010. Aneuploidies detection in miscarriages and fetal deaths using multiplex ligation-dependent probe amplification: an alternative for speeding up results? *Eur J Obstet Gynecol Reprod Biol,* 153**,** 151-5.

CATALANO, P. M. & SHANKAR, K. 2017. Obesity and pregnancy: mechanisms of short term and long term adverse consequences for mother and child. *Bmj,* 356**,** j1.

CHAN, K. C., DING, C., GEROVASSILI, A., YEUNG, S. W., CHIU, R. W., LEUNG, T. N., LAU, T. K., CHIM, S. S., CHUNG, G. T., NICOLAIDES, K. H. & LO, Y. M. 2006. Hypermethylated RASSF1A in maternal plasma: A universal fetal DNA marker that improves the reliability of noninvasive prenatal diagnosis. *Clin Chem,* 52**,** 2211-8.

CHAN, K. C., ZHANG, J., HUI, A. B., WONG, N., LAU, T. K., LEUNG, T. N., LO, K. W., HUANG, D. W. & LO, Y. M. 2004. Size distributions of maternal and fetal DNA in maternal plasma. *Clin Chem,* 50**,** 88-92.

CHAN, Y. Y., JAYAPRAKASAN, K., ZAMORA, J., THORNTON, J. G., RAINE-FENNING, N. & COOMARASAMY, A. 2011. The prevalence of congenital uterine anomalies in unselected and high-risk populations: a systematic review. *Hum Reprod Update,* 17**,** 761-71.

CHIU, R. W., AKOLEKAR, R., ZHENG, Y. W., LEUNG, T. Y., SUN, H., CHAN, K. C., LUN, F. M., GO, A. T., LAU, E. T., TO, W. W., LEUNG, W. C., TANG, R. Y., AU-YEUNG, S. K., LAM, H., KUNG, Y. Y., ZHANG, X., VAN VUGT, J. M., MINEKAWA, R., TANG, M. H., WANG, J., OUDEJANS, C. B., LAU, T. K., NICOLAIDES, K. H. & LO, Y. M. 2011. Non-invasive prenatal assessment of trisomy 21 by multiplexed maternal plasma DNA sequencing: large scale validity study. *Bmj,* 342**,** c7401.

CHIU, R. W., LAU, T. K., LEUNG, T. N., CHOW, K. C., CHUI, D. H. & LO, Y. M. 2002. Prenatal exclusion of beta thalassaemia major by examination of maternal plasma. *Lancet,* 360**,** 998-1000.

CHOI, M., SCHOLL, U. I., JI, W., LIU, T., TIKHONOVA, I. R., ZUMBO, P., NAYIR, A., BAKKALOGLU, A., OZEN, S., SANJAD, S., NELSON-WILLIAMS, C., FARHI, A., MANE, S. & LIFTON, R. P. 2009. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci U S A,* 106**,** 19096-101.

CHOI, T. Y., LEE, H. M., PARK, W. K., JEONG, S. Y. & MOON, H. S. 2014. Spontaneous abortion and recurrent miscarriage: A comparison of cytogenetic diagnosis in 250 cases. *Obstet Gynecol Sci,* 57**,** 518-25.

CHU, T., BUNCE, K., HOGGE, W. A. & PETERS, D. G. 2010. A novel approach toward the challenge of accurately quantifying fetal DNA in maternal plasma. *Prenat Diagn,* 30**,** 1226-9.

CLARK-GANHEART, C. A., FRIES, M. H., LEIFHEIT, K. M., JENSEN, T. J., MORENO-RUIZ, N. L., YE, P. P., JENNINGS, J. M. & DRIGGERS, R. W. 2015. Use of cell-free DNA in the investigation of intrauterine fetal demise and miscarriage. *Obstet Gynecol,* 125**,** 1321-9.

CLIFFORD, K., RAI, R., WATSON, H. & REGAN, L. 1994. An informative protocol for the investigation of recurrent miscarriage: preliminary experience of 500 consecutive cases. *Hum Reprod,* 9**,** 1328-32.

COCKSEDGE, K. A., SARAVELOS, S. H., WANG, Q., TUCKERMAN, E., LAIRD, S. M. & LI, T. C. 2008. Does free androgen index predict subsequent pregnancy outcome in women with recurrent miscarriage? *Hum Reprod,* 23**,** 797-802.

COLIN, E., HUYNH CONG, E., MOLLET, G., GUICHET, A., GRIBOUVAL, O., ARRONDEL, C., BOYER, O., DANIEL, L., GUBLER, M. C., EKINCI, Z., TSIMARATOS, M., CHABROL, B., BODDAERT, N., VERLOES, A., CHEVROLLIER, A., GUEGUEN, N., DESQUIRET-DUMAS, V., FERRE, M., PROCACCIO, V., RICHARD, L., FUNALOT, B., MONCLA, A., BONNEAU, D. & ANTIGNAC, C. 2014. Loss-of-function mutations in WDR73 are responsible for microcephaly and steroid-resistant nephrotic syndrome: Galloway-Mowat syndrome. *Am J Hum Genet,* 95**,** 637-48.

CRISTOFOLI, F., DE KEERSMAECKER, B., DE CATTE, L., VERMEESCH, J. R. & VAN ESCH, H. 2017. Novel STIL Compound Heterozygous Mutations Cause Severe Fetal Microcephaly and Centriolar Lengthening. *Molecular Syndromology.,* 27.

DALEY, R., HILL, M. & CHITTY, L. S. 2014. Non-invasive prenatal diagnosis: progress and potential. *Arch Dis Child Fetal Neonatal Ed,* 99**,** F426-30.

DARMOCHWAL-KOLARZ, D., LESZCZYNSKA-GORZELAK, B., ROLINSKI, J. & OLESZCZUK, J. 2002. The immunophenotype of patients with recurrent pregnancy loss. *Eur J Obstet Gynecol Reprod Biol,* 103**,** 53-7.

DEVANEY, S. A., PALOMAKI, G. E., SCOTT, J. A. & BIANCHI, D. W. 2011. Noninvasive fetal sex determination using cell-free fetal DNA: a systematic review and meta-analysis. *Jama,* 306**,** 627-36.

DIEGO-ALVAREZ, D., RODRIGUEZ DE ALBA, M., CARDERO-MERLO, R., DIAZ-RECASENS, J., AYUSO, C., RAMOS, C. & LORDA-SANCHEZ, I. 2007. MLPA as a screening method of aneuploidy and unbalanced chromosomal rearrangements in spontaneous miscarriages. *Prenat Diagn,* 27**,** 765-71.

DOCHERTY, L. E., REZWAN, F. I., POOLE, R. L., TURNER, C. L., KIVUVA, E., MAHER, E. R., SMITHSON, S. F., HAMILTON-SHIELD, J. P., PATALAN, M., GIZEWSKA, M., PEREGUD-POGORZELSKI, J., BEYGO, J., BUITING, K., HORSTHEMKE, B., SOELLNER, L., BEGEMANN, M., EGGERMANN, T., BAPLE, E., MANSOUR, S., TEMPLE, I. K. & MACKAY, D. J. 2015. Mutations in NLRP5 are associated with reproductive wastage and multilocus imprinting disorders in humans. *Nature communications,* 6**,** 8086.

DOHRN, N., LE, V. Q., PETERSEN, A., SKOVBO, P., PEDERSEN, I. S., ERNST, A., KRARUP, H. & PETERSEN, M. B. 2015. ECEL1 mutation causes fetal arthrogryposis multiplex congenita. *Am J Med Genet A,* 167a**,** 731-43.

DONAGHUE, C., DAVIES, N., AHN, J. W., THOMAS, H., OGILVIE, C. M. & MANN, K. 2017. Efficient and cost-effective genetic analysis of products of conception and fetal tissues using a QF-PCR/array CGH strategy; five years of data. *Mol Cytogenet,* 10**,** 12.

DONAGHUE, C., MANN, K., DOCHERTY, Z., MAZZASCHI, R., FEAR, C. & OGILVIE, C. 2010. Combined QF-PCR and MLPA molecular analysis of miscarriage products: an efficient and robust alternative to karyotype analysis. *Prenat Diagn,* 30**,** 133-7.

DONDERS, G. G., VAN CALSTEREN, K., BELLEN, G., REYBROUCK, R., VAN DEN BOSCH, T., RIPHAGEN, I. & VAN LIERDE, S. 2009. Predictive value for preterm birth of abnormal vaginal flora, bacterial vaginosis and aerobic vaginitis during the first trimester of pregnancy. *Bjog,* 116**,** 1315-24.

DONG, Z., YAN, J., XU, F., YUAN, J., JIANG, H., WANG, H., CHEN, H., ZHANG, L., YE, L., XU, J., SHI, Y., YANG, Z., CAO, Y., CHEN, L., LI, Q., ZHAO, X., LI, J., CHEN, A., ZHANG, W., WONG, H. G., QIN,

Y., ZHAO, H., CHEN, Y., LI, P., MA, T., WANG, W. J., KWOK, Y. K., JIANG, Y., PURSLEY, A. N., CHUNG, J. P. W., HONG, Y., KRISTIANSEN, K., YANG, H., PIÑA-AGUILAR, R. E., LEUNG, T. Y., CHEUNG, S. W., MORTON, C. C., CHOY, K. W. & CHEN, Z. J. 2019. Genome Sequencing Explores Complexity of Chromosomal Abnormalities in Recurrent Miscarriage. *Am J Hum Genet,* 105**,** 1102-1111.

DRURY, S., HILL, M. & CHITTY, L. S. 2016. Cell-Free Fetal DNA Testing for Prenatal Diagnosis. *Advances in Clinical Chemistry,* 76**,** 1-35.

DRURY, S., WILLIAMS, H., TRUMP, N., BOUSTRED, C., LENCH, N., SCOTT, R. H. & CHITTY, L. S. 2015. Exome sequencing for prenatal diagnosis of fetuses with sonographic abnormalities. *Prenat Diagn,* 35**,** 1010-7.

EHRICH, M., DECIU, C., ZWIEFELHOFER, T., TYNAN, J. A., CAGASAN, L., TIM, R., LU, V., MCCULLOUGH, R., MCCARTHY, E., NYGREN, A. O., DEAN, J., TANG, L., HUTCHISON, D., LU, T., WANG, H., ANGKACHATCHAI, V., OETH, P., CANTOR, C. R., BOMBARD, A. & VAN DEN BOOM, D. 2011. Noninvasive detection of fetal trisomy 21 by sequencing of DNA in maternal blood: a study in a clinical setting. *Am J Obstet Gynecol,* 204**,** 205.e1-11.

EIBEN, B., BARTELS, I., BAHR-PORSCH, S., BORGMANN, S., GATZ, G., GELLERT, G., GOEBEL, R., HAMMANS, W., HENTEMANN, M., OSMERS, R. & ET AL. 1990. Cytogenetic analysis of 750 spontaneous abortions with the direct-preparation method of chorionic villi and its implications for studying genetic causes of pregnancy wastage. *Am J Hum Genet,* 47**,** 656-63.

ELLARD, S., KIVUVA, E., TURNPENNY, P., STALS, K., JOHNSON, M., XIE, W., CASWELL, R. & LANGO ALLEN, H. 2015. An exome sequencing strategy to diagnose lethal autosomal recessive disorders. *European Journal of Human Genetics,* 23**,** 401-404.

ELLARD, S. B., E.; BERRY, I.; FORRESTER, N.; TURNBULL, C.; OWENS,; M. ECCLES,; D. ABBS,; S. SCOTT,; R. DEANS, Z.; LESTER, T.; CAMPBELL, J.; NEWMAN, W.; MCMULLAN, D. 2019. ACGS Best Practise Guidelines for Variant Classification 2019. *Association for CLinical Genomic Science*.

FAN, H. C., BLUMENFELD, Y. J., CHITKARA, U., HUDGINS, L. & QUAKE, S. R. 2008. Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. *Proc Natl Acad Sci U S A,* 105**,** 16266-71.

FAN, H. C., BLUMENFELD, Y. J., CHITKARA, U., HUDGINS, L. & QUAKE, S. R. 2010. Analysis of the size distributions of fetal and maternal cell-free DNA by paired-end sequencing. *Clin Chem,* 56**,** 1279-86.

FILGES, I., NOSOVA, E., BRUDER, E., TERCANLI, S., TOWNSEND, K., GIBSON, W. T., ROTHLISBERGER, B., HEINIMANN, K., HALL, J. G., GREGORY-EVANS, C. Y., WASSERMAN, W. W., MINY, P. & FRIEDMAN, J. M. 2014. Exome sequencing identifies mutations in KIF14 as a novel cause of an autosomal recessive lethal fetal ciliopathy phenotype. *Clinical Genetics,* 86**,** 220-8.

FINNING, K. M., MARTIN, P. G., SOOTHILL, P. W. & AVENT, N. D. 2002. Prediction of fetal D status from maternal plasma: introduction of a new noninvasive fetal RHD genotyping service. *Transfusion,* 42**,** 1079-85.

FLORY, M. R., MOSER, M. J., MONNAT, R. J., JR. & DAVIS, T. N. 2000. Identification of a human centrosomal calmodulin-binding protein that shares homology with pericentrin. *Proc Natl Acad Sci U S A,* 97**,** 5919-23.

FRANSSEN, M. T., MUSTERS, A. M., VAN DER VEEN, F., REPPING, S., LESCHOT, N. J., BOSSUYT, P. M., GODDIJN, M. & KOREVAAR, J. C. 2011. Reproductive outcome after PGD in couples with recurrent miscarriage carrying a structural chromosome abnormality: a systematic review. *Hum Reprod Update,* 17**,** 467-75.

GERULL, B., HEUSER, A., WICHTER, T., PAUL, M., BASSON, C. T., MCDERMOTT, D. A., LERMAN, B. B., MARKOWITZ, S. M., ELLINOR, P. T., MACRAE, C. A., PETERS, S., GROSSMANN, K. S., DRENCKHAHN, J., MICHELY, B., SASSE-KLAASSEN, S., BIRCHMEIER, W., DIETZ, R., BREITHARDT, G., SCHULZE-BAHR, E. & THIERFELDER, L. 2004. Mutations in the desmosomal

protein plakophilin-2 are common in arrhythmogenic right ventricular cardiomyopathy. *Nat Genet,* 36**,** 1162-4.
GIAKOUMELOU, S., WHEELHOUSE, N., CUSCHIERI, K., ENTRICAN, G., HOWIE, S. E. & HORNE, A. W. 2016. The role of infection in miscarriage. *Hum Reprod Update,* 22**,** 116-33.
GODDIJN, M., JOOSTEN, J. H., KNEGT, A. C., VAN DERVEEN, F., FRANSSEN, M. T., BONSEL, G. J. & LESCHOT, N. J. 2004. Clinical relevance of diagnosing structural chromosome abnormalities in couples with repeated miscarriage. *Hum Reprod,* 19**,** 1013-7.
GODDIJN, M. & LESCHOT, N. J. 2000. Genetic aspects of miscarriage. *Baillieres Best Pract Res Clin Obstet Gynaecol,* 14**,** 855-65.
GOHRING, I., TAGARIELLO, A., ENDELE, S., STOLT, C. C., GHASSIBE, M., FISHER, M., THIEL, C. T., TRAUTMANN, U., VIKKULA, M., WINTERPACHT, A., FITZPATRICK, D. R. & RAUCH, A. 2010. Disruption of ST5 is associated with mental retardation and multiple congenital anomalies. *J Med Genet,* 47**,** 91-8.
GONZALEZ-GONZALEZ, M. C., GARCIA-HOYOS, M., TRUJILLO, M. J., RODRIGUEZ DE ALBA, M., LORDA-SANCHEZ, I., DIAZ-RECASENS, J., GALLARDO, E., AYUSO, C. & RAMOS, C. 2002. Prenatal detection of a cystic fibrosis mutation in fetal DNA from maternal plasma. *Prenat Diagn,* 22**,** 946-8.
GROSSMANN, K. S., GRUND, C., HUELSKEN, J., BEHREND, M., ERDMANN, B., FRANKE, W. W. & BIRCHMEIER, W. 2004. Requirement of plakophilin 2 for heart morphogenesis and cardiac junction formation. *J Cell Biol,* 167**,** 149-60.
GUIBERT, J., BENACHI, A., GREBILLE, A. G., ERNAULT, P., ZORN, J. R. & COSTA, J. M. 2003. Kinetics of SRY gene appearance in maternal serum: detection by real time PCR in early pregnancy after assisted reproductive technique. *Hum Reprod,* 18**,** 1733-6.
HARDELIN, J. P., LEVILLIERS, J., BLANCHARD, S., CAREL, J. C., LEUTENEGGER, M., PINARD-BERTELLETTO, J. P., BOULOUX, P. & PETIT, C. 1993. Heterogeneity in the mutations responsible for X chromosome-linked Kallmann syndrome. *Hum Mol Genet,* 2**,** 373-7.
HARDY, K. & HARDY, P. J. 2015. 1(st) trimester miscarriage: four decades of study. *Transl Pediatr,* 4**,** 189-200.
HASSOLD, T., CHEN, N., FUNKHOUSER, J., JOOSS, T., MANUEL, B., MATSUURA, J., MATSUYAMA, A., WILSON, C., YAMANE, J. A. & JACOBS, P. A. 1980. A cytogenetic study of 1000 spontaneous abortions. *Ann Hum Genet,* 44**,** 151-78.
HASSOLD, T. & CHIU, D. 1985. Maternal age-specific rates of numerical chromosome abnormalities with special reference to trisomy. *Hum Genet,* 70**,** 11-7.
HESTAND, M. S., BESSEM, M., VAN RIJN, P., DE MENEZES, R. X. & SIE, D. 2019. Fetal fraction evaluation in non-invasive prenatal screening (NIPS). 27**,** 198-202.
HILL, M., TAFFINDER, S., CHITTY, L. S. & MORRIS, S. 2011. Incremental cost of non-invasive prenatal diagnosis versus invasive prenatal diagnosis of fetal sex in England. *Prenat Diagn,* 31**,** 267-73.
HIRSHFELD-CYTRON, J., SUGIURA-OGASAWARA, M. & STEPHENSON, M. D. 2011. Management of recurrent pregnancy loss associated with a parental carrier of a reciprocal translocation: a systematic review. *Semin Reprod Med,* 29**,** 470-81.
HUANG, J. Y., SU, M., LIN, S. H. & KUO, P. L. 2013. A genetic association study of NLRP2 and NLRP7 genes in idiopathic recurrent miscarriage. *Hum Reprod,* 28**,** 1127-34.
HUDECOVA, I., SAHOTA, D., HEUNG, M. M., JIN, Y., LEE, W. S., LEUNG, T. Y., LO, Y. M. & CHIU, R. W. 2014. Maternal plasma fetal DNA fractions in pregnancies with low and high risks for fetal chromosomal aneuploidies. *PLoS One,* 9**,** e88484.
HUI, R., CAMPBELL, D. H., LEE, C. S., MCCAUL, K., HORSFALL, D. J., MUSGROVE, E. A., DALY, R. J., SESHADRI, R. & SUTHERLAND, R. L. 1997. EMS1 amplification can occur independently of CCND1 or INT-2 amplification at 11q13 and may identify different phenotypes in primary breast cancer. *Oncogene,* 15**,** 1617-23.

HUNG, L. Y., TANG, C. J. & TANG, T. K. 2000. Protein 4.1 R-135 interacts with a novel centrosomal protein (CPAP) which is associated with the gamma-tubulin complex. *Mol Cell Biol,* 20**,** 7813-25.

ILLUMINA. 2017. *Noninvasive Prenatal Testing (NIPT)* [Online]. Available: https://www.illumina.com/clinical/reproductive-genetic-health/nipt.html [Accessed 18/05/2017.

JENDERNY, J. 2014. Chromosome aberrations in a large series of spontaneous miscarriages in the German population and review of the literature. *Mol Cytogenet,* 7**,** 38.

JIANG, P., CHAN, K. C., LIAO, G. J., ZHENG, Y. W., LEUNG, T. Y., CHIU, R. W., LO, Y. M. & SUN, H. 2012. FetalQuant: deducing fractional fetal DNA concentration from massively parallel sequencing of DNA in maternal plasma. *Bioinformatics,* 28**,** 2883-90.

JINKS, R. N., PUFFENBERGER, E. G., BAPLE, E., HARDING, B., CRINO, P., FOGO, A. B., WENGER, O., XIN, B., KOEHLER, A. E., MCGLINCY, M. H., PROVENCHER, M. M., SMITH, J. D., TRAN, L., AL TURKI, S., CHIOZA, B. A., CROSS, H., HARLALKA, G. V., HURLES, M. E., MAROOFIAN, R., HEAPS, A. D., MORTON, M. C., STEMPAK, L., HILDEBRANDT, F., SADOWSKI, C. E., ZARITSKY, J., CAMPELLONE, K., MORTON, D. H., WANG, H., CROSBY, A. & STRAUSS, K. A. 2015. Recessive nephrocerebellar syndrome on the Galloway-Mowat syndrome spectrum is caused by homozygous protein-truncating mutations of WDR73. *Brain,* 138**,** 2173-90.

JOERGENSEN, M. W., NIEMANN, I., RASMUSSEN, A. A., HINDKJAER, J., AGERHOLM, I., BOLUND, L. & SUNDE, L. 2014. Triploid pregnancies: genetic and clinical features of 158 cases. *Am J Obstet Gynecol,* 211**,** 370.e1-19.

KALOUSEK, D. K. & DILL, F. J. 1983. Chromosomal mosaicism confined to the placenta in human conceptions. *Science,* 221**,** 665-7.

KANG, M. Y., CHO, I. A. & BAEK, J. C. *Clin Case Rep*.

KANTAPUTRA, P., TANPAIBOON, P., PORNTAVEETUS, T., OHAZAMA, A., SHARPE, P., RAUCH, A., HUSSADALOY, A. & THIEL, C. T. 2011. The smallest teeth in the world are caused by mutations in the PCNT gene. *Am J Med Genet A,* 155a**,** 1398-403.

KAPRANOS, N. C. & KOTRONIAS, D. C. 2009. Detection of herpes simplex virus in first trimester pregnancy loss using molecular techniques. *In Vivo,* 23**,** 839-42.

KARIM, S., JAMAL, H. S., ROUZI, A., ARDAWI, M. S. M., SCHULTEN, H. J., MIRZA, Z., ALANSARI, N. A., AL-QUAITI, M. M., ABUSAMRA, H., NASEER, M. I., TURKI, R., CHAUDHARY, A. G., GARI, M., ABUZENADAH, A. M. & AL-QHATANI, M. H. 2017. Genomic answers for recurrent spontaneous abortion in Saudi Arabia: An array comparative genomic hybridization approach. *Reprod Biol,* 17**,** 133-143.

KASER, D. 2018. The Status of Genetic Screening in Recurrent Pregnancy Loss. *Obstet Gynecol Clin North Am,* 45**,** 143-154.

KESMODEL, U., WISBORG, K., OLSEN, S. F., HENRIKSEN, T. B. & SECHER, N. J. 2002. Moderate alcohol intake in pregnancy and the risk of spontaneous abortion. *Alcohol Alcohol,* 37**,** 87-92.

KHALID, A. S., JOYCE, C. & O'DONOGHUE, K. 2013. Prevalence of subclinical and undiagnosed overt hypothyroidism in a pregnancy loss clinic. *Ir Med J,* 106**,** 107-10.

KIM, J. & KIM, J. 2019. PCNT is critical for the association and conversion of centrioles to centrosomes during mitosis. 132.

KIM, S. K., HANNUM, G., GEIS, J., TYNAN, J., HOGG, G., ZHAO, C., JENSEN, T. J., MAZLOOM, A. R., OETH, P., EHRICH, M., VAN DEN BOOM, D. & DECIU, C. 2015. Determination of fetal DNA fraction from the plasma of pregnant women using sequence read counts. *Prenat Diagn,* 35**,** 810-5.

KING, A., HIBY, S. E., GARDNER, L., JOSEPH, S., BOWEN, J. M., VERMA, S., BURROWS, T. D. & LOKE, Y. W. 2000. Recognition of trophoblast HLA class I molecules by decidual NK cell receptors--a review. *Placenta,* 21 Suppl A**,** S81-5.

KOWALIK, C. R., GODDIJN, M., EMANUEL, M. H., BONGERS, M. Y., SPINDER, T., DE KRUIF, J. H., MOL, B. W. & HEINEMAN, M. J. 2011. Metroplasty versus expectant management for women with recurrent miscarriage and a septate uterus. *Cochrane Database Syst Rev*, Cd008576.

KRISHNA, I., BADELL, M., LOUCKS, T. L., LINDSAY, M. & SAMUEL, A. 2016. Adverse perinatal outcomes are more frequent in pregnancies with a low fetal fraction result on noninvasive prenatal testing. *Prenat Diagn,* 36**,** 210-5.

KWAK-KIM, J. Y., CHUNG-BANG, H. S., NG, S. C., NTRIVALAS, E. I., MANGUBAT, C. P., BEAMAN, K. D., BEER, A. E. & GILMAN-SACHS, A. 2003. Increased T helper 1 cytokine responses by circulating T cells are present in women with recurrent pregnancy losses and in infertile women with multiple implantation failures after IVF. *Hum Reprod,* 18**,** 767-73.

KWAK, J. Y., BEAMAN, K. D., GILMAN-SACHS, A., RUIZ, J. E., SCHEWITZ, D. & BEER, A. E. 1995. Up-regulated expression of CD56+, CD56+/CD16+, and CD19+ cells in peripheral blood lymphocytes in pregnant women with recurrent pregnancy losses. *Am J Reprod Immunol,* 34**,** 93-9.

KWAK, J. Y., BEER, A. E., KIM, S. H. & MANTOUVALOS, H. P. 1999. Immunopathology of the implantation site utilizing monoclonal antibodies to natural killer cells in women with recurrent pregnancy losses. *Am J Reprod Immunol,* 41**,** 91-8.

LACHAPELLE, M. H., MIRON, P., HEMMINGS, R. & ROY, D. C. 1996. Endometrial T, B, and NK cells in patients with recurrent spontaneous abortion. Altered profile and pregnancy outcome. *J Immunol,* 156**,** 4027-34.

LAM, K. W., JIANG, P., LIAO, G. J., CHAN, K. C., LEUNG, T. Y., CHIU, R. W. & LO, Y. M. 2012. Noninvasive prenatal diagnosis of monogenic diseases by targeted massively parallel sequencing of maternal plasma: application to beta-thalassemia. *Clin Chem,* 58**,** 1467-75.

LARSEN, E. C., CHRISTIANSEN, O. B., KOLTE, A. M. & MACKLON, N. 2013a. New insights into mechanisms behind miscarriage. *BMC Med,* 11**,** 154.

LARSEN, E. C., CHRISTIANSEN, O. B., KOLTE, A. M. & MACKLON, N. 2013b. New insights into mechanisms behind miscarriage. [Review]. 1**,** 154.

LATA, K., DUTTA, P., SRIDHAR, S., ROHILLA, M., SRINIVASAN, A., PRASHAD, G. R., SHAH, V. N. & BHANSALI, A. 2013. Thyroid autoimmunity and obstetric outcomes in women with recurrent miscarriage: a case-control study. *Endocr Connect,* 2**,** 118-24.

LAVIALLE, C., CORNELIS, G., DUPRESSOIR, A., ESNAULT, C., HEIDMANN, O., VERNOCHET, C. & HEIDMANN, T. 2013. Paleovirology of 'syncytins', retroviral env genes exapted for a role in placentation. *Philos Trans R Soc Lond B Biol Sci,* 368**,** 20120507.

LE GALLO, M., O'HARA, A. J., RUDD, M. L., URICK, M. E., HANSEN, N. F., O'NEIL, N. J., PRICE, J. C., ZHANG, S., ENGLAND, B. M., GODWIN, A. K., SGROI, D. C., HIETER, P., MULLIKIN, J. C., MERINO, M. J. & BELL, D. W. 2012. Exome sequencing of serous endometrial tumors identifies recurrent somatic mutations in chromatin-remodeling and ubiquitin ligase complex genes. *Nat Genet,* 44**,** 1310-5.

LEAL, G. F., ROBERTS, E., SILVA, E. O., COSTA, S. M., HAMPSHIRE, D. J. & WOODS, C. G. 2003. A novel locus for autosomal recessive primary microcephaly (MCPH6) maps to 13q12.2. *J Med Genet,* 40**,** 540-2.

LEVY, B., SIGURJONSSON, S., PETTERSEN, B., MAISENBACHER, M. K., HALL, M. P., DEMKO, Z., LATHI, R. B., TAO, R., AGGARWAL, V. & RABINOWITZ, M. 2014. Genomic imbalance in products of conception: single-nucleotide polymorphism chromosomal microarray analysis. *Obstet Gynecol,* 124**,** 202-9.

LI, Y., ZIMMERMANN, B., RUSTERHOLZ, C., KANG, A., HOLZGREVE, W. & HAHN, S. 2004. Size separation of circulatory DNA in maternal plasma permits ready detection of fetal DNA polymorphisms. *Clin Chem,* 50**,** 1002-11.

LEINFINGER, P.; Lohmann, L.; Luscan, A.; Trost, D.; Bidat, L.; Debarge, V.; Castaigne, V.; Senat, M.-V.;

Brechard, M.-P.; Guilbaud, L.; et al. Strategy for Use of Genome-Wide Non-Invasive Prenatal Testing for Rare Autosomal Aneuploidies and Unbalanced Structural Chromosomal Anomalies. J. Clin. Med. 2020, 9, 2466.

LO, K. K., BOUSTRED, C., CHITTY, L. S. & PLAGNOL, V. 2014. RAPIDR: an analysis package for non-invasive prenatal testing of aneuploidy. *Bioinformatics,* 30**,** 2965-7.

LO, Y. M., CHAN, K. C., SUN, H., CHEN, E. Z., JIANG, P., LUN, F. M., ZHENG, Y. W., LEUNG, T. Y., LAU, T. K., CANTOR, C. R. & CHIU, R. W. 2010. Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus. *Sci Transl Med,* 2**,** 61ra91.

LO, Y. M., CORBETTA, N., CHAMBERLAIN, P. F., RAI, V., SARGENT, I. L., REDMAN, C. W. & WAINSCOAT, J. S. 1997. Presence of fetal DNA in maternal plasma and serum. *Lancet,* 350**,** 485-7.

LO, Y. M., TEIN, M. S., LAU, T. K., HAINES, C. J., LEUNG, T. N., POON, P. M., WAINSCOAT, J. S., JOHNSON, P. J., CHANG, A. M. & HJELM, N. M. 1998. Quantitative analysis of fetal DNA in maternal plasma and serum: implications for noninvasive prenatal diagnosis. *Am J Hum Genet,* 62**,** 768-75.

LO, Y. M., ZHANG, J., LEUNG, T. N., LAU, T. K., CHANG, A. M. & HJELM, N. M. 1999. Rapid clearance of fetal DNA from maternal plasma. *Am J Hum Genet,* 64**,** 218-24.

LOPEZ-CARRASCO, A., OLTRA, S., MONFORT, S., MAYO, S., ROSELLO, M., MARTINEZ, F. & ORELLANA, C. 2013. Mutation screening of AURKB and SYCP3 in patients with reproductive problems. *Mol Hum Reprod,* 19**,** 102-8.

LORD, J., MCMULLAN, D. J., EBERHARDT, R. Y., RINCK, G., HAMILTON, S. J., QUINLAN-JONES, E., PRIGMORE, E., KEELAGHER, R., BEST, S. K., CAREY, G. K., MELLIS, R., ROBART, S., BERRY, I. R., CHANDLER, K. E., CILLIERS, D., CRESSWELL, L., EDWARDS, S. L., GARDINER, C., HENDERSON, A., HOLDEN, S. T., HOMFRAY, T., LESTER, T., LEWIS, R. A., NEWBURY-ECOB, R., PRESCOTT, K., QUARRELL, O. W., RAMSDEN, S. C., ROBERTS, E., TAPON, D., TOOLEY, M. J., VASUDEVAN, P. C., WEBER, A. P., WELLESLEY, D. G., WESTWOOD, P., WHITE, H., PARKER, M., WILLIAMS, D., JENKINS, L., SCOTT, R. H., KILBY, M. D., CHITTY, L. S., HURLES, M. E. & MAHER, E. R. 2019. Prenatal exome sequencing analysis in fetal structural anomalies detected by ultrasonography (PAGE): a cohort study. *Lancet,* 393**,** 747-757.

LUN, F. M., CHIU, R. W., SUN, K., LEUNG, T. Y., JIANG, P., CHAN, K. C., SUN, H. & LO, Y. M. 2013. Noninvasive prenatal methylomic analysis by genomewide bisulfite sequencing of maternal plasma DNA. *Clin Chem,* 59**,** 1583-94.

LUO, M. L., SHEN, X. M., ZHANG, Y., WEI, F., XU, X., CAI, Y., ZHANG, X., SUN, Y. T., ZHAN, Q. M., WU, M. & WANG, M. R. 2006. Amplification and overexpression of CTTN (EMS1) contribute to the metastasis of esophageal squamous cell carcinoma by promoting cell migration and anoikis resistance. *Cancer Res,* 66**,** 11690-9.

MA, D., GE, H., LI, X., JIANG, T., CHEN, F., ZHANG, Y., HU, P., CHEN, S., ZHANG, J., JI, X., XU, X., JIANG, H., CHEN, M., WANG, W. & XU, Z. 2014. Haplotype-based approach for noninvasive prenatal diagnosis of congenital adrenal hyperplasia by maternal plasma DNA sequencing. *Gene,* 544**,** 252-8.

MACKIE, F. L., HEMMING, K., ALLEN, S., MORRIS, R. K. & KILBY, M. D. 2017. The accuracy of cell-free fetal DNA-based non-invasive prenatal testing in singleton pregnancies: a systematic review and bivariate meta-analysis. *Bjog,* 124**,** 32-46.

MACKLON, N. S., GERAEDTS, J. P. & FAUSER, B. C. 2002. Conception to ongoing pregnancy: the 'black box' of early pregnancy loss. *Hum Reprod Update,* 8**,** 333-43.

MANOKHINA, I., HANNA, C. W., STEPHENSON, M. D., MCFADDEN, D. E. & ROBINSON, W. P. 2013. Maternal NLRP7 and C6orf221 variants are not a common risk factor for androgenetic moles, triploidy and recurrent miscarriage. *Molecular Human Reproduction,* 19**,** 539-544.

MCCLELLAND, L. S., ALLEN, S. K., LARKINS, S. A., HAMILTON, S. J., MARTON, T., COX, P. M., HARGITAI, B., JOHNSTON, E. H., MORGAN, C. & HARDY, G. 2011. Implementation and experience of an

alternative QF-PCR and MLPA diagnostic strategy to detect chromosomal abnormalities in fetal and neonatal pathology samples. *Pediatr Dev Pathol,* 14**,** 460-8.

MCKIE, A. B., ALSAEDI, A., VOGT, J., STUURMAN, K. E., WEISS, M. M., SHAKEEL, H., TEE, L., MORGAN, N. V., NIKKELS, P. G., VAN HAAFTEN, G., PARK, S. M., VAN DER SMAGT, J. J., BUGIANI, M. & MAHER, E. R. 2014. Germline mutations in RYR1 are associated with foetal akinesia deformation sequence/lethal multiple pterygium syndrome. *Acta Neuropathologica Communications,* 2**,** 148.

MENASHA, J., LEVY, B., HIRSCHHORN, K. & KARDON, N. B. 2005. Incidence and spectrum of chromosome abnormalities in spontaneous abortions: new insights from a 12-year study. *Genet Med,* 7**,** 251-63.

MENG, M., LI, X., GE, H., CHEN, F., HAN, M., ZHANG, Y., KANG, D., XIE, W., GAO, Z., PAN, X., DAI, P., CHI, F., CHEN, S., LIU, P., ZHANG, C., CAO, J., JIANG, H., XU, X., WANG, W. & DUAN, T. 2014. Noninvasive prenatal testing for autosomal recessive conditions by maternal plasma sequencing in a case of congenital deafness. *Genet Med,* 16**,** 972-6.

METWALLY, M., ONG, K. J., LEDGER, W. L. & LI, T. C. 2008. Does high body mass index increase the risk of miscarriage after spontaneous and assisted conception? A meta-analysis of the evidence. *Fertil Steril,* 90**,** 714-26.

METWALLY, M., SARAVELOS, S. H., LEDGER, W. L. & LI, T. C. 2010. Body mass index and risk of miscarriage in women with recurrent miscarriage. *Fertil Steril,* 94**,** 290-5.

MILLAR, L. K., WING, D. A., LEUNG, A. S., KOONINGS, P. P., MONTORO, M. N. & MESTMAN, J. H. 1994. Low birth weight and preeclampsia in pregnancies complicated by hyperthyroidism. *Obstet Gynecol,* 84**,** 946-9.

MILLER, D. T., ADAM, M. P., ARADHYA, S., BIESECKER, L. G., BROTHMAN, A. R., CARTER, N. P., CHURCH, D. M., CROLLA, J. A., EICHLER, E. E., EPSTEIN, C. J., FAUCETT, W. A., FEUK, L., FRIEDMAN, J. M., HAMOSH, A., JACKSON, L., KAMINSKY, E. B., KOK, K., KRANTZ, I. D., KUHN, R. M., LEE, C., OSTELL, J. M., ROSENBERG, C., SCHERER, S. W., SPINNER, N. B., STAVROPOULOS, D. J., TEPPERBERG, J. H., THORLAND, E. C., VERMEESCH, J. R., WAGGONER, D. J., WATSON, M. S., MARTIN, C. L. & LEDBETTER, D. H. 2010. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet,* 86**,** 749-64.

MISSION, J. F., MARSHALL, N. E. & CAUGHEY, A. B. 2015. Pregnancy risks associated with obesity. *Obstet Gynecol Clin North Am,* 42**,** 335-53.

NAGIRNAJA, L., PALTA, P., KASAK, L., RULL, K., CHRISTIANSEN, O. B., NIELSEN, H. S., STEFFENSEN, R., ESKO, T., REMM, M. & LAAN, M. 2014. Structural genomic variation as risk factor for idiopathic recurrent miscarriage. *Human Mutation,* 35**,** 972-82.

NAMBA, T., TIAN, F., CHU, K., HWANG, S. Y., YOON, K. W., BYUN, S., HIRAKI, M., MANDINOVA, A. & LEE, S. W. 2013. CDIP1-BAP31 complex transduces apoptotic signals from endoplasmic reticulum to mitochondria under endoplasmic reticulum stress. *Cell Rep,* 5**,** 331-9.

NATERA. 2016. *Panorama natera prenatal screen* [Online]. Available: http://www.panoramatest.com/panorama-test.

NEW, M. I., TONG, Y. K., YUEN, T., JIANG, P., PINA, C., CHAN, K. C., KHATTAB, A., LIAO, G. J., YAU, M., KIM, S. M., CHIU, R. W., SUN, L., ZAIDI, M. & LO, Y. M. 2014. Noninvasive prenatal diagnosis of congenital adrenal hyperplasia using cell-free fetal DNA in maternal plasma. *J Clin Endocrinol Metab,* 99**,** E1022-30.

NICOLAIDES, K. H., SYNGELAKI, A., DEL MAR GIL, M., QUEZADA, M. S. & ZINEVICH, Y. 2014. Prenatal detection of fetal triploidy from cell-free DNA testing in maternal blood. *Fetal Diagn Ther,* 35**,** 212-7.

NTRIVALAS, E. I., KWAK-KIM, J. Y., GILMAN-SACHS, A., CHUNG-BANG, H., NG, S. C., BEAMAN, K. D., MANTOUVALOS, H. P. & BEER, A. E. 2001. Status of peripheral blood natural killer cells in

women with recurrent spontaneous abortions and infertility of unknown aetiology. *Hum Reprod,* 16**,** 855-61.

OEPKES, D., PAGE-CHRISTIAENS, G. C., BAX, C. J., BEKKER, M. N., BILARDO, C. M., BOON, E. M., SCHURING-BLOM, G. H., COUMANS, A. B., FAAS, B. H., GALJAARD, R. H., GO, A. T., HENNEMAN, L., MACVILLE, M. V., PAJKRT, E., SUIJKERBUIJK, R. F., HUIJSDENS-VAN AMSTERDAM, K., VAN OPSTAL, D., VERWEIJ, E. J., WEISS, M. M. & SISTERMANS, E. A. 2016. Trial by Dutch laboratories for evaluation of non-invasive prenatal testing. Part I-clinical impact. *Prenat Diagn,* 36**,** 1083-1090.

OLIVEIRA, L. M., SEMINARA, S. B., BERANOVA, M., HAYES, F. J., VALKENBURGH, S. B., SCHIPANI, E., COSTA, E. M., LATRONICO, A. C., CROWLEY, W. F., JR. & VALLEJO, M. 2001. The importance of autosomal genes in Kallmann syndrome: genotype-phenotype correlations and neuroendocrine characteristics. *J Clin Endocrinol Metab,* 86**,** 1532-8.

PAPAGEORGIOU, E. A., FIEGLER, H., RAKYAN, V., BECK, S., HULTEN, M., LAMNISSOU, K., CARTER, N. P. & PATSALIS, P. C. 2009. Sites of differential DNA methylation between placenta and peripheral blood: molecular markers for noninvasive prenatal diagnosis of aneuploidies. *Am J Pathol,* 174**,** 1609-18.

PAPAGEORGIOU, E. A., KARAGRIGORIOU, A., TSALIKI, E., VELISSARIOU, V., CARTER, N. P. & PATSALIS, P. C. 2011. Fetal-specific DNA methylation ratio permits noninvasive prenatal diagnosis of trisomy 21. *Nat Med,* 17**,** 510-3.

PAPAGEORGIOU, E. A. & PATSALIS, P. C. 2012. Non-invasive prenatal diagnosis of aneuploidies: new technologies and clinical applications. *Genome Med,* 4**,** 46.

PARK, J. E. & PARK, J. K. 2019. Counting-based cell-free DNA screening test fails to identify triploidy-A case report. 7**,** 90-93.

PARRY, D. A., LOGAN, C. V., HAYWARD, B. E., SHIRES, M., LANDOLSI, H., DIGGLE, C., CARR, I., RITTORE, C., TOUITOU, I., PHILIBERT, L., FISHER, R. A., FALLAHIAN, M., HUNTRISS, J. D., PICTON, H. M., MALIK, S., TAYLOR, G. R., JOHNSON, C. A., BONTHRON, D. T. & SHERIDAN, E. G. 2011. Mutations causing familial biparental hydatidiform mole implicate c6orf221 as a possible regulator of genomic imprinting in the human oocyte. *American Journal of Human Genetics,* 89**,** 451-8.

PENG, X. L. & JIANG, P. 2017. Bioinformatics Approaches for Fetal DNA Fraction Estimation in Noninvasive Prenatal Testing. *Int J Mol Sci,* 18.

PERGAMENT, E., CUCKLE, H., ZIMMERMANN, B., BANJEVIC, M., SIGURJONSSON, S., RYAN, A., HALL, M. P., DODD, M., LACROUTE, P., STOSIC, M., CHOPRA, N., HUNKAPILLER, N., PROSEN, D. E., MCADOO, S., DEMKO, Z., SIDDIQUI, A., HILL, M. & RABINOWITZ, M. 2014. Single-nucleotide polymorphism-based noninvasive prenatal screening in a high-risk and low-risk cohort. *Obstet Gynecol,* 124**,** 210-8.

PERTILE, M.; Flowers, N.; Vavrek, D. P1-86. Performance of a genome-wide PCR-free, paired-end sequencing-based noninvasive prenatal screening test, VeriSeq NIPT Solution v2, in the detection of fetal chromosomal anomalies. Poster Abstracts of the ISPD 23rd International Conference on Prenatal Diagnosis and Therapy, Singapore, 7–11 September 2019. Prenat Diagn. 2020, 40 (Suppl. 1), 74–75.

POON, L. L., LEUNG, T. N., LAU, T. K., CHOW, K. C. & LO, Y. M. 2002. Differential DNA methylation between fetus and mother as a strategy for detecting fetal DNA in maternal plasma. *Clin Chem,* 48**,** 35-41.

QIAN, J., CHENG, Q., MURDOCH, S., XU, C., JIN, F., CHEBARO, W., ZHANG, X., GAO, H., ZHU, Y., SLIM, R. & XIE, X. 2011. The genetics of recurrent hydatidiform moles in China: correlations between NLRP7 mutations, molar genotypes and reproductive outcomes. *Molecular Human Reproduction,* 17**,** 612-9.

QIAO, Y., WEN, J., TANG, F., MARTELL, S., SHOMER, N., LEUNG, P. C. K., STEPHENSON, M. D. & RAJCAN-SEPAROVIC, E. 2016. Whole exome sequencing in recurrent early pregnancy loss. *Molecular Human Reproduction,* 22**,** 364-372.

QUINTERO-RONDEROS, P., MERCIER, E., FUKUDA, M., GONZALEZ, R., SUAREZ, C. F., PATARROYO, M. A., VAIMAN, D., GRIS, J. C. & LAISSUE, P. 2017. Novel genes and mutations in patients affected by recurrent pregnancy loss. *PLoS ONE,* 12 (10) (no pagination).

RAE, W., GAO, Y., BUNYAN, D., HOLDEN, S., GILMOUR, K., PATEL, S., WELLESLEY, D. & WILLIAMS, A. 2015. A novel FOXP3 mutation causing fetal akinesia and recurrent male miscarriages. *Clin Immunol,* 161**,** 284-5.

RAI, R., BACKOS, M., BAXTER, N., CHILCOTT, I. & REGAN, L. 2000a. Recurrent miscarriage--an aspirin a day? *Hum Reprod,* 15**,** 2220-3.

RAI, R., BACKOS, M., RUSHWORTH, F. & REGAN, L. 2000b. Polycystic ovaries and recurrent miscarriage--a reappraisal. *Hum Reprod,* 15**,** 612-5.

RAMAN, L., DHEEDENE, A., DE SMET, M., VAN DORPE, J. & MENTEN, B. 2019. WisecondorX: improved copy number detection for routine shallow whole-genome sequencing. *Nucleic Acids Res,* 47**,** 1605-1614.

RAMOND, F., JANIN, A., DI FILIPPO, S., CHANAVAT, V., CHALABREYSSE, L., ROUX-BUISSON, N., SANLAVILLE, D., TOURAINE, R. & MILLAT, G. 2017. Homozygous PKP2 deletion associated with neonatal left ventricle noncompaction. *Clin Genet,* 91**,** 126-130.

RAUCH, A., THIEL, C. T., SCHINDLER, D., WICK, U., CROW, Y. J., EKICI, A. B., VAN ESSEN, A. J., GOECKE, T. O., AL-GAZALI, L., CHRZANOWSKA, K. H., ZWEIER, C., BRUNNER, H. G., BECKER, K., CURRY, C. J., DALLAPICCOLA, B., DEVRIENDT, K., DORFLER, A., KINNING, E., MEGARBANE, A., MEINECKE, P., SEMPLE, R. K., SPRANGER, S., TOUTAIN, A., TREMBATH, R. C., VOSS, E., WILSON, L., HENNEKAM, R., DE ZEGHER, F., DORR, H. G. & REIS, A. 2008. Mutations in the pericentrin (PCNT) gene cause primordial dwarfism. *Science,* 319**,** 816-9.

REGAN, L., BACKOS, M. & RAI, R. 2011. RCOG Green-top Guideline No.17: The Investigation and Treatment of Couples with Recurrent First-trimester and Second-Trimester Miscarriage.

REGAN, L., OWEN, E. J. & JACOBS, H. S. 1990. Hypersecretion of luteinising hormone, infertility, and miscarriage. *Lancet,* 336**,** 1141-4.

RICHARDS, S., AZIZ, N., BALE, S., BICK, D., DAS, S., GASTIER-FOSTER, J., GRODY, W. W., HEGDE, M., LYON, E., SPECTOR, E., VOELKERDING, K. & REHM, H. L. 2015. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. 17**,** 405-24.

RIKKEN, J. F. W., KOWALIK, C. R., EMANUEL, M. H., BONGERS, M. Y., SPINDER, T., DE KRUIF, J. H., BLOEMENKAMP, K. W. M., JANSEN, F. W., VEERSEMA, S., MULDERS, A., THURKOW, A. L., HALD, K., MOHAZZAB, A., KHALAF, Y., CLARK, T. J., FARRUGIA, M., VAN VLIET, H. A., STEPHENSON, M. S., VAN DER VEEN, F., VAN WELY, M., MOL, B. W. J. & GODDIJN, M. 2018. The randomised uterine septum transsection trial (TRUST): design and protocol. *BMC Womens Health,* 18**,** 163.

ROBINSON, L., GALLOS, I. D., CONNER, S. J., RAJKHOWA, M., MILLER, D., LEWIS, S., KIRKMAN-BROWN, J. & COOMARASAMY, A. 2012. The effect of sperm DNA fragmentation on miscarriage rates: a systematic review and meta-analysis. *Hum Reprod,* 27**,** 2908-17.

ROBINSON, W. P., BERNASCONI, F., LAU, A. & MCFADDEN, D. E. 1999. Frequency of meiotic trisomy depends on involved chromosome and mode of ascertainment. *Am J Med Genet,* 84**,** 34-42.

ROCHE. 2016. *Harmony Prenatal Test* [Online]. Available: http://www.harmonytestusa.com/.

RYAN, A., HUNKAPILLER, N., BANJEVIC, M., VANKAYALAPATI, N., FONG, N., JINNETT, K. N., DEMKO, Z., ZIMMERMANN, B., SIGURJONSSON, S., GROSS, S. J. & HILL, M. 2016. Validation of an Enhanced Version of a Single-Nucleotide Polymorphism-Based Noninvasive Prenatal Test for Detection of Fetal Aneuploidies. *Fetal Diagn Ther,* 40**,** 219-223.

SAKKAS, D. & ALVAREZ, J. G. 2010. Sperm DNA fragmentation: mechanisms of origin, impact on reproductive outcome, and analysis. *Fertil Steril,* 93**,** 1027-36.

SALOMON, L. J. & SOTIRIADIS, A. 2019. Risk of miscarriage following amniocentesis or chorionic villus sampling: systematic review of literature and updated meta-analysis. 54**,** 442-451.

SCHISTERMAN, E. F., SILVER, R. M., LESHER, L. L., FARAGGI, D., WACTAWSKI-WENDE, J., TOWNSEND, J. M., LYNCH, A. M., PERKINS, N. J., MUMFORD, S. L. & GALAI, N. 2014. Preconception low-dose aspirin and pregnancy outcomes: results from the EAGeR randomised trial. *Lancet,* 384**,** 29-36.

SEHNERT, A. J., RHEES, B., COMSTOCK, D., DE FEO, E., HEILEK, G., BURKE, J. & RAVA, R. P. 2011. Optimal detection of fetal chromosomal abnormalities by massively parallel DNA sequencing of cell-free fetal DNA from maternal blood. *Clin Chem,* 57**,** 1042-9.

SEQUENOM. 2017. *Materni21plus* [Online]. Available: https://www.sequenom.com/tests/reproductive-health/maternit21-plus [Accessed 18/05/2017.

SEYEDHASSANI, S. M., HOUSHMAND, M., KALANTAR, S. M., AFLATOONIAN, A., MODABBER, G., HADIPOUR, F. & FALLAHZADEH, M. H. 2010a. The point mutations of mitochondrial tRNA threonine and proline in idiopathic repeated pregnancy loss. *Iranian Journal of Reproductive Medicine,* 8**,** 45-50.

SEYEDHASSANI, S. M., HOUSHMAND, M., KALANTAR, S. M., MODABBER, G. & AFLATOONIAN, A. 2010b. No mitochondrial DNA deletions but more D-loop point mutations in repeated pregnancy loss. *J Assist Reprod Genet,* 27**,** 641-8.

SHAMSELDIN, H. E., KURDI, W., ALMUSAFRI, F., ALNEMER, M., ALKAFF, A., BABAY, Z., ALHASHEM, A., TULBAH, M., ALSAHAN, N., KHAN, R., SALLOUT, B., AL MARDAWI, E., SEIDAHMED, M. Z., MERIKI, N., ALSABER, Y., QARI, A., KHALIFA, O., EYAID, W., RAHBEENI, Z., KURDI, A., HASHEM, M., ALSHIDI, T., AL-OBEID, E., ABDULWAHAB, F., IBRAHIM, N., EWIDA, N., EL-AKOURI, K., AL MULLA, M., BEN-OMRAN, T., PERGANDE, M., CIRAK, S., AL TALA, S., SHAHEEN, R., FAQEIH, E. & ALKURAYA, F. S. 2018. Molecular autopsy in maternal-fetal medicine. *Genet Med,* 20**,** 420-427.

SHAMSELDIN, H. E., SWAID, A. & ALKURAYA, F. S. 2013. Lifting the lid on unborn lethal Mendelian phenotypes through exome sequencing. *Genet Med,* 15**,** 307-9.

SHAMSELDIN, H. E., TULBAH, M., KURDI, W., NEMER, M., ALSAHAN, N., AL MARDAWI, E., KHALIFA, O., HASHEM, A., KURDI, A., BABAY, Z., BUBSHAIT, D. K., IBRAHIM, N., ABDULWAHAB, F., RAHBEENI, Z., HASHEM, M. & ALKURAYA, F. S. 2015. Identification of embryonic lethal genes in humans by autozygosity mapping and exome sequencing in consanguineous families. *Genome Biol,* 16**,** 116.

SKRZYPEK, H. & HUI, L. 2017. Noninvasive prenatal testing for fetal aneuploidy and single gene disorders. *Best Pract Res Clin Obstet Gynaecol,* 42**,** 26-38.

SOLER, A., MORALES, C., MADEMONT-SOLER, I., MARGARIT, E., BORRELL, A., BOROBIO, V., MUNOZ, M. & SANCHEZ, A. 2017. Overview of Chromosome Abnormalities in First Trimester Miscarriages: A Series of 1,011 Consecutive Chorionic Villi Sample Karyotypes. *Cytogenet Genome Res,* 152**,** 81-89.

SALOMON LJ, Sotiriadis A, Wulff CB, Odibo A, Akolekar R. Risk of miscarriage following amniocentesis or chorionic villus sampling: systematic review of literature and updated meta-analysis. Ultrasound Obstet Gynecol. 2019 Oct;54(4):442-451.

SONG, Y., ZHOU, X., HUANG, S., LI, X., QI, Q., JIANG, Y., LIU, Y., MA, C., LI, Z., XU, M., CRAM, D. S. & LIU, J. 2016. Quantitation of fetal DNA fraction in maternal plasma using circulating single molecule amplification and re-sequencing technology (cSMART). *Clin Chim Acta,* 456**,** 151-156.

SPARKS, A. B., WANG, E. T., STRUBLE, C. A., BARRETT, W., STOKOWSKI, R., MCBRIDE, C., ZAHN, J., LEE, K., SHEN, N., DOSHI, J., SUN, M., GARRISON, J., SANDLER, J., HOLLEMON, D., PATTEE, P.,

TOMITA-MITCHELL, A., MITCHELL, M., STUELPNAGEL, J., SONG, K. & OLIPHANT, A. 2012. Selective analysis of cell-free DNA in maternal blood for evaluation of fetal trisomy. *Prenat Diagn,* 32**,** 3-9.

SPENCER, K., LIAO, A. W., SKENTOU, H., CICERO, S. & NICOLAIDES, K. H. 2000. Screening for triploidy by fetal nuchal translucency and maternal serum free beta-hCG and PAPP-A at 10-14 weeks of gestation. *Prenat Diagn,* 20**,** 495-9.

STEPHENSON, M. D., AWARTANI, K. A. & ROBINSON, W. P. 2002. Cytogenetic analysis of miscarriages from couples with recurrent miscarriage: a case-control study. *Hum Reprod,* 17**,** 446-51.

STOUFFS, K., VANDERMAELEN, D., TOURNAYE, H., LIEBAERS, I. & LISSENS, W. 2011. Mutation analysis of three genes in patients with maturation arrest of spermatogenesis and couples with recurrent miscarriages. *Reproductive Biomedicine Online,* 22**,** 65-71.

STRAVER, R., SISTERMANS, E. A., HOLSTEGE, H., VISSER, A., OUDEJANS, C. B. & REINDERS, M. J. 2014a. WISECONDOR: detection of fetal aberrations from shallow sequencing maternal plasma based on a within-sample comparison scheme. *Nucleic Acids Res,* 42**,** e31.

STRAVER, R., SISTERMANS, E. A. & REINDERS, M. J. 2014b. Introducing WISECONDOR for noninvasive prenatal diagnostics. *Expert Rev Mol Diagn,* 14**,** 513-5.

SUGIURA-OGASAWARA, M., OZAKI, Y., KATANO, K., SUZUMORI, N., KITAORI, T. & MIZUTANI, E. 2012. Abnormal embryonic karyotype is the most frequent cause of recurrent miscarriage. *Hum Reprod,* 27**,** 2297-303.

SUI, Y., CHEN, Q. & SUN, X. 2015. Association of skewed X chromosome inactivation and idiopathic recurrent spontaneous abortion: a systematic review and meta-analysis. *Reprod Biomed Online,* 31**,** 140-8.

SUN, K., JIANG, P., CHAN, K. C., WONG, J., CHENG, Y. K., LIANG, R. H., CHAN, W. K., MA, E. S., CHAN, S. L., CHENG, S. H., CHAN, R. W., TONG, Y. K., NG, S. S., WONG, R. S., HUI, D. S., LEUNG, T. N., LEUNG, T. Y., LAI, P. B., CHIU, R. W. & LO, Y. M. 2015. Plasma DNA tissue mapping by genome-wide methylation sequencing for noninvasive prenatal, cancer, and transplantation assessments. *Proc Natl Acad Sci U S A,* 112**,** E5503-12.

SUZUKI, T., BEHNAM, M., RONASIAN, F., SALEHI, M., SHIINA, M., KOSHIMIZU, E., FUJITA, A., SEKIGUCHI, F., MIYATAKE, S., MIZUGUCHI, T., NAKASHIMA, M., OGATA, K., TAKEDA, S., MATSUMOTO, N. & MIYAKE, N. 2018. A homozygous NOP14 variant is likely to cause recurrent pregnancy loss. *J Hum Genet,* 63**,** 425-430.

SUZUMORI, N., EBARA, T., YAMADA, T., SAMURA, O., YOTSUMOTO, J., NISHIYAMA, M., MIURA, K., SAWAI, H., MUROTSUKI, J., KITAGAWA, M., KAMEI, Y., MASUZAKI, H., HIRAHARA, F., SALDIVAR, J. S., DHARAJIYA, N., SAGO, H. & SEKIZAWA, A. 2016. Fetal cell-free DNA fraction in maternal plasma is affected by fetal trisomy. *J Hum Genet,* 61**,** 647-52.

TALY, J. F., MAGIS, C., BUSSOTTI, G., CHANG, J. M., DI TOMMASO, P., ERB, I., ESPINOSA-CARRASCO, J., KEMENA, C. & NOTREDAME, C. 2011. Using the T-Coffee package to build multiple sequence alignments of protein, RNA, DNA sequences and 3D structures. *Nat Protoc,* 6**,** 1669-82.

TJOA, M. L., CINDROVA-DAVIES, T., SPASIC-BOSKOVIC, O., BIANCHI, D. W. & BURTON, G. J. 2006. Trophoblastic oxidative stress and the release of cell-free feto-placental DNA. *Am J Pathol,* 169**,** 400-4.

TOGNERI, F. S., KILBY, M. D., YOUNG, E., COURT, S., WILLIAMS, D., GRIFFITHS, M. J. & ALLEN, S. K. 2019. Implementation of cell-free DNA-based non-invasive prenatal testing in a National Health Service Regional Genetics Laboratory. *Genet Res (Camb),* 101**,** e11.

TONG, M., VIALL, C. A. & CHAMLEY, L. W. 2015. Antiphospholipid antibodies and the placenta: a systematic review of their in vitro effects and modulation by treatment. *Hum Reprod Update,* 21**,** 97-118.

TSURUSAKI, Y., YONEZAWA, R., FURUYA, M., NISHIMURA, G., POOH, R. K., NAKASHIMA, M., SAITSU, H., MIYAKE, N., SAITO, S. & MATSUMOTO, N. 2014. Whole exome sequencing revealed

biallelic IFT122 mutations in a family with CED1 and recurrent pregnancy loss. *Clin Genet,* 85**,** 592-4.

ULKER, V., GURKAN, H., TOZKIR, H., KARAMAN, V., OZGUR, H., NUMANOGLU, C., GEDIKBASI, A., AKBAYIR, O. & UYGUNER, Z. O. 2013. Novel NLRP7 mutations in familial recurrent hydatidiform mole: are NLRP7 mutations a risk for recurrent reproductive wastage? *Eur J Obstet Gynecol Reprod Biol,* 170**,** 188-92.

VAN BEEK, D. M., STRAVER, R., WEISS, M. M., BOON, E. M. J., HUIJSDENS-VAN AMSTERDAM, K., OUDEJANS, C. B. M., REINDERS, M. J. T. & SISTERMANS, E. A. 2017. Comparing methods for fetal fraction determination and quality control of NIPT samples. 37**,** 769-773.

VAN DEN BERG, M. M., VAN MAARLE, M. C., VAN WELY, M. & GODDIJN, M. 2012. Genetics of early miscarriage. *Biochim Biophys Acta,* 1822**,** 1951-9.

VERHAGEN, J. M. A., VAN DEN BORN, M., KURUL, S., ASIMAKI, A., VAN DE LAAR, I., FROHN-MULDER, I. M. E., KAMMERAAD, J. A. E., YAP, S. C., BARTELINGS, M. M., VAN SLEGTENHORST, M. A., VON DER THÜSEN, J. H. & WESSELS, M. W. 2018. Homozygous Truncating Variant in PKP2 Causes Hypoplastic Left Heart Syndrome. *Circ Genom Precis Med,* 11**,** e002397.

VIAGGI, C. D., CAVANI, S., MALACARNE, M., FLORIDDIA, F., ZEREGA, G., BALDO, C., MOGNI, M., CASTAGNETTA, M., PIOMBO, G., COVIELLO, D. A., CAMANDONA, F., LIJOI, D., INSEGNO, W., TRAVERSA, M. & PIERLUIGI, M. 2013. First-trimester euploid miscarriages analysed by array-CGH. *Journal of Applied Genetics,* 54**,** 353-9.

VODOPIUTZ, J., SEIDL, R., PRAYER, D., KHAN, M. I., MAYR, J. A., STREUBEL, B., STEISS, J. O., HAHN, A., CSAICSICH, D., CASTRO, C., ASSOUM, M., MULLER, T., WIECZOREK, D., MANCINI, G. M., SADOWSKI, C. E., LEVY, N., MEGARBANE, A., GODBOLE, K., SCHANZE, D., HILDEBRANDT, F., DELAGUE, V., JANECKE, A. R. & ZENKER, M. 2015. WDR73 Mutations Cause Infantile Neurodegeneration and Variable Glomerular Kidney Disease. *Hum Mutat,* 36**,** 1021-8.

WANG, B. T., CHONG, T. P., BOYAR, F. Z., KOPITA, K. A., ROSS, L. P., EL-NAGGAR, M. M., SAHOO, T., WANG, J. C., HEMMAT, M., HADDADIN, M. H., OWEN, R. & ANGUIANO, A. L. 2014. Abnormalities in spontaneous abortions detected by G-banding and chromosomal microarray analysis (CMA) at a national reference laboratory. *Mol Cytogenet,* 7**,** 33.

WANG, E., BATEY, A., STRUBLE, C., MUSCI, T., SONG, K. & OLIPHANT, A. 2013. Gestational age and maternal weight effects on fetal cell-free DNA in maternal plasma. *Prenat Diagn,* 33**,** 662-6.

WANG, Y., CHENG, Q., MENG, L., LUO, C., HU, H., ZHANG, J., CHENG, J., XU, T., JIANG, T., LIANG, D., HU, P. & XU, Z. 2017. Clinical application of SNP array analysis in first-trimester pregnancy loss: a prospective study. 1**,** 849-858.

WARREN, J. E., TUROK, D. K., MAXWELL, T. M., BROTHMAN, A. R. & SILVER, R. M. 2009. Array comparative genomic hybridization for genetic evaluation of fetal loss between 10 and 20 weeks of gestation. *Obstet Gynecol,* 114**,** 1093-102.

WEISS, M. M. *Eur J Hum Genet*.

WILBE, M., EKVALL, S., EURENIUS, K., ERICSON, K., CASAR-BOROTA, O., KLAR, J., DAHL, N., AMEUR, A., ANNEREN, G. & BONDESON, M. L. 2015. MuSK: a new target for lethal fetal akinesia deformation sequence (FADS). *Journal of Medical Genetics,* 52**,** 195-202.

WILLEMS, M., GENEVIEVE, D., BORCK, G., BAUMANN, C., BAUJAT, G., BIETH, E., EDERY, P., FARRA, C., GERARD, M., HERON, D., LEHEUP, B., LE MERRER, M., LYONNET, S., MARTIN-COIGNARD, D., MATHIEU, M., THAUVIN-ROBINET, C., VERLOES, A., COLLEAUX, L., MUNNICH, A. & CORMIER-DAIRE, V. 2010. Molecular analysis of pericentrin gene (PCNT) in a series of 24 Seckel/microcephalic osteodysplastic primordial dwarfism type II (MOPD II) families. *J Med Genet,* 47**,** 797-802.

WILSON, R., GEYER, S. H., REISSIG, L., ROSE, J., SZUMSKA, D., HARDMAN, E., PRIN, F., MCGUIRE, C., RAMIREZ-SOLIS, R., WHITE, J., GALLI, A., TUDOR, C., TUCK, E., MAZZEO, C., SMITH, J. C., ROBERTSON, E., ADAMS, D. J., MOHUN, T. & WENINGER, W. J. 2016a. Highly variable

penetrance of abnormal phenotypes in embryonic lethal knockout mice. *Wellcome Open Res,* 1**,** 1.

WILSON, R., GEYER, S. H., REISSIG, L., ROSE, J., SZUMSKA, D., HARDMAN, E., PRIN, F., MCGUIRE, C., RAMIREZ-SOLIS, R., WHITE, J., GALLI, A., TUDOR, C., TUCK, E., MAZZEO, C. I., SMITH, J. C., ROBERTSON, E., ADAMS, D. J., MOHUN, T. & WENINGER, W. J. 2016b. Highly variable penetrance of abnormal phenotypes in embryonic lethal knockout mice. *Wellcome Open Res,* 1.

WINDHAM, G. C., VON BEHREN, J., FENSTER, L., SCHAEFER, C. & SWAN, S. H. 1997. Moderate maternal alcohol consumption and risk of spontaneous abortion. *Epidemiology,* 8**,** 509-14.

WRIGHT, C. F., FITZGERALD, T. W., JONES, W. D., CLAYTON, S., MCRAE, J. F., VAN KOGELENBERG, M., KING, D. A., AMBRIDGE, K., BARRETT, D. M., BAYZETINOVA, T., BEVAN, A. P., BRAGIN, E., CHATZIMICHALI, E. A., GRIBBLE, S., JONES, P., KRISHNAPPA, N., MASON, L. E., MILLER, R., MORLEY, K. I., PARTHIBAN, V., PRIGMORE, E., RAJAN, D., SIFRIM, A., SWAMINATHAN, G. J., TIVEY, A. R., MIDDLETON, A., PARKER, M., CARTER, N. P., BARRETT, J. C., HURLES, M. E., FITZPATRICK, D. R. & FIRTH, H. V. 2015. Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet,* 385**,** 1305-14.

XU, Y., LI, X., GE, H. J., XIAO, B., ZHANG, Y. Y., YING, X. M., PAN, X. Y., WANG, L., XIE, W. W., NI, L., CHEN, S. P., JIANG, W. T., LIU, P., YE, H., CAO, Y., ZHANG, J. M., LIU, Y., YANG, Z. J., CHEN, Y. W., CHEN, F., JIANG, H. & JI, X. 2015. Haplotype-based approach for noninvasive prenatal tests of Duchenne muscular dystrophy using cell-free fetal DNA in maternal plasma. 17**,** 889-96.

YAMAGUCHI, H. & CONDEELIS, J. 2007. Regulation of the actin cytoskeleton in cancer cell migration and invasion. *Biochim Biophys Acta,* 1773**,** 642-52.

YARON, Y., OCHSHORN, Y., TSABARI, S. & SHIRA, A. B. 2004. First-trimester nuchal translucency and maternal serum free beta-hCG and PAPP-A can detect triploidy and determine the parental origin. *Prenat Diagn,* 24**,** 445-50.

YARON, Y., PAUTA, M., BADENAS, C., SOLER, A., BOROBIO, V., ILLANES, C., PAZ, Y. M. F., MARTINEZ-PORTILLA, R. & BORRELL, A. 2020. Maternal plasma genome-wide cell-free DNA can detect fetal aneuploidy in early and recurrent pregnancy loss and can be used to direct further workup. *Hum Reprod,* 35**,** 1222-1229.

YOO, S. K., LIM, B. C., BYEUN, J., HWANG, H., KIM, K. J., HWANG, Y. S., LEE, J., PARK, J. S., LEE, Y. S., NAMKUNG, J., PARK, J., LEE, S., SHIN, J. Y., SEO, J. S., KIM, J. I. & CHAE, J. H. 2015. Noninvasive prenatal diagnosis of duchenne muscular dystrophy: comprehensive genetic diagnosis in carrier, proband, and fetus. *Clin Chem,* 61**,** 829-37.

YU, S. C., CHAN, K. C., ZHENG, Y. W., JIANG, P., LIAO, G. J., SUN, H., AKOLEKAR, R., LEUNG, T. Y., GO, A. T., VAN VUGT, J. M., MINEKAWA, R., OUDEJANS, C. B., NICOLAIDES, K. H., CHIU, R. W. & LO, Y. M. 2014. Size-based molecular diagnostics using plasma DNA for noninvasive prenatal testing. *Proc Natl Acad Sci U S A,* 111**,** 8583-8.

ZHANG, H., LIU, W., CHEN, M., LI, Z., SUN, X. & WANG, C. 2016. Implementation of a High-Resolution Single-Nucleotide Polymorphism Array in Analyzing the Products of Conception. *Genetic Testing & Molecular Biomarkers,* 20**,** 352-8.

ZHANG, Y., LI, G., FAN, Y., CUI, Y., HUANG, S., MA, J., YAN, J. & CHEN, Z. J. 2015. Novel missense mutation in WNT6 in 100 couples with unexplained recurrent miscarriage. *Hum Reprod,* 30**,** 994-9.

ZHANG, Y. X., ZHANG, Y. P., GU, Y., GUAN, F. J., LI, S. L., XIE, J. S., SHEN, Y., WU, B. L., JU, W., JENKINS, E. C., BROWN, W. T. & ZHONG, N. 2009. Genetic analysis of first-trimester miscarriages with a combination of cytogenetic karyotyping, microsatellite genotyping and arrayCGH. *Clin Genet,* 75**,** 133-40.

ZHENG, H. Y., TANG, Y., NIU, J., LI, P., YE, D. S., CHEN, X., SHI, X. Y., LI, L. & CHEN, S. L. 2013. Aberrant DNA methylation of imprinted loci in human spontaneous abortions after assisted reproduction techniques and natural conception. *Human Reproduction,* 28**,** 265-273.

ZHONG, X. Y., HOLZGREVE, W. & HAHN, S. 2001. Risk free simultaneous prenatal identification of fetal Rhesus D status and sex by multiplex real-time PCR using cell free fetal DNA in maternal plasma. *Swiss Med Wkly,* 131**,** 70-4.

ZHOU, Q., WU, S. Y., AMATO, K., DIADAMO, A. & LI, P. 2016. Spectrum of Cytogenomic Abnormalities Revealed by Array Comparative Genomic Hybridization on Products of Conception Culture Failure and Normal Karyotype Samples. *J Genet Genomics,* 43**,** 121-31.

ZHOU, Y., ZHU, Z., GAO, Y., YUAN, Y., GUO, Y., ZHOU, L., LIAO, K., WANG, J., DU, B., HOU, Y., CHEN, Z., CHEN, F., ZHANG, H., YU, C., ZHAO, L., LAU, T. K., JIANG, F. & WANG, W. 2015. Effects of Maternal and Fetal Characteristics on Cell-Free Fetal DNA Fraction in Maternal Plasma. *Reprod Sci,* 22**,** 1429-35.

ZIMMERMANN, B., EL-SHEIKHAH, A., NICOLAIDES, K., HOLZGREVE, W. & HAHN, S. 2005. Optimized real-time quantitative PCR measurement of male fetal DNA in maternal plasma. *Clin Chem,* 51**,** 1598-604.

ZIMMERMANN, B., HILL, M., GEMELOS, G., DEMKO, Z., BANJEVIC, M., BANER, J., RYAN, A., SIGURJONSSON, S., CHOPRA, N., DODD, M., LEVY, B. & RABINOWITZ, M. 2012. Noninvasive prenatal aneuploidy testing of chromosomes 13, 18, 21, X, and Y, using targeted sequencing of polymorphic loci. *Prenat Diagn,* 32**,** 1233-41.

ZUHLKE, K. A., JOHNSON, A. M., TOMLINS, S. A., PALANISAMY, N., CARPTEN, J. D., LANGE, E. M., ISAACS, W. B. & COONEY, K. A. 2014. Identification of a novel germline SPOP mutation in a family with hereditary prostate cancer. *Prostate,* 74**,** 983-90.

Appendices and Supplementary Tables

Appendix 1- published article- Cell-Free DNA in the Investigation of Miscarriage

Journal of *Clinical Medicine*

MDPI

Article

Cell-Free DNA in the Investigation of Miscarriage

Emily Colley [1,2,3,*], Adam J. Devall [1,2], Helen Williams [1,4], Susan Hamilton [3], Paul Smith [1,2], Neil V. Morgan [5], Siobhan Quenby [6,7], Arri Coomarasamy [1,2] and Stephanie Allen [3]

1 Tommy's National Centre for Miscarriage Research, Birmingham Women's and Children's Hospital, Birmingham B15 2TG, UK; A.J.Devall@bham.ac.uk (A.J.D.); H.M.Williams.1@bham.ac.uk (H.W.); paul.smith@doctors.org.uk (P.S.); A.Coomarasamy@bham.ac.uk (A.C.)
2 Institute of Metabolism and Systems Research, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK
3 West Midlands Regional Genetics Laboratory, Birmingham Women's and Children's Hospital, Birmingham B15 2TG, UK; susan.hamilton15@nhs.net (S.H.); stephanie.allen13@nhs.net (S.A.)
4 Institute of Clinical Sciences, College of Medical and Dental Sciences, University of Birmingham, Edgbaston B15 2TT, UK
5 Institute of Cardiovascular Sciences, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK; N.V.Morgan@bham.ac.uk
6 Division of Biomedical Sciences, Warwick Medical School, University of Warwick, Coventry CV4 7HL, UK; S.Quenby@warwick.ac.uk
7 Tommy's National Centre for Miscarriage Research, University Hospitals Coventry & Warwickshire NHS Trust, Coventry CV2 2DX, UK
* Correspondence: emily.colley1@nhs.net

Received: 1 October 2020; Accepted: 22 October 2020; Published: 26 October 2020

check for updates

Abstract: Approximately one in four pregnancies result in pregnancy loss, and ~50% of these miscarriages are caused by chromosomal abnormalities. Genetic investigations are recommended after three consecutive miscarriages on products of conception (POC) tissue. Cell-free DNA (cfDNA) has been utilised for prenatal screening, but very little work has been carried out in nonviable pregnancies. We investigated the use of cfDNA from maternal blood to identify chromosomal abnormalities in miscarriage. One hundred and two blood samples from women experiencing a first trimester miscarriage were collected and stored. The mean gestational age was 7.1 weeks (range: 5–11 weeks). In this research, samples without a genetic test result from POC were not analysed. CfDNA was extracted and analysed using a modified commercial genome-wide non-invasive prenatal test. No results were provided to the patient. In 57 samples, cytogenetic results from POC analysis were available. Chromosomal abnormalities were identified in 47% (27/57) of POC analyses, and cfDNA analysis correctly identified 59% (16/27) of these. In total, 75% (43/57) of results were correctly identified. The average cfDNA fetal fraction was 6% (2–19%). In conclusion, cfDNA can be used to detect chromosomal abnormalities in miscarriages where the 'fetal fraction' is high enough; however, more studies are required to identify variables that can affect the overall results.

Keywords: miscarriage; cell-free DNA; cytogenetic analysis; chromosomal abnormalities

1. Introduction

Early pregnancy loss is the most common complication during pregnancy [1], and is defined as miscarriage. One in five pregnancies ends in spontaneous miscarriage [2], and 50% of these are due to chromosomal abnormalities [3]. It is important to identify whether a chromosomal abnormality was the underlying etiology of the pregnancy loss because this may have an indication for the prognosis of future pregnancies. If a sporadic chromosomal abnormality is the cause of the pregnancy loss, the prognosis for future pregnancies is better than if the chromosome complement is normal. In which

case, there may be another non-chromosomal, reason for the miscarriage. If there is an unbalanced chromosomal rearrangement in the pregnancy loss, it could mean that one of the parents carries a balanced chromosomal rearrangement. This would mean that future pregnancies would be susceptible to the same or other unbalanced rearrangement. In these cases, it is important to obtain blood samples for parental karyotyping for assessment of recurrence risk.

The Royal College of Obstetricians and Gynaecologists (RCOG) Green-top Guideline No. 17 [4] recommends cytogenetic analysis of pregnancy tissue after the third and subsequent miscarriages, or karyotyping of parental samples if pregnancy tissue is not available. Traditionally, cell culture and G banded chromosomal analysis were used to detect abnormalities in pregnancy tissue. However, there is often a high failure rate, due to the poor quality of tissue received, the difficulty with culturing cells from such tissues and a limited resolution in detecting micro-deletion and duplication syndromes. Therefore, molecular-based approaches, such as quantitative fluorescent PCR (QF-PCR) and microarray have been implemented across laboratories.

Currently, genetic testing for miscarriage is completed on pregnancy tissue, which comprises of placental and fetal components, referred to as products of conception (POC). This tissue needs to be fresh, uncontaminated, and unfixed so that the fetal tissues can be identified and have DNA extraction or cell culture performed. This comes with the risk of potential maternal cell contamination (MCC) which could lead to misdiagnosis of the sample. The POC samples contain maternal tissues intertwined with fetal tissues. Maternal cells can be carried over during the selection of fetal tissues resulting in maternal DNA during DNA extraction or an overgrowth of maternal cells during cell culture. Moreover, in many cases, POC are unavailable, or unreturned by the patient.

Cell-free DNA (cfDNA) was first identified by Dennis Lo [5] who demonstrated that small fragments of cfDNA from the plasma of pregnant women represent the entire fetal genome. Although cfDNA is already utilised for prenatal screening, very little work has been carried out in nonviable pregnancies to date. Only two studies by Clark-Ganheart el al. and Yaron et al. [6,7] have evaluated the use of cfDNA in a miscarriage setting.

A prospective cohort study Clark-Ganheart et al. [6] analysed 50 cfDNA samples of non-viable pregnancies. Gestational age determined by ultrasound scan ranged from 6.1 to 38.4 weeks. Among these, 38 of the 50 samples had a reportable result, including eight samples which demonstrated trisomies. The study by Yaron et al. [7] tested cfDNA to analyse pregnancy loss at less than 14 weeks. In total, 86 pregnancies had cfDNA results with comparable POC (from CVS sampling). The median fetal fraction was 5%. Out of the 86 samples, 55 (64%) had a chromosomal abnormality and 30 of those (55%) were detected using standard non-invasive prenatal testing (NIPT) log-likelihood ratio (LLR) cut-offs. To increase the sensitivity, a pregnancy-loss specific threshold was developed using a 50 sample 'training set'. This increased the detection rate to 82%.

CfDNA would be extremely useful to ascertain chromosomal causes of miscarriages at the point of miscarriage diagnosis by a simple blood test. This study investigates how cfDNA can be utilised to detect chromosomal abnormalities in miscarriage and to compare the results with those of POC testing.

2. Materials and Methods

2.1. Ethical Approval

The study was completed at Tommy's National Centre for Miscarriage Research, with IRAS project ID, 215646, that received Research Ethics Approval (REC reference: 16/WM/0423, 23/11/2016, West Midlands-South Birmingham Research Ethics Committee) and Health Research Authority (HRA) approval.

2.2. Patient Samples

Informed consent was obtained from patients experiencing early miscarriage and seen at Tommy's National Centre for Miscarriage Research hosted by Birmingham Women's and Children's Hospital

b

NHS Trust and University Hospital Coventry & Warwickshire NHS Trust between February 2017 and July 2019. The consent explicitly included consent to work with the patient's POC and genetic material. Samples were collected as donations to medical research and the tissue(s) were handled in accordance with the Human Tissue Act (HTA). The donors maintained their ability to withdraw consent for further use but did not retain any rights to the samples after acquisition.

Eligibility criteria included maternal age over 16 years and a gestational age of <12 weeks confirmed by ultrasound scan at the time of miscarriage diagnosis with pregnancy tissue remaining in situ. Samples were included in analysis in cases where there was a cytogenetic result from corresponding POC analysis, except in the case of seven known triploid cases, which were excluded.

Blood samples were taken for cfDNA analysis and to assess βhCG levels. Up to 10 mL of maternal blood was collected for cfDNA in cell-free DNA BCT (STRECK) tubes, and crown-rump length (CRL) measurements were taken by ultrasound where possible to assess the fetal gestation. Chromosomal abnormalities obtained from POC testing were communicated to the patient via standard patient care. CfDNA results were not shared with the patient.

2.3. Sample Processing

Plasma was isolated from whole blood using double centrifugation and transferred into a DNA LoBind tubes (Eppendorf) in 1 mL aliquots. These aliquots were stored at −80 °C until use.

2.4. Cytogenetic Analysis

Products of conception (POC) were collected as routine clinical samples and targeted QF-PCR and chromosomal microarray analysis (CMA) was completed on POC after the third and subsequent consecutive miscarriages(s) according to the RCOG Green-top Guideline No. 17 [4].

QF-PCR trisomy screen was first performed on DNA from POC to test for trisomies 13, 18 or 21, triploidy and sex chromosome aneuploidy. If the QF-PCR was abnormal, then it was reported. If it was normal, then CMA testing was carried out using OGT CytoSure 8 × 60 k Constitutional v3 design; exon/gene level resolution of ~500 DDD/ClinGen curated developmental genes and syndromic regions; tiered backbone resolution ~120–500 kb; analysis in build GRCh37 using CytoSure v4.9 and CBS algorithm. The microarray analysis detected copy number imbalances >1 Mb and in some cases had higher resolution.

2.5. Cell-free DNA Testing

Plasma (1 mL) from patients who had consented to testing by an external laboratory were submitted to the Illumina laboratory in Cambridge and processed in a 24-sample batch through a modified Illumina VeriSeq NIPT solution v2 workflow as previously described [8,9], but using the latest analysis platform [10] and with a few small modifications.

3. Results

In total, 102 cfDNA samples were collected once a miscarriage had been confirmed. All 102 samples were analysed by VeriSeq NIPT v2 Solution analysis on a NextSeq500. Eighty-five corresponding POC samples were received. In total, 64 pregnancies had a corresponding cytogenetic result from POC analysis, 21 POC samples were not suitable for analysis and 17 cfDNA samples did not receive corresponding POC samples. The 17 unreceived POC samples and the 21 POC samples not suitable for analysis were excluded from cfDNA analysis, and known triploid pregnancies were excluded (Figure 1).

Chromosomal abnormalities were identified by POC analysis in 27/57 (47%) cases. Patient baseline characteristics are summarised in Table 1. From the 57 samples with corresponding cytogenetic analysis, the average age was 34 years (20–43 years), with a clinical gestation of 7.1 weeks (5–11 weeks) and a fetal fraction of 6% (2–19%). In total, 70% (40/57) of samples including 16/27 (59%) of genetic abnormalities and 27/30 (90%) of genetically normal samples were identified correctly using VeriSeq.

C

This corresponds to a sensitivity of 59% (16/27), specificity of 90% (27/30) and accuracy of 75% (43/57), although the sample cohort was relatively small.

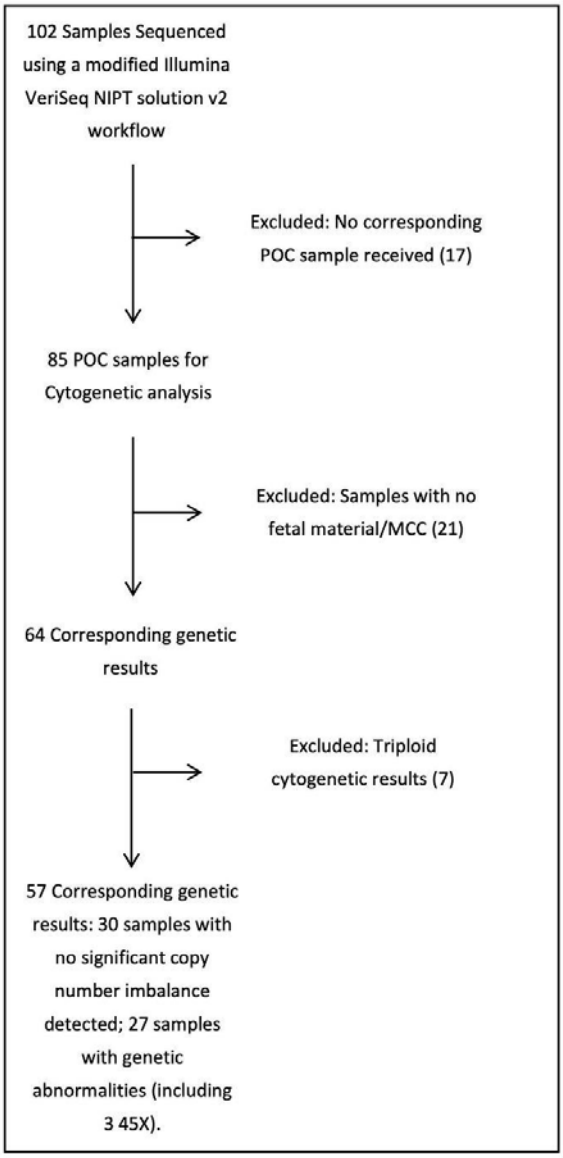


Figure 1. Flow chart describing inclusion/exclusion of cfDNA samples. MCC: maternal cell contamination; NIPT: on-invasive prenatal testing; POC: products of conception.

d

Table 1. Characteristics of cfDNA samples with corresponding products of conception (POC) results suitable for analysis (excluding triploid pregnancies).

	Total (n = 57)	Chromosomally Normal (n = 30)	Chromosomally Abnormal (n = 27)
Maternal age (years) (mean and range)	34 (20–43)	31 (20–41)	37 (24–43)
Previous losses (mean and range)	3.8 (0–14)	4.1 (0–14)	3.3 (0–6)
Gestational age (weeks) (mean and range)	7.1 (5–11)	7.4 (5–11)	6.4 (5–9.3)
βhCG (mIU/mL) (mean and range)	38,356 (69–263,766)	50,632 (69–263,766)	21,538 (491–100,638)
Fetal fraction (%) (mean and range)	6 (2–19)	7 (2–19)	5 (3–12)

CfDNA analysis correctly identified 43/57 (75%) of samples including 16 abnormal and 27 normal samples. Table 2 compares POC results and cfDNA test results of the cases where an abnormal result was detected in POC. The following anomalies were detected in POC: common trisomies (3), monosomy X (2), common trisomy combined with 45, X (1), monosomy 21 (1), full rare trisomies (14), mosaic rare trisomies (2) and copy number variations (4). Amongst the rare trisomies, trisomy 22 and trisomy 15 were the most frequent. Fetal fractions were from 3–12% (mean 5%). CfDNA results were fully concordant with POC results in 40/57 samples. CfDNA results generated normal results in 27/30 cases, and discrepant results in 3/30 cases of known normal cases. The two mosaic samples (sample IDs 51 and 586) were not correctly identified with cfDNA testing; however, other imbalances were detected in those samples. Sample ID 228 gave a monosomy 21 result on POC, but cfDNA testing identified CNVs in several other chromosomes.

Four of the samples were from miscarriages where subchromosomal deletions and duplications were identified by POC analysis (sample IDs 133, 202, 303 and 319). A 56 Mb duplication at 7q22.1q36.3 and a 21 Mb terminal duplication at 19q13.12q13.43 were detected by cfDNA analysis. A 70 Mb deletion at 13q13.3q34, a 6 Mb terminal deletion at 7q36.2q36.3, a 9 Mb terminal duplication at 4q34.3q35.2 and a 30 Mb terminal deletion at 5q33.1q35.3 were not detected by cfDNA analysis.

The results were grouped into three categories using different gestations, βhCG values and fetal fraction cut offs, to see if this could improve the result calling between cfDNA and POC cytogenetic analysis (Table 3). The gestation was split into four groups, <7 weeks, 7–8 weeks, ≥8 and unknown gestation. As the gestation increased in these groups, the correctly identified chromosomal abnormalities from cfDNA testing increased. The βhCG value was split into three groups of <8000, 8000–35,000 and >35,000 mIU/mL. As the βhCG value increased in these groups, the correctly identified chromosomal abnormalities from cfDNA testing also increased. The fetal fraction groups were split into three groups: <5, 5–8 and ≥9. Again, as fetal fraction groups increased so did the percentage of correctly identified chromosomal abnormalities from cfDNA testing.

e

Table 2. Analysis of cfDNA using a modified Illumina VeriSeq non-invasive prenatal testing (NIPT) solution v2 workflow compared to the genetic outcomes of microarray analysis of positive POC results.

Sample ID	Gestation (Weeks)	CRL (mm)	BhCG (mIU/mL)	Maternal Age (Years)	No. of Previous Losses	POC Results	POC Sex	CfDNA Results	CfDNA Sex	Fetal Fraction
4	7 + 2	2.8	19,247	40	2	Trisomy 22	Female	NO ANOMALY DETECTED	Female	4%
51	5 + 4	2	12,725	34	6	Mosaic trisomy 4	Male	DETECTED: del (10) (p15.3q21.1)	Male	4%
99	6 + 0	not recorded	5111	27	4	Trisomy 5	Male	NO ANOMALY DETECTED	Male	6%
133	not recorded	not recorded	57,348	34	4	Terminal deletion at 7q36.2q36.3 (6 Mb) and terminal duplication at 19q13.12q13.43 (21 Mb)	Female	DETECTED: dup (19) (q13.12q13.43)	Female	7%
163	8 + 4	20.1	13,466	24	3	Turners, 45 X	Female	DETECTED: XO	Female	7%
164	6 + 0	9	6774	42	2	Trisomy 15	Female	DETECTED: +15	Female	5%
175	8 + 0	No FP seen	5323	34	6	Turners, 45 X	Female	DETECTED: XO	Female	5%
176	7 + 0	14	491	43	2	Trisomy 15	Female	DETECTED: +15	Female	5%
202	6 + 0	5	13,819	29	4	Terminal duplication at 4q34.3q35.2 (9 Mb) and terminal deletion at 5q33.1q35.3 (30 Mb)	Female	NO ANOMALY DETECTED	Female	4%
228	7 + 0	12	6220	41	5	Monosomy 21	Male	DETECTED: dup (15) (q21.3q23); dup (20) (q11.21q13.12)	Male	3%
245	not recorded	6.5	14,762	33	3	Trisomy 22	Male	NO ANOMALY DETECTED	Male	4%
260	7 + 0	not recorded	5194	43	2	Trisomy 22	Male	NO ANOMALY DETECTED	Male	4%

Table 2. *Cont.*

Sample ID	Gestation (Weeks)	CRL (mm)	BhCG (mIU/mL)	Maternal Age (Years)	No. of Previous Losses	POC Results	POC Sex	CfDNA Results	CfDNA Sex	Fetal Fraction
264	6 + 0	4	14,002	33	Not recorded	Trisomy 13.	Female	DETECTED: +13; +16	Female	4%
279	6 + 0	6	8429	40	4	Trisomy 7	Female	DETECTED: +7	Female	12%
287	not recorded	4	not recorded	42	0	Trisomy 12	Female	DETECTED: +12	Female	4%
290	5 + 0	5.2	44,313	39	0	Trisomy 16	Female	DETECTED: +16	Female	7%
303	6 + 0	1.9	34,087	35	3	Deletion at 13q13.3q34 (70 Mb)	Female	NO ANOMALY DETECTED	Female	5%
319	7 + 0	9.3	13,642	40	6	Duplication 7q22.1q36.3 (56 Mb)	Female	DETECTED: dup (7) (q22.1q31.1)	Female	7%
328	not recorded	4.4	2983	42	3	Trisomy 11	Female	NO ANOMALY DETECTED	Female	4%
400	5 + 6	4.6	29,052	42	5	Trisomy 22	Male	DETECTED: +22	Male	3%
462	6 + 1	3	22,429	40	2	Trisomy 15	Male	DETECTED: +15	Male	5%
519	9 + 3	26.31	100,638	28	6	Trisomy 21	Male	DETECTED: +21	Male	10%
529	7 + 0	not recorded	21,171	39	5	Trisomy 21 and monosomy X.	Female	DETECTED: +21	Female	5%
541	7 + 3	12.02	42,333	33	2	Trisomy 15	Female	DETECTED: +15	Female	8%
586	6 + 0	5.28	22,435	41	2	Mosaic trisomy 17	Female	DETECTED: del (6) (p25.1p22.3); +18	Female	4%
816	5 + 0	no FP seen	5852	40	3	Trisomy 18.	Female	NO ANOMALY DETECTED	Female	4%
965	7 + 6	14.7	73,962	42	3	Trisomy 15	Male	DETECTED: +15	Male	INVALIDATED

CfDNA results labelled in green are concordant with POC results and results labelled in red are discordant with POC results.

Table 3. CfDNA vs karyotype of POC.

		CfDNA Results	
		Correctly Identified (%)	Not Identified (%)
	Total	43 (75.4)	14 (324.6)
Gestation (weeks)	<7	14 (66.7)	7 (33.3)
	7–8	10 (76.9)	3 (23.1)
	≥8	11 (100.0)	0 (0.0)
	Unknown	8 (53.3)	7 (46.7)
βhCG (mIU/mL)	<8000	9 (60.0)	6 (40.0)
	8000–35,000	14 (66.7)	7 (33.3)
	>35,000	19 (95.0)	1 (5.0)
Fetal Fraction (%)	<5	13 (59.1)	9 (40.9)
	5–8	19 (79.2)	5 (20.8)
	≥9	10 (100.0)	0 (0.0)

4. Discussions

Our cfDNA study cohort was recruited through Tommy's National Centre for Miscarriage Research, at Birmingham Women's Hospital and University Hospital Coventry and Warwickshire. In total, 102 samples were evaluated using modified VeriSeq NIPT V2 (Illumina), and 57 samples were analysed with corresponding POC cytogenetic analysis.

The cfDNA analysis was separated into three categories for analysis (Table 3). Whilst some chromosomal abnormalities were identified at lower fetal fraction, at <5% fetal fraction, only 60% of samples were correctly identified, and of those, most were from euploid pregnancies. In contrast, where the fetal fraction was ≥9%, 100% of cytogenetic results were correctly identified. In our study, we note that the majority of abnormalities can be detected above 5% fetal fraction. However, it is difficult to define an exact cut off due to the low sample numbers and biological variation.

The discrepancies we observed between the POC genetic results and the cfDNA testing could be caused by confined placental mosaicism. CfDNA analysis tests DNA derived from the placenta/cytotrophoblasts only, whereas the POC testing may consist of fetal tissue and whole placental tissue. This could result in a discrepancy between the results. In two cases, mosaic genetic abnormalities were identified in the POC analysis which cfDNA testing did not identify. These results could be due to confined placental mosaicism for the trisomic cells or due to the current limitation of the method. Mosaicism is difficult to diagnose with any methodology, and it is possible that cfDNA analysis could become a helpful adjunct to current POC testing in detecting biologically relevant abnormal cell lines.

Tommy's National Centre for Miscarriage Research is specialised in the care of families undergoing recurrent miscarriage. These families are very aware of when they first become pregnant and benefit from careful monitoring during their first trimester. Consequently, the miscarriages in our study cohort were diagnosed earlier than in other studies. Clark-Ganheart et al. [6] recorded gestational ages of 16.9 (6.1–37.2) weeks, and Yaron et al. [7] recorded gestational ages 9.6 (5.1–13.6) weeks (Figure 2).

This study and others have demonstrated that in the majority of cases of pregnancy loss where the pregnancy tissue is still in situ, it is possible to detect chromosomal abnormalities using cfDNA. This study correctly identified 59% of chromosomal abnormalities with a 75% concordance to POC results. In comparison, Clark-Ganheart et al. [6] had 87.5% concordant results where there was an available cytogenetic result, and Yaron et al. [7] had 82% concordant results using pregnancy loss-specific LLR thresholds. Using 50 cases as a training set, Yaron et al. [7] established a pregnancy loss-specific LLR threshold. Overall detection was 82% on 86 non-mosaic cases. This was achieved after identifying a pregnancy-loss-specific LLR based on a training set. This indicates that the LLR needed for this cohort may need to be different from singleton pregnancies. In comparison, our study used the standard NIPT LLR cut-offs to analyse cfDNA, and it is feasible that having a pregnancy-loss-specific LLR would improve the detection rate. The next step in this study may be to conduct a trial using

h

an algorithm similar to the one proposed by Yaron et al. [7], using a pipeline with modified LLRs to optimise the detection rate of all autosomal trisomies for this cohort of patients.

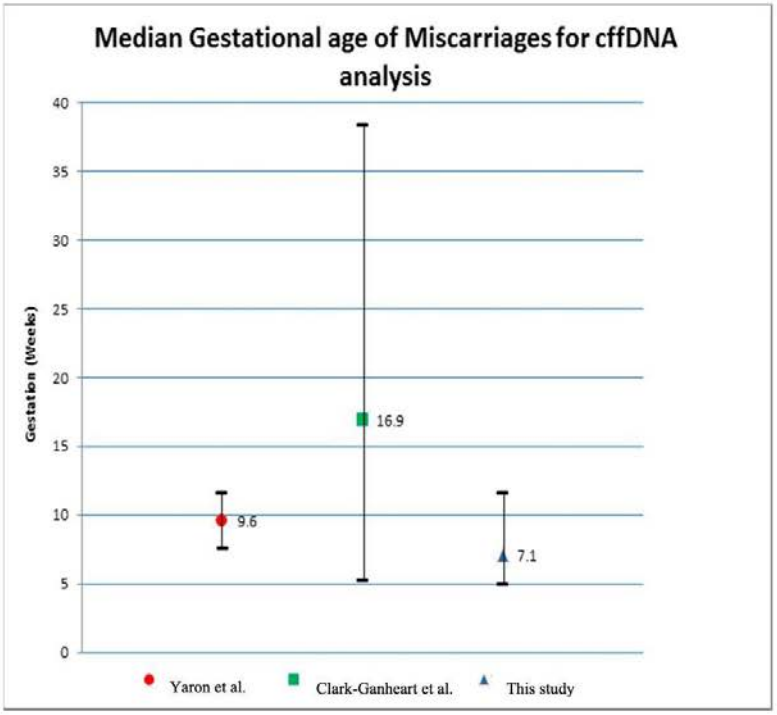


Figure 2. Gestations of pregnancy loss by ultrasound scan [6,7].

This study and others [6,7] have shown that cfDNA can be utilised to assess the genetic contribution to miscarriage. However, there are still some genetic abnormalities that can be missed, dependent on the assay used (e.g., triploid and mosaic samples, and autosomal trisomies at low fetal fraction/low gestation). Triploid cases were excluded in both this study and Yaron et al. [7] as they were not detectable by the modified Illumina VeriSeq NIPT solution v2 workflow. However, a single nucleotide polymorphism-based platform for analysis of cfDNA should be able to identify triploid cases.

CfDNA cannot completely replace current cytogenetic testing. A recurrent pregnancy loss algorithm was proposed by Yaron et al. [7] which would utilise cfDNA testing in recurrent pregnancy loss.

When a third or subsequent pregnancy loss has been diagnosed, current guidelines by the Royal College of Obstetricians and Gynaecologists (RCOG) Green-top Guidelines No. 17 [4] are to test the pregnancy tissue for fetal aneuploidies. Alongside this routine testing, a maternal blood sample could be collected to complete cfDNA testing. If an aneuploidy is detected in cfDNA testing and explains the reason for the miscarriage, no further work is required as numerical errors usually occur sporadically and the likelihood of a successful subsequent pregnancy is not negatively affected. As cfDNA only detects an unbalanced chromosomal abnormality, if no chromosomal abnormality is identified using cfDNA testing, then cytogenetic analysis on POC should be recommended to see if there is a chromosomal abnormality that is not detectable by cfDNA testing (e.g., CNVs, triploid or mosaic samples). This would reduce the number of POC tests required and could achieve a result

for more patients where there is no POC available. It is important to note that some chromosomal abnormalities would still be missed if POC is not available. In cases where an unbalanced rearrangement is identified that could be due to an inherited or de novo Robertsonian or reciprocal translocation, parental karyotyping should be recommended to assess whether one (or both) of the parents is a carrier of this translocation.

5. Conclusions

Knowing the genetic result of a pregnancy loss can be applied during counselling patients for the prognosis of future pregnancies. It may also be helpful to provide psychological support and relief from the guilt that can be associated with pregnancy loss.

Using cfDNA to identify whether a miscarriage was caused by chromosomal abnormalities would have a huge clinical impact upon patients for whom conventional cytogenetic testing may not be available, either due to the unavailability of pregnancy tissue for testing or patient preferences. However, cfDNA testing is only feasible where the pregnancy remains in situ at the time of miscarriage diagnosis.

We have demonstrated that in some cases, cfDNA can be used to detect a genetic aberration in miscarriages providing the maternal plasma sample is collected when the pregnancy tissue is still in situ and in cases where there is enough fetal fraction. Further work is required to improve this testing and to identify variables that can affect the overall results so that it may be applied clinically.

Author Contributions: E.C.—Data analysis and writing; A.J.D.—Study design and ethical approval; H.W.—Study design, ethical approval, grant administration and editing; S.H.—Data analysis and critical appraisal; P.S.—Study design and critical appraisal; N.V.M.—Supervision and critical appraisal; S.Q.—Critical appraisal; A.C.—Study design, supervision and critical appraisal; S.A.—Study design, supervision, critical appraisal, data analysis, writing and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Tommy's Charity, who funds Tommy's National Centre for Miscarriage Research.

Acknowledgments: The authors thank Lieve Page-Christians, Philippa Burns, Patrizia Di Pietro, Cosmin Deciu, Sarah Kinnings, Eoin Brown, Victoria Corey, Kathryn Robinson, Sven Schaffer, Sucheta Bhatt, Barbara Baggiani, Matt Smith, Marion Burdin and Jasmin Dehnhardt from Illumina. The authors thank the research teams at Tommy's National Centre for Miscarriage Research at Birmingham Women's and Children's Hospital and University Hospital Coventry & Warwickshire for patient recruitment and sample collection.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Roberts, C.; Lowe, C. Where have all the conceptions gone? *Lancet* **1975**, *305*, 498–499. [CrossRef]
2. Macklon, N.S.; Geraedts, J.P.M.; Fauser, B.C.J.M. Conception to ongoing pregnancy: The 'black box' of early pregnancy loss. *Hum. Reprod. Updat.* **2002**, *8*, 333–343. [CrossRef] [PubMed]
3. Choi, T.Y.; Lee, H.M.; Park, W.K.; Jeong, S.Y.; Moon, H.S. Spontaneous abortion and recurrent miscarriage: A comparison of cytogenetic diagnosis in 250 cases. *Obstet. Gynecol. Sci.* **2014**, *57*, 518–525. [CrossRef] [PubMed]
4. Regan, L.; Backos, M.; Rai, R. *The Investigation and Treatment of Couples with Recurrent First-Trimester and Second-Trimester Miscarriage*; RCOG Green-top Guideline No.17; Royal College of Obstetricians and Gynaecologists: London, UK, 2011.
5. Lo, Y.M.D.; Corbetta, N.; Chamberlain, P.F.; Rai, V.; Sargent, I.L.; Redman, C.W.; Wainscoat, J.S. Presence of fetal DNA in maternal plasma and serum. *Lancet* **1997**, *350*, 485–487. [CrossRef]
6. Clark-Ganheart, C.A.; Fries, M.H.; Leifheit, K.M.; Jensen, T.J.; Moreno-Ruiz, N.L.; Ye, P.P.; Jennings, J.M.; Driggers, R.W. Use of Cell-Free DNA in the Investigation of Intrauterine Fetal Demise and Miscarriage. *Obstet. Gynecol.* **2015**, *125*, 1321–1329. [CrossRef]
7. Yaron, Y.; Pauta, M.; Badenas, C.; Soler, A.; Borobio, V.; Illanes, C.; Paz-Y-Miño, F.; Martinez-Portilla, R.; Borrell, A. Maternal plasma genome-wide cell-free DNA can detect fetal aneuploidy in early and recurrent pregnancy loss and can be used to direct further workup. *Hum. Reprod.* **2020**, *35*, 1222–1229. [CrossRef] [PubMed]

j

8. Pertile, M.; Flowers, N.; Vavrek, D. P1-86. Performance of a genome-wide PCR-free, paired-end sequencing-based noninvasive prenatal screening test, VeriSeq NIPT Solution v2, in the detection of fetal chromosomal anomalies. Poster Abstracts of the ISPD 23rd International Conference on Prenatal Diagnosis and Therapy, Singapore, 7–11 September 2019. *Prenat Diagn.* **2020**, *40* (Suppl. 1), 74–75.
9. Andrews, D.; Padhukasahasram, B.; Kim, S. P1-66. Validation of VeriSeq NIPT Solution v2, a paired-end PCR-free WGS-based methodology for noninvasive prenatal screening of fetal aneuploidies and CNVs on a genome-wide scale. Poster Abstracts of the ISPD 23rd International Conference on Prenatal Diagnosis and Therapy, Singapore, 7–11 September 2019. *Prenat Diagn.* **2020**, *40* (Suppl. 1), 62–63.
10. Leinfinger, P.; Lohmann, L.; Luscan, A.; Trost, D.; Bidat, L.; Debarge, V.; Castaigne, V.; Senat, M.-V.; Brechard, M.-P.; Guilbaud, L.; et al. Strategy for Use of Genome-Wide Non-Invasive Prenatal Testing for Rare Autosomal Aneuploidies and Unbalanced Structural Chromosomal Anomalies. *J. Clin. Med.* **2020**, *9*, 2466. [CrossRef] [PubMed]

k

Appendix 2- published article- Potential genetic causes of miscarriage in euploid pregnancies: a systematic review

Human Reproduction Update, Vol.25, No.4, pp. 452–472, 2019
Advance Access Publication on May 31, 2019 doi:10.1093/humupd/dmz015

human reproduction update

Potential genetic causes of miscarriage in euploid pregnancies: a systematic review

Emily Colley [1,2,*], Susan Hamilton [1,3], Paul Smith [1,2], Neil V. Morgan [4], Arri Coomarasamy [1,2], and Stephanie Allen [1,3]

[1]Tommy's National Centre for Miscarriage Research, Birmingham Women's and Children's Hospital, Birmingham B15 2TG, UK, [2]Institute of Metabolism and Systems Research, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK, [3]West Midlands Regional Genetics Laboratory, Birmingham Women's and Children's Hospital, Birmingham B15 2TG, UK and [4]Institute of Cardiovascular Sciences, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

*Correspondence address. Tommy's National Centre for Miscarriage Research, Birmingham Women's and Children's Hospital, Birmingham, B15 2TG, UK. E-mail: ECC669@bham.ac.uk orcid.org/0000-0001-6373-750X

Submitted on February 28, 2018; resubmitted on March 7, 2019;; editorial decision on March 17, 2019

TABLE OF CONTENTS

BACKGROUND: Approximately 50% of pregnancy losses are caused by chromosomal abnormalities, such as aneuploidy. The remainder has an apparent euploid karyotype, but it is plausible that there are cases of pregnancy loss with other genetic aberrations that are not currently routinely detected. Studies investigating the use of exome sequencing and chromosomal microarrays in structurally abnormal pregnancies and developmental disorders have demonstrated their clinical application and/or potential utility in these groups of patients. Similarly, there have been several studies that have sought to identify genes that are potentially causative of, or associated with, spontaneous pregnancy loss, but the evidence has not yet been synthesized.

OBJECTIVE AND RATIONALE: The objective was to identify studies that have recorded monogenic genetic contributions to pregnancy loss in euploid pregnancies, establish evidence for genetic causes of pregnancy loss, identify the limitations of current evidence, and make recommendations for future studies. This evidence is important in considering additional research into Mendelian causes of pregnancy loss and appropriate genetic investigations for couples experiencing recurrent pregnancy loss.

SEARCH METHODS: A systematic review was conducted in MEDLINE (1946 to May 2018) and Embase (1974 to May 2018). The search terms 'spontaneous abortion', 'miscarriage', 'pregnancy loss', or 'lethal' were used to identify pregnancy loss terms. These were combined with search terms to identify the genetic contribution including 'exome', 'human genome', 'sequencing analysis', 'sequencing', 'copy number variation', 'single-nucleotide polymorphism', 'microarray analysis', and 'comparative genomic hybridization'. Studies were limited to pregnancy loss up to 20 weeks in humans and excluded if the genetic content included genes that are not lethal *in utero*, PGD studies, infertility studies, expression studies, aneuploidy with no recurrence risk, methodologies where there is no clinical relevance, and complex genetic studies. The quality of the studies was assessed using a modified version of the Newcastle–Ottawa scale.

OUTCOMES: A total of 50 studies were identified and categorized into three themes: whole-exome sequencing studies; copy number variation studies; and other studies related to pregnancy loss including recurrent molar pregnancies, epigenetics, and mitochondrial DNA aberrations. Putatively causative variants were found in a range of genes, including *CHRNA1* (cholinergic receptor, nicotinic, alpha polypeptide 1), *DYNC2H1* (dynein, cytoplasmic 2, heavy chain 1), and *RYR1* (ryanodine receptor 1), which were identified in multiple studies. Copy number variants were also identified to have a causal or associated link with recurrent miscarriage.

WIDER IMPLICATIONS: Identification of genes that are causative of or predisposing to pregnancy loss will be of significant individual patient impact with respect to counselling and treatment. In addition, knowledge of specific genes that contribute to pregnancy loss could also be of importance in designing a diagnostic sequencing panel for patients with recurrent pregnancy loss and also in understanding the biological pathways that can cause pregnancy loss.

Key words: genetic causes / pregnancy loss / euploid miscarriage / exome sequencing / chromosomal array / single-nucleotide variation / copy number variant

Introduction

Pregnancy loss is the most common complication of pregnancy and multiple studies have sought to identify potentially causative genes involved in pregnancy loss. In this systematic review, we have synthesized the evidence to establish a genetic causality of pregnancy loss.

Miscarriage and recurrent pregnancy loss

Approximately 15% of clinically recognized pregnancies end in pregnancy loss, with the majority occurring during the first trimester. Of these, 50% are caused by chromosomal abnormalities such as aneuploidy (Hassold *et al.*, 1980) and can be detected by conventional cytogenetic analysis. It is suggested that 86% of these abnormalities are numerical, 6% are structural abnormalities, and 8% are due to other genetic mechanisms, such as chromosomal mosaicism and molar pregnancies (Goddijn and Leschot, 2000).

Recurrent miscarriage (RM) is defined by the Royal College of Obstetricians and Gynaecologists (RCOG) as at least three consecutive miscarriages before 24 weeks gestation (RCOG, 2011) and recurrent pregnancy loss (RPL) by the ESHRE November 2017 guidelines as the loss of two or more pregnancies (Bender Atik *et al.*, 2018). In addition to genetic aetiology, a spectrum of non-genetic causes of RPL have also been identified, including thrombophilic factors, endocrinological causes, immunological and immunogenetic causes, sperm DNA fragmentation, uterine malformations, and lifestyle factors such as smoking, reviewed by (Larsen *et al.*, 2013).

Cytogenetic and chromosomal microarray analysis

Traditionally, cytogenetic analysis of pregnancy tissue has been performed to identify genetic causes of RPL and to indicate the need for further analysis of parental samples where there is the possibility of a balanced chromosome rearrangement (e.g. translocation) in one of the parents. It is important to identify any numeric chromosome errors, such as trisomy, monosomy, or polyploidy, since these are causes of pregnancy loss that usually occur sporadically, and the likelihood of a successful pregnancy outcome is not negatively affected in subsequent pregnancies. Where there is a balanced translocation in one of the parents, genetic counselling is important as there is likely to be a recurrence risk in future pregnancies and pre-implantation genetic testing, chorionic villus sampling, or amniocentesis can be used to detect an abnormality in the conceptus. However, for couples with a translocation, medical management (e.g. natural conception and observation) has been reviewed to have a higher live birth rate than IVF/PGD (Franssen *et al.*, 2011; Hirshfeld-Cytron *et al.*, 2011).

The most recent ESHRE guidelines for genetic analysis of products of conception (POC) give a conditional recommendation for genetic analysis but recommend that testing is carried out by array comparative genomic hybridization (CGH) instead of traditional karyotyping (Bender Atik *et al.*, 2018). Conventional karyotype analysis identifies balanced and unbalanced chromosomal rearrangements and copy number variants (CNVs) to a ~5 Mb resolution. Chromosomal microarray analysis can now identify unbalanced CNVs below 1 Mb, with a resolution at the level of individual exons of genes in targeted regions of the genome (Miller *et al.*, 2010). Microarray analysis is also less labour intensive as it is based on DNA analysis rather than cultured cells and has a higher success rate in poor quality tissue samples; however, the quality of tissue will impact the success and failure rate of both conventional karyotyping and array CGH. Array CGH has become the gold standard for genetic CNV analysis. It should, however, be noted that array CGH may miss some balanced chromosomal rearrangements and may also fail to identify maternal cell contamination.

Other genetic causes

In the case of pregnancy loss, with an apparently euploid karyotype, there may be genetic aberrations causative of pregnancy loss that are not currently known or routinely assessed. These could include single-nucleotide variants (SNVs) that affect individual genes and are detectable by sequencing or small sub-microscopic aberrations that affect a cluster of genes and are detectable by microarray analysis. In the case of SNVs, this is particularly important as many may follow a recessive or X-linked pattern of inheritance and therefore have a high recurrence risk. CNVs detected in cases of pregnancy loss may unmask a recessive mutation in a relevant gene or involve dosage-sensitive genes, where loss or gain of copies affects the gene function. These regions may also represent benign CNVs seen frequently with no recorded effect on phenotype, although it remains possible that some may be involved in RPL. Evidence in humans and other species (Wilson *et al.*, 2016) suggests that many genes are important in early development and can lead to embryonic lethality when functionally 'knocked out', resulting in pregnancy loss. More widespread genetic analysis of embryonic pregnancy loss may provide an opportunity to identify genes that are essential in early human development or where a lack of function leads to pregnancy loss.

Molar pregnancies

A molar pregnancy or hydatidiform mole (HM) is an abnormal pregnancy, which has cystic degeneration of the chorionic villi, abnormal proliferation of the trophoblast, and abnormal development of the fetus. These can either be complete HM or partial HM, distinguishable by the extent of trophoblast proliferation and presence of embryonic tissue.

Complete HMs are usually diploid with all chromosomes of paternal origin. The majority arise from an anuclear ovum being fertilized by a haploid sperm and replicating its own chromosomes (uniparental paternal isodisomy) or rarely from an anuclear ovum fertilized by two sperm (uniparental paternal heterodisomy). Partial HMs are mostly triploid with 23 chromosomes of maternal origin and 46 of paternal origin.

While HMs are usually triploid and sporadic and therefore outside the scope of this review; a minority of molar pregnancies are diploid and biparental, usually being recurrent and familial. These may be caused by maternal autosomal recessive mutations in genes, such as *NLRP7* and *KHDC3L*, resulting in an abnormal epigenotype of imprinted loci. This results in abnormal gene expression, which causes abnormal placental trophoblast development and manifests as HM (Carey *et al.*, 2015).

Whole-exome sequencing

Advances in sequencing technology, including whole-exome sequencing (WES) and whole-genome sequencing (WGS), are increasingly providing the opportunity to detect genetic sequence variation and to characterize genetic mutations causing disease. WGS is the most extensive sequencing method and targets the entire genome, whereas WES targets the exome, which is the protein-coding region of the DNA. The exome makes up ~1% of the human genome, and it is estimated to contain 85% of the genetic mutations associated with disease (Choi *et al.*, 2009). Generally, WES is the preferred method of sequencing because it is cheaper than WGS and has a smaller, more manageable data set while still comprehensively covering the coding regions of DNA. WGS has the advantage of analying and giving a comprehensive view of the whole genome and has the potential to detect large structural variants, insertions/deletions, SNVs, and copy number changes. However, we still understand relatively little about the non-coding regions of the genome.

Studies investigating the use of WES in structurally abnormal pregnancies, late pregnancy losses, and developmental disorders (Carss *et al.*, 2014; Wright *et al.*, 2015; Shamseldin *et al.*, 2018) have demonstrated the clinical application in these patients. However, very few WES studies have reported analysis in pregnancy loss or lethal genes that could contribute to RPL. The few studies using WES to look for genetic aberrations in RPL have also tended to represent only small patient cohorts. The ability to recognize and detect genetic mutations may have implications for routine genetic testing and clinical practice, especially when a pathogenic aberration is identified that can be reliably detected in future pregnancies.

Aims

There are several studies that have sought to identify genes causative of or associated with pregnancy loss, but the evidence has not yet been synthesized. We propose to review these studies and establish evidence of genetic causality of RPL, including reviewing appropriate methodologies. We will evaluate studies investigating Mendelian inheritance patterns, including autosomal recessive and dominant X-linked inheritance, and also *de novo* genetic causes, but we have excluded studies investigating more complex genetic associations, which have recently been systematically reviewed (Pereza *et al.*, 2017).

Methods

Registration

This systematic review has been registered with PROSPERO (CRD42017073910).

n

Search

A systematic literature review to assess the studies investigating the genetic contribution to RPL was conducted in MEDLINE (1946 to May 2018) and Embase (1974 to May 2018) using Ovid (https://ovidsp.tx.ovid.com). The search terms used to identify pregnancy loss were 'spontaneous abortion', 'miscarriage', 'pregnancy loss', or 'lethal, and the search terms to identify the genetic contributions are 'exome', 'human genome', 'sequencing analysis', 'sequencing', 'copy number variation', 'single-nucleotide polymorphism', 'microarray analysis', and 'comparative genomic hybridization'. The search terms and corresponding Mesh terms are shown in Supplementary Table SI. Additional studies were also identified from references of selected studies.

Study selection

Studies were selected by two independent reviewers. Studies were first screened for eligibility using article titles and then by screening the study abstracts. Studies were included if they had pregnancy loss up to 20 weeks, but were not restricted if they also included some later losses, providing the genetic aberrations were defined. Studies were excluded if the genetic content included genes that were not lethal *in utero*, PGD studies, infertility studies, expression studies, aneuploidy with no recurrence risk, methodologies where there is no clinical relevance, and complex genetics. Both recurrent and sporadic pregnancy losses were included. The full inclusion and exclusion criteria are presented in Supplementary Table SII.

Data extraction process

Data on publication date, country, study objective, sample, phenotype and gestation, methods and analysis, study outcome, and quality scores were extracted. Data extraction was checked by a second reviewer. Each of the identified genes were found in Online Mendelian Inheritance in Man and the Mendelian Inheritance in Man (MIM) number, gene name, gene function, associated disease/phenotype, and cytogenetic location were ascertained.

Quality assessment

The quality of each study was assessed using a modified Newcastle–Ottawa scale (Supplementary Table SIII). Each study was scored out of 12 and was judged on the sample size, inclusion/exclusion criteria, the genetic analysis method, statistical analysis, case definition, controls, and comparability. The breakdown of each score is included in Supplementary Table SIV.

Results

A total of 50 studies were included in the review. The initial search of the Medline and Embase databases identified 3404 potentially relevant articles. After screening the titles and abstracts, 74 full texts were obtained for detailed review. A total of 30 full articles were excluded because they were either not related to pregnancy loss, were more than 20 weeks gestation, or contained no genetic content. Examination

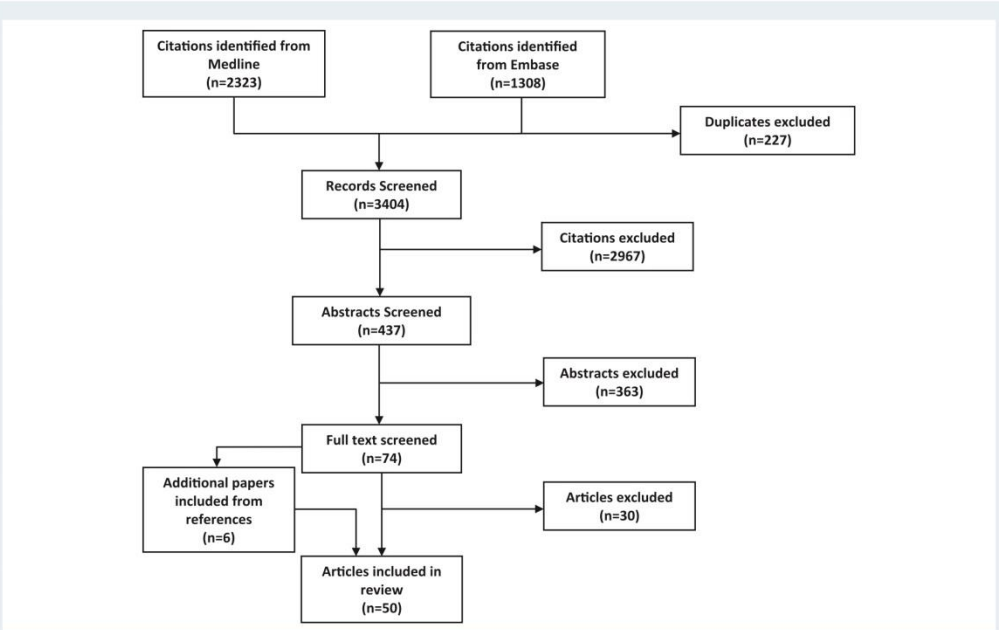


Figure 1 PRISM flow diagram for a systematic review of the potential genetic causes of miscarriage in euploid pregnancies.

of the bibliographies and journal indices generated six additional studies for the review. Fig. 1 illustrates the study selection. The papers identified were categorized into three themes: WES studies, CNV studies, and other studies related to pregnancy loss including recurrent molar pregnancies. The gene/locus names and MIM number of all genes discussed are displayed in Table I.

The 50 studies that met the inclusion and exclusion criteria were all published in English between 2009 and 2018. Out of the studies identified, 21 were from Europe, 14 were from North America, 13 were from Asia, and there was one study each from South America and Africa.

Whole-exome sequencing

Thirteen studies were identified (Table II) that used WES to identify SNVs in families with multiple pregnancy losses or a combination of pregnancy losses and terminations. Eight of these studies focused on a single couple only (Shamseldin *et al*., 2013; Filges *et al*., 2014; Tsurusaki *et al*., 2014; Dohrn *et al*., 2015; Rae *et al*., 2015; Wilbe *et al*., 2015; Bondeson *et al*., 2017; Cristofoli *et al*., 2017). Six studies used WES analysis of trios (Filges *et al*., 2014; Dohrn *et al*., 2015; Wilbe *et al*., 2015; Qiao *et al*., 2016; Bondeson *et al*., 2017; Cristofoli *et al*., 2017).

Studies using WES identified variants in genes from both fetal and parental samples, thus allowing for the inheritance to be identified. One study identified compound heterozygous mutations in *KIF14* in a family with unexplained euploid miscarriages (Filges *et al*., 2014). The other studies included pregnancies terminated for a fetal abnormality including a homozygous missense mutation in *ECEL1* from a consanguineous couple with pregnancies terminated due to arthrogryposis multiplex congenita (Dohrn *et al*., 2015), a novel homozygous mutation in the *MuSK* gene in a non-consanguineous couple with a history of fetal akinesia deformation sequence (FADS) (Wilbe *et al*., 2015), compound heterozygous mutations in *STIL* from a non-consanguineous couple with fetal microcephaly (Cristofoli *et al*., 2017), a homozygous nonsense mutation in *CEP55* in a non-consanguineous family with two fetuses with Meckel-like syndrome (Bondeson *et al*., 2017), and compound heterozygous mutations in *IFT122* in a couple experiencing both RPL and later losses with scan abnormalities (Tsurusaki *et al*., 2014).

Two studies (Shamseldin *et al*., 2013; Rae *et al*., 2015) identified pathogenic variants by WES of fetuses affected with hydrops fetalis. The first identified pathogenic variant in the gene *FOXP3* was from a non-consanguineous couple whom had multiple male pregnancy terminations. *FOXP3* is an X-linked gene that is known to cause fetal akinesia syndrome (Rae *et al*., 2015). The second identified novel mutation in the gene *CHRNA1* was identified in a consanguineous couple (Shamseldin *et al*., 2013). Autosomal recessive mutations in this gene are also known to cause fetal akinesia.

A single study identified a homozygous missense variant in *NOP14* in pregnancy loss material from two consanguineous Iranian couples experiencing RPL. WES was completed on fetal tissue samples and the heterozygous copies of the variant were confirmed in the parents using Sanger sequencing (Suzuki *et al*., 2018).

Studies also used WES in larger cohorts. One study (Shamseldin *et al*., 2015) looked at consanguineous couples with two or more pregnancies diagnosed with non-immune hydrops fetalis (NIHF). Seven pathogenic variants previously known to cause NIHF (Shamseldin *et al*., 2015) were identified from 24 consanguineous couples with lethal NIHF.

Two studies (Ellard *et al*., 2015, Qiao *et al*., 2016) analysed non-consanguineous couples with RPL. Variants in *GLE1*, *RYR1*, and *DYNC2H1* were identified using WES of parental samples only (Ellard *et al*., 2015). Compound heterozygous variants were also identified in *DYNC2HI* and *ALOX15* in seven euploid pregnancy losses from four families (Qiao *et al*., 2016).

The final study used a slightly different approach and analysed a panel of 234 pre-selected RPL candidate genes from women affected by RPL. Using WES and bioinformatic filtering of non-synonymous sequence variants, 27 variants were identified from the previously selected genes (Quintero-Ronderos *et al*., 2017). The genes in which variants were identified in the described sequencing studies are detailed in Table III. However, genes from Quintero-Ronderos *et al*. (2017) have been excluded because they were from a pre-selected gene panel and therefore would introduce bias.

Copy number variants

Thirteen studies and one meta-analysis (Bagheri *et al*., 2015) (Table IV) were identified which looked for CNVs in fetal tissue, parental samples, or both by chromosomal microarray analysis. Three different microarray platforms were used for analysis, either single-nucleotide polymorphism (SNP) array, oligonucleotide (oligo) array, or bacterial artificial chromosome (BAC) array.

Six studies reported CNVs in pregnancy loss (Zhang *et al*., 2009; Viaggi *et al*., 2013; Levy *et al*., 2014; Zhang *et al*., 2016; Zhou *et al*., 2016; Donaghue *et al*., 2017), four studies in RPL (Rajcan-Separovic *et al*., 2010a; Nagirnaja *et al*., 2014; Karim *et al*., 2017; Robberecht *et al*., 2012), and three studies with a mixture of both pregnancy loss and RPL (Warren *et al*., 2009; Rajcan-Separovic *et al*., 2010b; Wang *et al*., 2017). Seven of the studies included parental samples and therefore the inheritance of reported CNVs was determined. Six of the studies did not include parental samples, and therefore the inheritance pattern of the CNVs reported in these studies could not be determined.

The pregnancy losses reported were pregnancies of varying gestational age, with the majority of pregnancy losses at less than 20 weeks. In three studies (Rajcan-Separovic *et al*., 2010a; Robberecht *et al*., 2012; Viaggi *et al*., 2013), all pregnancy losses tested were less than 12 weeks gestation. Two papers (Rajcan-Separovic *et al*., 2010b; Robberecht *et al*., 2012) also identified pregnancies with developmental abnormalities and used hystero-embryoscopy to allow morphological examination of the fetus *in utero* prior to genetic analysis.

Of the studies which determined the inheritance of the CNVs, there were 30 *de novo* and 43 inherited CNVs (Warren *et al*., 2009; Rajcan-Separovic *et al*., 2010a; Rajcan-Separovic *et al*., 2010b; Robberecht *et al*., 2012; Levy *et al*., 2014; Wang *et al*., 2017). In general, the studies showed a 2.2–13% detection rate (DR) of pathogenic CNVs (Warren *et al*., 2009; Zhang *et al*., 2009; Levy *et al*., 2014; Zhang *et al*., 2016; Donaghue *et al*., 2017; Wang *et al*., 2017) plus a 0.9 % to 3.3 % DR of variants of unknown significance (VOUS) (Qiao *et al*., 2016; Zhang *et al*., 2016; Donaghue *et al*., 2017; Wang *et al*., 2017). An additional meta-analysis study (Bagheri *et al*., 2015) compared the characteristics

Table I Gene/locus names and MIM number for all genes discussed in the text.

Gene	Gene/locus name	Gene/locus MIM number
ALOX15	Arachidonate 15-lipoxygenase	152392
AURKB	Aurora kinase B	604970
C6orf221	Chromosome 6 open reading frame 221	611687
CDH1	Cadherin 1	192090
CEP55	Centrosomal protein, 55-KD	610000
CHD11	Cadherin 11	600023
CHRNA1	Cholinergic receptor, nicotinic, alpha polypeptide 1	100690
CR1	Complement component receptor 1	120620
DNMT3L	DNA methyltransferase 3-like protein	606588
DYNC2H1	DYNEIN, cytoplasmic 2, heavy chain 1	603297
ECEL1	Endothelin-converting enzyme-like 1	605896
F5	Coagulation factor V	612309
FGA	Fibrinogen, A alpha polypeptide	134820
FLT1	FMS-related tyrosine kinase 1	165070
FOXP3	Forkhead box P3	300292
FZD6	Frizzled, drosophila, homolog of, 6	603409
GBE1	Glycogen branching enzyme	607839
GLE1	RNA export mediator	603371
GOLPH3	Golgi phosphoprotein 3	612207
H19	H19, imprinted maternally expressed noncoding transcript	103280
HERG	Human ether-a-go-go-related gene	152427
IFT122	Intraflagellar transport 122	606045
KHDC3L	KHDC3-like protein, subcortical maternal complex member	611687
KIF14	Kinesin family member 14	611279
LIFR	Leukemia inhibitory factor receptor	151443
LIT1	Long QT intronic transcript 1	604115
MMP10	Matrix metalloproteinase 10	185260
MMP9	Matrix metalloproteinase 9	120361
MSH4	MutS, E. coli, homolog of, 4	602105
MuSK	Muscle, skeletal, receptor tyrosine kinase	601296
MYOM1	Myomesin 1	603508
NLRP2	NLR family, pyrin domain-containing 2	609364
NLRP5	NLR family, pyrin domain-containing 5	609658
NLRP7	NLR family, pyrin domain-containing 7	609661
NOP14	Nucleolar protein 14	611526
PDZD2	PDZ domain-containing 2	610697
RYR1	Ryanodine receptor 1	180901
SNRPN	Small nuclear ribonucleoprotein polypeptide N	182279
STIL	SCL/TAL1-interrupting locus	181590
SYCP3	Synaptonemal complex protein 3	604759
THBD	Thrombomodulin	188040
TLR3	TOLL-LIKE receptor 3	603029
TRO	Trophinin	300132
UBN1	Ubinuclein 1	609771
WNT6	Wingless-type MMTV integration site family, member 6	604663

Table II Characteristics of studies using whole-exome sequencing to identify single-nucleotide variants.

Author, date and country	Study objective	Sample, phenotype and gestation	Methods and analysis	Study outcome	Quality scores
Rae *et al*., 2015 *Southampton, UK*	Assessment of genes causing pregnancy loss of male pregnancies.	One woman with multiple male pregnancy loss (hydrops fetalis). Two fetal tissue samples at 18 and 20 weeks gestation.	Fetal DNA screening using NGS panel (Sure Select XT kit—no information on gene number given).	Identified pathogenic variant in X-linked *FOXP3* gene. Known to cause IPEX syndrome, but gene not previously linked to NIHF.	3
Shamseldin *et al*., 2013 *Riyadh, Saudi Arabia*	Identify causative mutation in a family with RPL due to NIHF.	Consanguineous couple with two previous fetal losses and a history of fetal hydrops. One fetal tissue at 19 weeks gestation.	WES of affected fetus and filtering of variants by autozygome.	Identified novel mutation in *CHRNA1*, known to cause multiple pterygium and fetal akinesia syndrome.	6
Filges *et al*., 2014 *Vancouver, Canada*	Identify causal variants for a recurrent pattern of an undescribed lethal fetal congenital anomaly syndrome.	Two pregnancy tissue samples with normal female karyotypes at 21+4 and 18+5 weeks gestation from one family (non-consanguineous).	WES of DNA from parents and chorionic villus samples. Variants from trio filtered for autosomal recessive inheritance (homozygous or compound heterozygous).	Compound heterozygosity for two non-synonymous truncating mutations in *KIF14* in the one fetus.	4
Suzuki *et al*., 2018 *Yokohama, Japan*	Identify causative variant(s) of two consanguineous families with RPL.	Two Iranian consanguineous families with RPL (nine and two pregnancy loss in each family).	WES on fetal tissue samples. Filtering for compound heterozygosity and homozygous variants.	Homozygous variant in *NOP14* gene in both pregnancy losses with parents having heterozygous variants confirmed by Sanger sequencing.	3
Dohrn *et al*., 2015 *Aalborg, Denmark*	Clinical report of consanguineous couple with a history of pregnancies with arthrogryposis multiplex congenita.	Two pregnancy tissue samples terminated at 14+2 and 13+4 weeks gestation from one consanguineous family.	WES of parents and fetuses. Variants filtered for autosomal recessive inheritance (homozygosity in fetus).	Homozygous missense variant in *ECEL1*.	3
Wilbe *et al*., 2015 *Uppsala, Sweden*	WES on family trio with a history of recurrent fetal loss and fetal akinesia deformation syndrome.	One family trio (mother, father and 17-week fetal tissue) with five affected fetuses with severe/lethal form of FADS.	WES on parental and one affected fetus. Variant filtering to identify potentially damaging variants following a recessive inheritance pattern.	Novel homozygous mutation in *MuSK* gene leading to a frameshift mutation and predicting a premature stop codon.	7
Cristofoli *et al*., 2017 *Leuven, Belgium*	Identify causative mutation in a family with profound fetal microcephaly associated with delayed gyrification of the corpus callosum.	One couple with five first trimester pregnancy losses. Two fetal tissue terminations sampled at 20 and 25 weeks gestation.	WES on fetus and parents using OneSeq Constitutional Research Panel for exome enrichment. Trio filtering of variants for recessive, *de novo* and hemizygous inheritance.	Identification of novel compound heterozygous mutation in *STIL* in both fetuses.	6
Bondeson *et al*., 2017 *Uppsala, Sweden*	Uncover the genetics of suspected autosomal recessive lethal fetal ciliopathy Meckel-like syndrome in a family with two affected fetuses with Meckel-like syndrome and one healthy sibling.	Family with four pregnancies (one pregnancy loss, one TOP, one IUFD, one healthy girl).	WES of trio (20 week TOP) with variant filtering for potentially damaging variants following recessive inheritance pattern. SNPs identified investigated for shared ancestral haplotype analysis.	Homozygous nonsense mutation in *CEP55* causing a premature stop codon in the affected fetus. The IUFD fetus was homozygous and the mother, father and healthy sibling were all heterozygous for the mutation.	5

Continued

Table II *Continued*

Author, date and country	Study objective	Sample, phenotype and gestation	Methods and analysis	Study outcome	Quality scores
Tsurusaki *et al*., 2014 *Yokohama, Japan*	Identify causative mutation in a family with RPL and scan abnormalities.	39-year-old women with two terminations (13 and 21 weeks), one intrauterine fetal death (13 weeks), four RPL (6–8 weeks) and one healthy son.	WES for maternal sample, healthy son and 21 week TOP due to skeletal anomalies.	Compound heterozygous mutation in *IFT122* identified in the TOP. Same heterozygous mutation was found in pregnancy loss tissue at 7 weeks by PCR analysis. One mutation identified in maternal sample. Mutation known to cause cranioectodermal dysplasia-1.	33
Shamseldin *et al*., 2015 *Riyadh, Saudi Arabia*	Analysis of consanguineous couples in which embryonic lethality appears to follow a Mendelian recessive pattern.	Twenty-four consanguineous couples with two or more pregnancies diagnosed with lethal NIHF.	Autozygosity mapping using genome-wide genotyping array from affected fetus. WES on affected fetus filtered for the autozygome and novel homozygous variants.	In eight fetuses, variants known to cause NIHF were detected in four genes (same gene affected in five families). Mutations in seven novel candidate genes were identified in 10 fetuses (same gene affected in four families).	5
Ellard *et al*., 2015 *Exeter, UK*	Investigate a strategy for diagnosis of autosomal recessive lethal disorders.	Three couples with multiple pregnancy terminations due to presumed autosomal recessive disorder.	WES of parental DNA samples only, filtering variants for autosomal recessive inheritance (homozygous or compound heterozygous). Pregnancy tissue subsequently tested for variants using sanger sequencing.	Proof of principle study detected variants in *GLE1*, *RYR1* and *DYNC2H1* (fetal akinesia syndrome). Demonstrates that WES of parental samples can be effective in diagnosing lethal recessive disorders.	8
Qiao *et al*., 2016 *British Columbia, Canada*	WES in recurrent pregnancy loss.	Seven fetal tissues from four families with RPL at <20 weeks gestation.	WES on couples and pregnancy tissue. Variant filtering for autosomal recessive, compound heterozygous and X-linked inheritance; pathway and disease association enrichment analysis.	Compound heterozygous variants identified in *DYNC2HI* and *ALOX15*.	5
Quintero-Ronderos *et al*., 2017 *Bogotá, Colombia*	Analyse non-synonymous sequence variants in RPL candidate genes for dissection of complex disease aetiology.	Forty-nine women with RPL at <20 weeks gestation.	WES of women (not partners), bioinformatic filtering of non-synonymous sequence variants in subset of 234 RPL candidate genes selected from review of literature and databases.	Twenty-seven variants, in 22 genes, affecting 20 patients were considered strong candidates for relating to RPL molecular aetiology.	7

Quality scores were assessed using a modified Newcastle–Ottawa scale.
FADS, fetal akinesia deformation sequence; IPEX, immunodysregulation polyendocrinopathy enteropathy X-linked; IUFD, intrauterine fetal death; NGS, next generation sequencing; NIHF, non-immune hydrops fetalis; RPL, recurrent pregnancy loss; SNP, single-nucleotide polymorphism; STIL, SCL/TAL1-interrupting locus; TOP, termination of pregnancy.
ALOX15, 15-lipoxygenase, reticulocyte arachidonate; *CEP55*, centrosomal protein, 55-KD; *CHRNA1*, cholinergic receptor, nicotinic, alpha polypeptide 1; *DYNC2H1*, dynein, cytoplasmic 2, heavy chain 1; *ECEL1*, endothelin-converting enzyme-like 1; *FOXP3*, forkhead box P3; *GLE1*, RNA export mediator; *IFT122*, intraflagellar transport 122; *KIF14*, kinesin family member 14; *MuSK*, muscle, skeletal, receptor tyrosine kinase, *NOP14*, nucleolar protein 14; RYR1, ryanodine receptor 1.

Table III **Genes identified by whole-exome sequencing studies.**

Gene	Paper author	Gene/locus name	Gene/locus MIM number	Gene function	Associated disease/phenotype	Cytogenetic location
ALOX15	Qiao *et al.*, 2016	15-LIPOXYGENASE, RETICULOCYTE ARACHIDONATE	152392	Implicated in anti-inflammation, membrane remodelling, and cancer development/metastasis.	None	17p13.2
CEP55	Bondeson *et al.*, 2017	CENTROSOMAL PROTEIN, 55-KD	610000	Regulator of the final stages of mitosis.	Multinucleated neurons, anhydramnios, renal dysplasia, cerebellar hypoplasia, and hydranencephaly	10q23.33
CHRNA1	Shamseldin *et al.*, 2013, Shamseldin *et al.*, 2015	CHOLINERGIC RECEPTOR, NICOTINIC, ALPHA POLYPEPTIDE 1	100690	Nicotinic acetylcholine receptor, electrical signalling between nerve and muscle cells.	Myasthenic syndrome, congenital, Multiple pterygium syndrome	2q31.1
CTSA	Shamseldin *et al.*, 2015	CATHEPSIN A	613111	Multifunctional enzyme, lysosomal multienzyme complex.	Galactosialidosis	20q13.12
DNAH14	Shamseldin *et al.*, 2015	DYNEIN, AXONEMAL, HEAVY CHAIN 14	603341	Microtubule-associated motor protein complexes.	None	1q42.12
DYNC2H1	Ellard *et al.*, 2015, Qiao *et al.*, 2016	DYNEIN, CYTOPLASMIC 2, HEAVY CHAIN 1	603297	Ciliary intraflagellar transport (IFT).	Short-rib thoracic dysplasia 3 with or without polydactyly	11q22.3
ECEL1	Dohrn *et al.*, 2015	ENDOTHELIN-CONVERTING ENZYME-LIKE 1	615065	Zinc metalloproteases.	Neuromuscular disease /arthrogryposis multiplex congenita, distal arthrogryposis	2q37.1
FOXP3	Rae *et al.*, 2015	FORKHEAD BOX P3	300292	Development and function of naturally occurring CD4 positive T regulatory cells.	Immunodysregulation, polyendocrinopathy, and enteropathy, X-linked	Xp11.23
FZD6	Shamseldin *et al.*, 2015	FRIZZLED, DROSOPHILA, HOMOLOG OF, 6	603409	Wnt family member, directional cues to align either individual cells within an epithelial sheet or multicellular clusters.	Nail disorder, nonsyndromic congenital, 10, (claw-shaped nails)	8q22.3
GALNT14	Shamseldin *et al.*, 2015	UDP-N-ACETYL-ALPHA-D-GALACTOSAMINE:POLYPEPTIDE N-ACETYLGALACTOSAMINYLTRANSFERASE 14	608225	Glycosyltransferases, enzymes catalyse the first step in the O-glycosylation.	None	2p23.1
GLE1	Ellard *et al.*, 2015	RNA EXPORT MEDIATOR	603371	Regulates gene expression.	Arthrogryposis, lethal congenital contracture syndrome 1	9q34.11
GUSB	Shamseldin *et al.*, 2015	BETA-GLUCURONIDASE	611499	Lysosomal hydrolase (involved in the stepwise degradation of glucuronic acid-containing glycosaminoglycans).	Mucopolysaccharidosis VII	7q11.21

Continued

Table III *Continued*

Gene	Paper author	Gene/locus name	Gene/locus MIM number	Gene function	Associated disease/phenotype	Cytogenetic location
KIF14	Filges *et al.*, 2014	KINESIN FAMILY MEMBER 14	611279	Intracellular transport and cell division.	Meckel syndrome 12, microcephaly	1q32.1
IFT122	Tsurusaki *et al.*, 2014	INTRAFLAGELLAR TRANSPORT 122	606045	Cell cycle progression, signal transduction, apoptosis, and gene regulation.	Cranioectodermal dysplasia 1 (DEC-1)	3q21.3-q22.1
MuSK	Wilbe *et al.*, 2015	MUSCLE, SKELETAL, RECEPTOR TYROSINE KINASE	601296	Muscle-specific kinase for neuromuscular junction formation.	FADS, Myasthenic syndrome, congenital, 9, associated with acetylcholine receptor deficiency	9q31.3
MYOM1	Shamseldin *et al.*, 2015	MYOMESIN 1	603508	Striated muscle.	none	18p11.31
NEB	Shamseldin *et al.*, 2015	NEBULIN	161650	Protein component of the cytoskeletal matrix within skeletal muscle.	Nemaline myopathy 2, autosomal recessive	2q23.3
NOP14	Suzuki *et al.*, 2018	NUCLEOLAR PROTEIN 14	611526	18S rRNA processing and 40S ribosome biogenesis.	None	4p16.3
PIGC	Shamseldin *et al.*, 2015	PHOSPHATIDYLINOSITOL GLYCAN ANCHOR BIOSYNTHESIS CLASS C PROTEIN	601730	Glycosylphosphatidylinositol biosynthesis.	None	1q24.3
RYR1	Ellard *et al.*, 2015, McKie *et al.*, 2014	RYANODINE RECEPTOR 1	180901	Calcium release channel of the sarcoplasmic reticulum.	Central core disease, King–Denborough syndrome, minicore myopathy with external ophthalmoplegia, neuromuscular disease, congenital, with uniform type 1 fiber	19q13.2
STIL	Cristofoli *et al.*, 2017	SCL/TAL1-INTERRUPTING LOCUS	181590	Centriole duplication during the cell cycle.	Microcephaly 7, primary, autosomal recessive	1p33
THSD1	Shamseldin *et al.*, 2015	THROMBOSPONDIN TYPE 1 DOMAIN-CONTAINING PROTEIN 1	616821	Adhesive glycoprotein.	None	13q14.3
UBN1	Shamseldin *et al.*, 2015	UBINUCLEIN 1	609771	Proliferation.	None	16p13.3

Table IV Characteristics of studies using microarray analysis to look for CNVs.

Reference	Sample size	Clinical information		Array	Results from POC analysis			Quality score
		Sample	Gestation		CNVs	Size	Inherited vs. De novo	
Warren *et al.*, 2009 *Utah, USA*	30	POC-miscarriage	10–20 weeks	BAC and Oligo	4 CNVs (3 gains, 1 loss)	All <300 kb	4 *de novo*	4
Zhang *et al.*, 2009 *Beijing, China*	58	POC-miscarriage	First trimester	Oligo	5 CNVs (4 gains, 1 loss)	108–1460 kb	No parental samples	5
Rajcan-Separovic *et al.*, 2010b *Vancouver, Canada*	17	POC-miscarriage with abnormal morphology	Not given (CRL 4–30 mm)	BAC and Oligo	6 CNVs (in 5 miscarriages) (4 gains, 2 losses)*	<250 kb	2 maternal, 1 paternal, 2 *de novo*, 1 unknown	4
Viaggi *et al.*, 2013 *Genova, Italy*	40	POC-miscarriage	<12 weeks	Oligo	14 CNVs (9 gains, 5 losses)*	<5 Mb	No parental samples	5
Levy *et al.*, 2014 *New York, USA*	2389	POC-miscarriage	<20 weeks (mean 7.7)	SNP	33 CNVs—12 clinically significant, 21 VOUS	All <10 Mb	Clinically significant—5 maternal, 6 paternal, 1 unknown. VOUS—17 maternal, 3 paternal, 1 unknown	6
Zhang *et al.*, 2016 *Guangzhou, China*	60	POC-miscarriage	<20 weeks (mean 17.2)	SNP and Oligo	4 pCNVs, 2 VOUS	3 >10 Mb, 3 <10 Mb	No parental samples	6
Zhou *et al.*, 2016 *New Haven, USA*	234	POC-128 culture failure (CF-POC) and 106 normal karyotype (NK-POC)	6–37 weeks	Oligo	5 pCNVs from CF-POC, 6 pCNVs from NK-POC	10 <10 Mb, 1 >10 Mb (from CF-POC)	No parental samples	9
Donaghue *et al.*, 2017 *London, UK*	1911	POC-miscarriage	8–40 weeks (mean 16)	Oligo	260 CNVs	197 (7.2%) >10 Mb, 25 (0.9%) <10 Mb, 24 (0.9) VOUS, 14 (0.5%) IF	No parental samples	5
Rajcan-Separovic *et al.*, 2010a *Vancouver, Canada*	26	POC-RPL	<20 weeks	Oligo	13 CNVs (6 gains, 7 losses)*	<1600 kb	6 maternal, 7 paternal	7
Robberecht *et al.*, 2012 *Leuven, Belgium*	32	POC-miscarriage with abnormal morphology (1–6 previous losses)	<11 weeks (mean 7.5 weeks)	BAC and SNP	11 CNVs (8 gains, 3 losses)*	<1 Mb	2 *de novo*, 9 inherited	5
Wang *et al.*, 2017 *Nanjing, China*	535	POC-SA and RM	<13 weeks (mean 9.8)	SNP	31 pCNVs	12 <10 Mb, 19 >10 Mb	1 maternal, 3 paternal, 22 *de novo*, 5 lost at follow up	9
Nagirnaja *et al.*, 2014 *Tartu, Estonia*	309 female, 249 male	Parental blood samples—RPL (≤3 miscarriages)	<22 weeks	SNP	N/A	N/A	N/A	8
Karim *et al.*, 2017 *Jeddah, Saudi Arabia*	44	Parental samples with RPL (29.5% >5 miscarriages)	79.5% first trimester, 16% second trimester, 4.5% third trimester	Oligo	36 CNVs (25 gains, 11 loses)	all <10 Mb	N/A	7
Bagheri *et al.*, 2015 *Vancouver, Canada*	101	Meta-analysis of studies			24 rare CNVs		14 familial, 1 *de novo*, 9 unknown origin	7

*These CNVs were subsequently reviewed and reclassified by Bagheri *et al.* (2015).
Quality scores were assessed using a modified Newcastle–Ottawa scale.
Array, chromosomal microarray analysis; BAC, bacterial artificial chromosome microarray; CNV, copy number variation; CRL, crown to rump length; IF, incidental finding; Oligo, oligonucleotide microarray; pCNV, pathogenic copy number variant; POC, products of conception; VOUS, variants of unknown significance.

and contributions of rare and common CNVs from four of the other studies by reclassifying CNVs according to the prevalence of healthy controls using Database of Genomic Variants (Rajcan-Separovic *et al.*, 2010a; Rajcan-Separovic *et al.*, 2010b; Robberecht *et al.*, 2012; Viaggi *et al.*, 2013; Bagheri *et al.*, 2015). They concluded that common CNVs were specifically enriched in immunological pathways and rare CNVs were not, although the small number of rare CNVs may have hampered this conclusion. However, both rare and common CNVs could have a role in pregnancy loss, as rare CNVs have a two times higher gene density and contain more genes studied in mouse knockouts and common CNVs contain more genes in biological pathways relevant to pregnancy. The studies that identified VOUS were in accordance with each other and suggested the rate of 2–3%.

Of particular interest is to find recurrent CNVs that are associated with pregnancy loss. Maisenbacher *et al.* (2017) determined the frequency of the 22q11.2 deletion in a large cohort of pregnancy loss samples using an SNP microarray. The 22q11.2 deletion was detected in 15 (0.07%) of 22 451 POCs, with an overall incidence of 1/1497. They concluded that this was higher than the reported general population prevalence (1/4000–1/6000). Likewise, Nagirnaja *et al.* (2014) identified CNV regions on chromosome 5 (5p13.3), disrupting the *PDZD2* and *GOLPH3* genes. There was significant association with an increased risk of RPL. *PDZD2* and *GOLPH3* are predominately expressed in the placenta, suggesting a functional relevance; however, neither of these genes have previously been linked to placental function or pregnancy complications (Nagirnaja *et al.*, 2014).

Recurrent molar pregnancies

Eleven studies (Table V) were identified that evaluated the genetics of diploid and biparental recurrent HM (RHM) pregnancies. One study (Parry *et al.*, 2011) identified biallelic mutations in *C6orf221* in three consanguineous families with familial biparental HM. Three studies (Abdalla *et al.*, 2012, Brown *et al.*, 2013, Ulker *et al.*, 2013) reported case studies of an individual consanguineous family, two non-consanguineous families, and two consanguineous families with RHM. Autosomal recessive mutations were identified in the *NLRP7* gene and were considered to be responsible for the occurrence of HM. Deveault *et al.* (2009) investigated 13 women experiencing RHM, some with a family history of molar pregnancies and 11 *NLRP7* variants were identified. Mutation analysis of the *NLRP7* gene in 35 women experiencing RPL with at least one HM revealed 17 different mutations (Qian *et al.*, 2011). Qian *et al.* (2011) also suggested that one defective allele in *NLRP7* causes diploid androgenic moles and two defective alleles cause diploid biparental moles.

Two studies (Messaed *et al.*, 2011; Huang *et al.*, 2013) investigated cohorts of women to see whether mutations in the *NLRP7* gene could also be responsible for RPL without history of molar pregnancy. Messaed *et al.* (2011) investigated 135 women with either RPL or at least one HM and sequencing of *NLRP7* exons identified two patients with RPL to have *NLRP7* mutations. Huang *et al.* (2013) also showed significant association between RPL and *NLRP7* polymorphisms. In contrast, two further studies (Andreasen *et al.*, 2013, Manokhina *et al.*, 2013) identified no disease-causing mutations in *NLRP7* in women with RPL, and similarly, Aghajanova *et al.* (2015) found no mutations in *NLRP7*, *NLRP2*, or *KHDC3L* (*C6orf221*).

Other genetic causes

Two studies (Seyedhassani *et al.*, 2010a, Seyedhassani *et al.*, 2010b) analysed and sequenced mitochondrial tDNA 96 women with RPL. Four variants in threonine tRNA and one variant in proline tRNA were observed, but in some cases these were also observed in controls (Seyedhassani *et al.*, 2010a), which calls into question the significance of these findings. Analysis of mitochondrial D-loop sequences showed a higher rate of point mutations in RPL patients than in controls. In total, 89 out of 153 variants were only identified in women with RPL and 22 of these mutations were considered to be significant (Seyedhassani *et al.*, 2010b).

X-chromosome inactivation occurs during early embryogenesis and has also been proposed to have an aetiological role in RPL. Skewed X-chromosome inactivation (XCI) status was compared between women with RPL and healthy controls. Extremely skewed XCI (defined as >90%) was identified in 17.7% of women with RPL compared to 1.6% of extremely skewed XCI in controls (Bagislar *et al.*, 2006).

Six further papers were identified that discussed specific genes and their contribution to pregnancy loss. Each paper (Bhuiyan *et al.*, 2008; Stouffs *et al.*, 2011; Lopez-Carrasco *et al.*, 2013; McKie *et al.*, 2014, Bendroth-Asmussen *et al.*, 2016; Zhang *et al.*, 2016) investigated an individual gene or genes. In a case study of a 30-year-old women with pregnancy loss from glycogen storage disease type IV (GSD-IV), DNA extracted from placental tissue identified compound heterozygous mutations in *GBE1* (Bendroth-Asmussen *et al.*, 2016).

Another case study, a consanguineous Arabian family with pregnancy losses, stillborn, fetal demise, and two live children, had homozygosity mapping. This led to the screening of the *HERG* gene in the live children, parents, and stillborn. Homozygous nonsense mutations in *HERG* were identified in the child with polymorphic ventricular tachycardia and the same heterozygous mutation in the parents and unaffected child. Amniotic fluid cells from the stillborn child were also homozygous for the same *HERG* mutation (Bhuiyan *et al.*, 2008).

Three rare homozygous *RYR1* variants were identified using genome-wide linkage studies and sequencing of *RYR1* coding exons. Initially, a *RYR1* homozygous nonsense mutation was detected in two fetuses with FADS/lethal multiple pterygium syndrome (LMPS). The parents were both homozygous for the same mutation. When 66 further probands with FADS/LMPS phenotype were screened for germline *RYR1* mutations, two further potential homozygous mutations were detected (McKie *et al.*, 2014).

In a larger study, 100 couples with at least three unexplained pregnancy losses had *WNT6* mutation analysis performed. *WNT6* has previously been shown to have an important role for stromal cell proliferation during decidualization in mice. Four novel mutations were identified in the women with RPL but not in the male partners or healthy controls (Zhang *et al.*, 2015), although there was no conclusive evidence for pathogenicity.

Ten aberrations were identified in *MSH4*, *DNMT3L*, and *SYCP3* in 23 couples with RPL. Six of these aberrations were predicted to alter the amino acid sequence. All but one of these aberrations was considered a likely SNV. The mutation in the *SYCP3* gene was shown to have a 78% likelihood of causing a deleterious effect on protein function due to an alteration in the amino acid sequence changing a non-polar isoleucine into a polar threonine (Stouffs *et al.*, 2011). Another study (Lopez-Carrasco *et al.*, 2013) targeted the two spindle checkpoint genes

Table V Characteristics of studies looking at other genetic findings in pregnancy loss including recurrent molar pregnancies.

	Author, date and country	Study objective	Sample, phenotype and gestation	Methods and analysis	Study outcome	Quality scores
Recurrent Molar	Deveault *et al.*, 2009 *Montreal, Canada*	Investigate the role of *NLRP7* in recurrent HMs.	Thirteen patients, each with RHM or recurrent reproductive wastage. Five with family history of molar pregnancy, eight with no fetal heart.	Sequencing of *NLRP7* exons.	Eleven *NLRP7* variants in parents with diploid biparental, diploid androgenic, triploid, and tetraploid moles.	6
Recurrent Molar	Parry *et al.*, 2011 *Leeds, UK*	Report biallelic mutations of *C6orf221* in three families with FBHM.	Three consanguineous families with FBHM, and 14 further probands.	Homozygosity sampling using genome-wide SNP array, NGS enriching for homozygous regions.	Biallelic mutations in *C6orf221* in three families.	4
Recurrent Molar	Abdalla *et al.*, 2012 *Alexandria, Egypt*	Mutation analysis of *NLRP7* and *C6orf221* genes in women with RHM.	Three women from two families with first trimester losses with repeat HMs.	Cytogenetic analysis; sequence analysis of *NLRP7* and *C6orf221* genes.	Two autosomal recessive mutations were identified in *NLRP7*.	4
Recurrent Molar	Brown *et al.*, 2013 *Vermont, USA*	Case study of one patient.	One woman with successive molar pregnancies in the first trimester. One complete HM and two partial moles.	Chromosomal microarray analysis of DNA from peripheral blood.	Homozygosity for clinically relevant *NLRP7* mutation. SNP array ruled out deletion.	4
Recurrent Molar	Ulker *et al.*, 2013 *Istanbul, Turkey*	Investigate genetic disposition of families with RHM pregnancies.	Two consanguineous Turkish families with recurrent HM.	Sequencing of *NLRP7* axons and SNP array.	One family with homozygous *NLRP7* nonsense mutation. Another family with heterozygous deletion affecting the *NLRP2* and *NLRP7* genes.	3
RPL/ Recurrent Molar	Qian *et al.*, 2011 *Zhejiang, China*	*NLRP7* mutation analysis in Chinese patients with recurrent reproductive wastage, including at least one HM.	Thirty-five women with RPL with at least one HM.	Mutation analysis of NLRP7 exons.	Founder mutations in *NLRP7*—diploid androgenic moles in patients with one defective allele, diploid biparental moles in patients with two defective alleles.	9
RPL / Recurrent Molar	Messaed *et al.*, 2011 *Montreal, Canada*	Investigate the role of *NLRP7* in sporadic moles and other forms of reproductive wastage.	One hundred thirty-five women with three spontaneous abortions or at least one HM.	Sequencing of *NLRP7* exons.	*NLRP7* is responsible for some cases of recurrent spontaneous abortion.	9
RPL/ Recurrent Molar	Andreasen *et al.*, 2013 *Aarhus, Denmark*	Investigate the association between molar pregnancy and RPL regarding changes in the *NLRP7* and *C6orf221/ KHDC3L* genes.	Nineteen women with RPL and one HM, five women with recurrent HM, seven women with HM with family members with HM.	DNA from maternal blood—NLRP7 and *KHDC3L* were sequenced using PCR.	No unequivocal pathogenic mutations in *NLRP7* or *KHDC3L*, although eight rare non-synonymous variants (NSVs) in *NLRP7* were observed and three NSVs in *KHDC3L*.	7
RPL	Huang *et al.*, 2013 *Tainan, Taiwan*	Do genetic variants of *NLRP2* and *NLRP7* confer susceptibility to idiopathic RPL?	One hundred forty-three women with a least two consecutive pregnancy losses at <12 weeks.	Genomic DNA extracted from lymphocytes and SNPs identified using end-point TaqMan assays from genomic DNA.	An SNP in *NLRP7* showed significant association with idiopathic RPL. An SNP in *NLRP2* showed a marginally significant association.	10

Continued

Table V *Continued*

	Author, date and country	Study objective	Sample, phenotype and gestation	Methods and analysis	Study outcome	Quality scores
RPL	Manokhina *et al.*, 2013 *Vancouver, Canada*	Elucidate which subpopulations of women with adverse reproductive outcomes should be screened for *NLRP7/ C6orf221*.	Seventeen women with RPL or complete HM.	Peripheral blood DNA sequencing of *NLRP7* and *C6orf221 exons*. Six non-synonymous *NLRP7* variants genotyped in larger clinical groups.	No association between *NLRP7* variants and RPL or partial HM.	8
RPL	Aghajanova *et al.*, 2015 *Houston, USA*	Are mutations in *NLRP2*, *NLRP7*, or *KHDC3L* associated with RPL or infertility?	Ninety-four women with unexplained infertility and 24 women with RPL.	Sequencing of coding exons of *NLRP7*, *NLRP2* and *KHDC3L* in genomic DNA.	No disease-causing mutations were identified in *NLRP2*, *NLRP7*, and *KHDC3L* in either unexplained infertility or RPL.	6
Imprinting/molar	Docherty *et al.*, 2015 *Southampton, UK*	Report genetic causes in multi-locus imprinting disturbances (MLID) patients, to seek genetic causes of the epigenetic errors in MLID.	Thirty-nine MLID patients and 33 mothers of patients with MLID.	WES and sanger sequencing to confirm NLRP5 variants, target methylation analysis using methylation-specific PCR.	Identification of *NLRP5* variants in five mothers of patients with MLID and hypomethylation of imprinted loci observed in MLID patients exposed to maternal *NLRP5* variants.	6
Mitochondrial DNA	Seyedhassani *et al.*, 2010a *Tehran, Iran*	Analysis of mitochondrial tRNA (Thr) and tRNA (Pro) in women with RPL.	Blood samples from 96 women with at least three pregnancy losses at <20 weeks gestation.	Mitochondrial proline and threonine tRNAs were sequenced.	Four mutations in tRNA (Thr) and one mutation in tRNA (Pro) were identified in some women.	5
Mitochondrial DNA	Seyedhassani *et al.*, 2010b *Tehran, Iran*	Examine mitochondrial D-loop deletions and nucleotide alterations in samples taken from RPL women.	Ninety-six blood samples from women with at least three pregnancy losses at <20 weeks gestation.	Mitochondrial D-loop DNA sequence analysis.	One hundred fifty-three different mutations in D-loop region. Higher rate of mutations in D-loop in maternal blood of women with RPL in comparison to control group.	10
X-Chromosome inactivation	Bagislar *et al.*, 2006 *Ankara, Turkey*	Is there a difference in X-chromosome inactivation patterns in patients with RPL?	Eighty patients with RPL and 160 controls.	Methylation-sensitive restriction enzyme digest and PCR. Radioactive labelling and densitometry analyses.	Skewed X chromosome inactivation in 20.9% RPL vs. 6.4% controls, with extreme skewing in 17.7 % RPL vs. 1.6 % controls.	8
GBE1	Bendroth-Asmussen *et al.*, 2016 *Copenhagen, Denmark*	Case report of 30-year-old woman with pregnancy loss in the first trimester. Presenting glycogen storage disease type IV (GSD-IV) as a cause of early pregnancy loss.	Pregnancy tissue at 10+1 weeks gestation.	Placental histopathology—with periodic acid—Schiff staining and immune-histochemical stains and sequencing of *GBE1*.	Compound heterozygosity for mutations in *GBE1*. GSD-IV can cause early pregnancy loss, and diagnosis can be made on histopathological examination.	2
HERG	Bhuiyan *et al.*, 2008 *Amsterdam, Netherlands (Multi centre)*	Identify causes of RPL and late fetal loss due to arrhythmias in a consanguineous Arabian couple.	Male child with polymorphic ventricular tachycardia.	Homozygosity mapping of cardiac ion channel genes.	Homozygous nonsense mutation in *HERG* (consanguineous parents both had the same heterozygous mutation).	3

Continued

Table V *Continued*

	Author, date and country	Study objective	Sample, phenotype and gestation	Methods and analysis	Study outcome	Quality scores
RYR1	McKie *et al.*, 2014 *Cambridge, UK*	Identify any novel genetic causes of the FADS syndrome and multiple pterygium syndrome (MPS) phenotypes.	Two pregnancy tissues affected by FADS/ MPS at 12+6 and 14+0 weeks gestation. Thirty-six families with FADS/LMPS and 30 families with Escobar variant of MPS.	Genome-wide linkage using DNA from fetal material and sequencing of *RYR1* coding exons in probands.	Linkage study identified three rare homozygous *RYR1* variants that were also identified in families with FADS/LMPS phenotypes.	5
WNT6	Zhang *et al.*, 2015 *Jinan, China*	Investigate whether mutations in WNT6 play a role in unexplained RPL.	One hundred couples with at least three pregnancy loss; 100 matched controls.	Mutation analysis of *WNT6* exons.	Four novel mutations in women with unexplained pregnancy loss, none in males or controls. No conclusive evidence for pathogenicity.	8
22q11.2	Maisenbacher *et al.*, 2017 *San Carlos, USA*	Determine the incidence of the 22q11.2 deletion in the pregnancy loss population.	Twenty-two thousand five hundred forty-one POC specimens received for pregnancy loss testing.	Genotypes using SNP arrays across the genome.	Fifteen samples with 22q11.2 deletion (0.07% of the sampled population), 12 from normal karyotypes and two from abnormal karyotypes.	6
Maturation Arrest	Stouffs *et al.*, 2011 *Brussels, Belgium*	Gain insight into maturation arrest and the relationship with mutations in genes essential for meiosis and the relevance to RPL.	Forty azoospermic patients and 23 couples with at least two consecutive pregnancy losses.	Genomic DNA sequencing of *DNMT3L*, *SYCP3* and *MSH4* genes.	Five and six aberrations affecting amino acid sequence in azoospermic men and RPL patients, respectively. Some aberrations were also identified in controls.	7
Spindle Checkpoint genes	Lopez-Carrasco *et al.*, 2013 *Valencia, Spain*	Investigate involvement of spindle checkpoint genes in patients with chromosomal instability and reproductive problems.	One hundred two patients (43 of which have RPL and 46 with spermiogram alterations).	Genomic DNA sequencing of *AURKB* and *SYCP3* genes.	Three rare heterozygous missense variants in *AURKB*.	10
Epigenetics	Zheng *et al.*, 2013 *Guangzhou, China*	Do assisted reproduction techniques affect DNA methylation of imprinted genes and can aberrant methylation of imprinted genes account for pregnancy loss?	Three hundred twenty pregnancy tissue samples from pregnancy loss after ART or natural conception between 6–26 weeks.	Paternally methylated gene, *H19*, and maternally methylated genes, *LITI* and *SNRPN*, were analysed using bisulfite pyrosequencing of genomic DNA and methylation analysis.	Two regions of hyper-methylation found in *H19*. Mean percentage of methylation higher in pregnancy loss samples. Positive correlation between percentages of methylation of all three genes and pregnancy loss.	11

Quality scores were assessed using a modified Newcastle–Ottawa scale.
FBHM, familial biparental hydatidiform mole; HM, hydatidiform mole; RHM, recurrent hydatidiform mole.
C6orf221, chromosome 6 open reading frame 221; *DNMT3L*, DNA methyltransferase 3-like protein; *GBE1*, glycogen branching enzyme; *H19*, H19, imprinted maternally expressed noncoding transcript; *HERG*, human ether-a-go-go-related gene; *KHDC3L*, KHDC3-like protein, subcortical maternal complex member; *LIT1*, long qt intronic transcript 1; *MSH4*, MutS, E. coli, homolog of, 4; *NLRP2*, NLR family, pyrin domain-containing 2; *NLRP5*, NLR family, pyrin domain-containing 5; *NLRP7*, NLR family, pyrin domain-containing 7; *SNRPN*, small nuclear ribonucleoprotein polypeptide N; *SYCP3*, synaptonemal complex protein 3; *WNT6*, wingless-type MMTV integration site family, member 6.

AURKB and *SYCP3* in 102 patients with either RPL or spermiogram alterations. One heterozygous intronic deletion was identified in *SYCP3* with no *in silico* causative indication. Six aberrations were identified in *AURKB*; however, a deletion and two nucleotide changes were considered to have no functional alteration or be frequent variants, respectively. Three rare missense variants were identified in *AURKB*, with two of these variants found in a couple with pregnancy loss.

Discussion

In this systematic review, we have identified 50 papers that investigated genetic contributions other than aneuploidy to pregnancy loss. The studies highlight some key areas, including identification of SNVs by WES, identification of CNVs by microarray analysis, and investigation of a group of genes associated with diploid and biparental recurrent molar pregnancies that are linked to pregnancy loss. Other genetic contributions, such as epigenetics and mitochondrial DNA (mtDNA), were also investigated in individual papers. There were also studies reporting sequencing of candidate genes already known to be associated with pregnancy loss with or without structural abnormalities.

We have summarized the current evidence below for each of these categories and then discuss the implications of these findings both for future studies and for genetic investigation of couples experiencing RPL.

Whole-exome sequencing

Advances in next generation sequencing are vastly improving and enabling a molecular diagnosis for a range of disorders and clinical pathways. As the cost of WES decreases, the technology is becoming more widely used and clinically applicable. This review identified a number of studies (Table II) over the last 4 years that have used WES to look for as yet unidentified genetic causes of pregnancy loss. The majority of these studies looked at individual patients or couples with RPL, some of which showed ultrasound scan abnormalities during the pregnancy (Tsurusaki *et al*., 2014; Wilbe *et al*., 2015; Bondeson *et al*., 2017; Cristofoli *et al*., 2017). More recently, a small number of studies have been published studying larger cohorts of patients and exploring possible strategies for genetic investigation of these patients (Ellard *et al*., 2015; Shamseldin *et al*., 2015; Qiao *et al*., 2016). This review included studies where patients suffered multiple pregnancy losses with phenotypic findings in all or some of their pregnancy losses. This included ultrasound scan abnormalities and post-mortem findings, and in some cases, where patients opted for termination of pregnancy. These were thought to be important to include because there could be a range of phenotypic effects caused by a genetic abnormality in a lethal gene, which could include abnormalities and late fetal death in some pregnancies, but pregnancy loss in others.

Bioinformatic filtering is required when studying the whole exome in order to provide a more manageable approach to interpretation of the data. In most of these studies, 'trios' of patients were sequenced, and bioinformatic modelling of inheritance patterns was used to limit the number of variants identified. In most cases, patterns of autosomal recessive inheritance (or X-linked recessive in male fetal losses) were modelled to look for variants. As might be expected, very often the couples investigated were consanguineous or possibly from populations isolated geographically. An alternative autozygosity mapping approach was used by Shamseldin *et al.* (2013) to restrict the genes that were analysed by WES (Shamseldin *et al*., 2013; Shamseldin *et al*., 2015) and a 'proof of principle' study (Ellard *et al*., 2015) developed a technique to identify autosomal recessive lethal disorders using WES in couples with RPL.

It is important to note that where autosomal recessive mutations are identified as a cause of pregnancy loss, this will guide counselling and treatment options for the couple as there is a 1:4 recurrence risk in future pregnancies, and prenatal diagnosis or PGD would be available to the couple.

Interestingly, genes that were identified from these WES studies are associated with processes that have an early role in developmental biology and are essential in embryogenesis. Some key processes include centrosome integrity, anti-inflammatory/immune responses, proliferation and maintenance of epithelial cells, maintenance and development of collagen and muscle tissues, and blood coagulation. The majority of WES studies focused on individual families. Therefore, the genes detected are limited to preselected cases and it is not possible to group them together for a meta-analysis to ascertain the DRs.

Immune cells present early during pregnancy, especially during implantation where the maternal immune system has to tolerate the implanting embryo. The immune response during implantation is not currently well understood. However, the maternal immunity shifts from cell-mediated immunity to humoral (antibody mediated) immunity to protect the embryo from rejection. Aberrations in several genes, *ALOX15* (Qiao *et al*., 2016), *CR1* (Quintero-Ronderos *et al*., 2017), *FOXP3* (Rae *et al*., 2015), and *TLR3* (Filges *et al*., 2014) were identified and are known to be involved in inflammatory and immune defenses. Mutations in these genes could be causing defects resulting in early pregnancy loss because the immune response is rejecting the embryo.

During embryogenesis, cells differentiate and proliferate. Potentially causative mutations were identified in *FLT1* (Quintero-Ronderos *et al*., 2017), *LIFR* (Quintero-Ronderos *et al*., 2017), and *UBN1* (Shamseldin *et al*., 2015) genes involved in cell differentiation and proliferation. Mutations in the two genes *TRO* and *CHD11* were both identified (Quintero-Ronderos *et al*., 2017) and are involved in cell adhesion. As cell differentiation, cell proliferation and cell adhesion are an important part of fetal growth during pregnancy, disruption in these genes could cause the pregnancy to fail.

Mutations in genes involved in tissue formation were also identified. In particular, *CDH1* (Quintero-Ronderos *et al*., 2017) and *FZD6* (Shamseldin *et al*., 2015) are specifically involved in cell adhesion, *MMP10* and *MMP9* (Quintero-Ronderos *et al*., 2017) for extracellular remodelling, and *MuSK* (Wilbe *et al*., 2015) and *MYOM1* (Shamseldin *et al*., 2015) for formation of neuromuscular junctions and striated muscle.

During pregnancy, blood passes through the placenta for the exchange of gases, nutrients, electrolytes, and waste products between the mother and fetus. Mutations in three genes, *F5*, *FGA*, and *THBD* (Quintero-Ronderos *et al*., 2017), were identified. These are involved in the coagulation pathway. The flow of blood is necessary for the fetus to grow and any disruption causing the blood to clot could result in loss of the pregnancy.

In summary, WES of POC or fetal DNA and parental DNA is a promising method to identify variants in genes that might be responsible for RPL and/or fetal abnormalities. Where aberrations are inherited from the parents, a genetic diagnosis may provide invaluable

information for pre-implantation screening or prenatal diagnosis in future pregnancies. However, studies with larger unbiased cohorts are needed to conclusively determine DRs and the clinical utility of WES in this group of patients.

Chromosomal microarray analysis

In some cases, CNVs either as gains or losses may be responsible for pregnancy loss of a fetus with an apparently normal karyotype. CNVs, both rare and common, may be impacting pregnancy-related genes or pathways, resulting in pregnancy loss. These may involve single genes or clusters of genes that are deleted, duplicated, or disrupted.

Studies identified by our systematic review are summarized in Table IV. Due to the diverse approaches taken, the studies are difficult to compare collectively. Cohorts reported sporadic pregnancy loss and RPL, different gestations, and different methods of analysis. Some studies (Warren *et al*., 2009; Rajcan-Separovic *et al*., 2010a; Rajcan-Separovic *et al*., 2010b; Robberecht *et al*., 2012; Levy *et al*., 2014; Bagheri *et al*., 2015; Wang *et al*., 2017) analysed both fetal tissue and parental DNA concurrently (i.e. a trio) to identify whether CNVs were *de novo* or inherited. This is important in assessing both the likely pathogenicity of the finding and the associated recurrence risk. Where the CNV is also detected in a parent, it is less likely to be causative of a pregnancy loss in isolation. It is possible that inherited CNVs could still cause RPL where the CNV co-occurs with an autosomal recessive gene mutation (SNV) on the other allele or where genes present within the CNV are relevant to genomic imprinting or embryonic/placental growth (Rajcan-Separovic *et al*., 2010a, Rajcan-Separovic *et al*., 2010b).

Relatively little is known about the genes and pathways involved in pregnancy loss, and therefore, many CNVs identified will be classed as having uncertain clinical significance. One study analysed CNVs in parents experiencing idiopathic RPL using functional enrichment analysis, identifying biological pathways that were significantly over-represented, such as antigen binding and immune signalling (Nagirnaja *et al*., 2014; Karim *et al*., 2017). Enrichment was identified in genes associated with immunoregulatory interactions at the feto-maternal interface and impaired immune signalling (Nagirnaja *et al*., 2014).

Identification of pregnancies with developmental abnormalities using hystero-embryoscopy enables genetic abnormalities to be compared with developmental abnormalities and growth disorganization of the embryo. CNVs identified where there is a developmental abnormality present are more likely to indicate genes important in early development. In addition to evaluating a genetic cause for pregnancy loss, such studies can provide an opportunity to identify and evaluate the function of the genes. Where variants are identified in genes, through analysis of an enriched cohort with developmental abnormalities, it is easier to interpret their clinical significance.

Several studies explored the possibility of uniparental disomy (UPD) and looked for regions of loss of heterozygosity in euploid embryos (Robberecht *et al*., 2012; Levy *et al*., 2014; Wang *et al*., 2017). The pathological relevance of UPD is difficult to evaluate as not all platforms are capable of detecting UPD (e.g. Oligo BAC array) and therefore are difficult to compare. Pregnancy loss could be due to UPD resulting in unmasking of an underlying lethal recessive disease gene(s) or imprinted genes.

CNVs were identified in the highly imprinted region 11p15.5. This region is abundant with imprinted genes and has an important role in the maternal–fetal exchange. Aberrant methylation or duplication of imprinted genes in this region could cause pregnancy loss (Zhang *et al*., 2016).

Recurrent molar pregnancies

Although the majority of HMs are sporadic, a small minority are recurrent and/or familial. A number of studies looked at the role of genes including *NLRP7*, *C6orf221* (*KHDC3L*), and *NLRP2* in pregnancy loss manifesting as recurrent molar pregnancy. In the cases reviewed, the HMs are euploid and are instead caused by autosomal recessive mutations in genes that code for the cell machinery that labels the parental origin of the two sets of chromosomes.

It is thought that *NLRP7* and *C6orf221* are components of an oocyte complex that forms during oogenesis and determines the epigenetic status of the oocyte genome by inactivating genes. It is likely that mutations in *NLRP7* cause HM by impairing the normal imprinting process causing maternal genes to be expressed when they should not be.

Studies have explored the role of *NLRP2*, *NLRP5*, *NLRP7*, and *C6orf221* in other forms of pregnancy loss such as partial moles, RPL, stillbirth, infertility, and multi-locus imprinting disturbance (Messaed *et al*., 2011; Andreasen *et al*., 2013; Huang *et al*., 2013; Manokhina *et al*., 2013; Aghajanova *et al*., 2015; Docherty *et al*., 2015). These have shown conflicting results, many showing no evidence of *NLRP7*, *NLRP2*, and *C6orf221* mutations in women with RPL (Andreasen *et al*., 2013; Manokhina *et al*., 2013; Aghajanova *et al*., 2015).

Evidence from several papers suggests that genes involved in oocyte development, maturation, and epigenetic reprogramming are likely to be important in a subset of pregnancy losses. One of the most studied epigenetic modifications is DNA methylation. DNA methylation is implicated in the regulation of imprinting and the expression of imprinted genes is thought to be important for the development and physiology of the placenta (Frost and Moore, 2010). Aberrant DNA methylation of several imprinted loci (*H19*, *LIT1*, and *SNRPN*) was demonstrated in pregnancy losses, with increasing methylation of these genes showing a positive correlation with pregnancy loss. It is possible that inappropriate DNA methylation may either be a contributing factor or consequence of the defect that led to pregnancy loss (Zheng *et al*., 2013). It also remains to be investigated as to whether there are wider epigenetic defects at other loci. Zheng *et al*. (2013) propose a multifactorial threshold model for pregnancy loss where additional genetic and environmental factors may also play a role.

Other genetic causes

Mitochondria have been hypothesized to have an important role in development. They predominantly regulate the production of ATP, used to regulate cellular metabolism. Processes such as cell proliferation and development require high energy giving the mitochondria an important role during pregnancy. Seyedhssani *et al*. (2010a,b) have identified mutations in mtDNA in women with RPL (Seyedhassani *et al*., 2010b). Furthermore, a significant number of mutations were identified in the D-loop of mtDNA. The D-loop contains essential elements for mtDNA transcription and disruption could

affect the transcription or translation of mtDNA, in turn compromising embryonic development or causing pregnancy loss.

It is hypothesized that skewed XCI could be involved in the pathogenesis of RPL. Bagislar *et al.* (2006) demonstrated extremely skewed XCI in 17.7 % of patients with RPL. It is suggested that skewed XCI could expose X-linked variants that are lethal in the hemizygous state. In addition, a more recent review (Sui *et al.*, 2015) included 12 case-control studies on skewed XCI with or without RPL. In patients with RPL, skewed XCI was significantly higher, although the significance drops with fewer losses and for less extreme skewing. Although the association between RPL and skewed XCI is unclear, two mechanisms have been proposed. Firstly, if a female carrier with a recessive lethal X-linked genetic mutation and skewed XCI has a male fetus who inherits the X-linked genetic mutation, it could lead to pregnancy loss. Secondly, an X-linked genetic mutation could cause follicular atresia and an increase in aneuploid embryos resulting in pregnancy loss (Sui *et al.*, 2015).

Six papers (Bhuiyan *et al.*, 2008; Stouffs *et al.*, 2011; Lopez-Carrasco *et al.*, 2013; McKie *et al.*, 2014; Bendroth-Asmussen *et al.*, 2016; Zhang *et al.*, 2016) describe targeted sequence analysis of specific candidate genes (*GBE1*, *RYR1*, *WNT6*, *DNMT3L*, SYCP3, *MSH4, HERG*, and AURKB) in either an individual case of pregnancy loss (Bhuiyan *et al.*, 2008; Bendroth-Asmussen *et al.*, 2016) or in patient cohorts (Stouffs *et al.*, 2011; Lopez-Carrasco *et al.*, 2013; McKie *et al.*, 2014; Zhang *et al.*, 2016). This targeting was informed by factors including histopathological examination of placental tissue observed in fetal arrhythmia, scan findings, and functional prediction of gene pathways.

Limitations of current evidence

This review was completed in a systematic manner by two independent reviewers making it reproducible. The limitation of this study, however, is the quality of the studies published to date. Each study was scored according to our modified Newcastle–Ottawa scale (Supplementary Table SIV) with a few of the studies being of poor quality and scoring as little as 3 or 4 on our scale.

The most common limitations in these studies related to the small size of the studied cohorts, with several focusing on a single family, and many of the studies lacking information on control populations or statistical analysis. Work on small groups, and in particular a single family, may detect genetic abnormalities that have occurred in isolation or are very rare. In many cases, this results in identification of variants in unique candidate genes with no definitive causal effect. Therefore, larger cohorts are needed to replicate these findings and to determine how relevant these findings are to other couples with RPL.

There was also limited availability of functional data in many of the studies. A few studies supplemented their cases with information on scan abnormalities or post-mortem abnormalities detected in cases of losses and hystero-embryoscopy to correlate genetic findings with findings in the embryo. The studies were also difficult to compare and collate as there were multiple variations in the cohorts studied and the methods of analysis.

Conclusion

It is evident that there are many genetic and environmental factors that result in a successful pregnancy and a disruption in any of these could contribute to pregnancy loss.

From the genetic perspective this includes both clearly pathogenic genetic causes, such as sporadic aneuploidy and translocations, and other potential genetic causes such as smaller CNVs and mutations in genes important in early fetal development. In addition, there are likely to be complex genetic contributions, such as multi-factorial inheritance, and changes in methylation (epigenetics) and mitochondrial function, which could be contributing to pregnancy loss. These more complex genetic mechanisms may be influenced by environmental factors, such as diet, medication, pollutants, and lifestyle, which could provide a cumulative effect resulting in pregnancy loss.

The papers we have identified have demonstrated that monogenic aetiologies could contribute to a proportion of pregnancy losses. However, as most studies have been carried out in highly selected families or small cohorts, additional studies are required to further assess if this technology is generalizable to more couples experiencing RPL.

It is plausible that cases of pregnancy loss (particularly in RPL) may have causative mutations not detectable with routine cytogenetic analysis or fetal scans, but are detectable by WES. Although WES is not currently recommended for routine diagnostic use for pregnancy losses, the identification of genes associated with pregnancy loss will be of significant individual patient impact with respect to treatment and availability of PGD. If monogenetic etiologies of RPL and the overall prevalence of monogenetic causes of pregnancy loss are better elucidated through larger, well-designed studies, the identification of non-aneuploid causes of RPL could be of significant patient impact.

Knowledge of specific genes that contribute to pregnancy loss could also be of importance in understanding the biological pathways that can cause pregnancy loss. However, much larger and more comparable cohort studies are required in all of these areas to determine causality of candidate genes and to dissect out these effects, as at present many of these findings are of uncertain clinical significance. Functional analysis, such as embryoscopy studies and *in vivo* animal modelling, may assist in further assessment of the mutation effect on early embryonic development.

RPL is a complex problem influenced by many different aetiologies. Currently, with the exception of aneuploidy and other chromosomal abnormalities, routine investigation for the genetic contributions causing pregnancy loss is limited. With increased knowledge of additional non-aneuploid contributions to RPL, additional genetic testing recommendations may be made in the future to couples experiencing RPL. These would have implications for diagnosis and recurrence risks.

Supplementary data

Supplementary data are available at *Human Reproduction Update* online.

Authors' roles

E.C.- study search, study selection, data extraction, quality assessment, and writing. S.H.- data extraction, quality assessment, and editing. P.S.-

study design and critical appraisal of manuscript. N.M.- critical appraisal of manuscript and editing. A.C.- study design and critical appraisal of manuscript. S.A.- supervision, study selection, writing, and editing.

Funding

Tommy's Baby Charity, registered charity (1060508) to E.C.

Conflict of interest

There are no conflicts of interest to declare.

References

Abdalla EM, Hayward BE, Shamseddin A, Nawar MM. Recurrent hydatidiform mole: detection of two novel mutations in the NLRP7 gene in two Egyptian families. *Eur J Obstet Gynecol Reprod Biol* 2012;**164**:211–215.

Aghajanova L, Mahadevan S, Altmae S, Stavreus-Evers A, Regan L, Sebire N, Dixon P, Fisher RA, Van den Veyver IB. No evidence for mutations in NLRP7, NLRP2 or KHDC3L in women with unexplained recurrent pregnancy loss or infertility. *Hum Reprod* 2015;**30**:232–238.

Andreasen L, Christiansen OB, Niemann I, Bolund L, Sunde L. NLRP7 or KHDC3L genes and the etiology of molar pregnancies and recurrent miscarriage. *Mol Hum Reprod* 2013;**19**:773–781.

Bagheri H, Mercier E, Qiao Y, Stephenson MD, Rajcan-Separovic E. Genomic characteristics of miscarriage copy number variants. *Mol Hum Reprod* 2015;**21**:655–661.

Bagislar S, Ustuner I, Cengiz B, Soylemez F, Akyerli CB, Ceylaner S, Ceylaner G, Acar A, Ozcelik T. Extremely skewed X-chromosome inactivation patterns in women with recurrent spontaneous abortion. *Aust N Z J Obstet Gynaecol* 2006;**46**:384–387.

Bender Atik R, Christiansen OB, Elson J, Kolte AM, Lewis S, Middeldorp S, Nelen W, Peramo B, Quenby S, Vermeulen N, Goddijn M. ESHRE guideline: recurrent pregnancy loss. *Hum Reprod Open* 2018;**2**:2018.

Bendroth-Asmussen L, Aksglaede L, Gernow AB, Lund AM. Glycogen storage disease type IV: a case with histopathologic findings in first-trimester placental tissue. *Int J Gynecol Pathol* 2016;**35**:38–40.

Bhuiyan ZA, Momenah TS, Gong Q, Amin AS, Ghamdi SA, Carvalho JS, Homfray T, Mannens MM, Zhou Z, Wilde AA. Recurrent intrauterine fetal loss due to near absence of HERG: clinical and functional characterization of a homozygous nonsense HERG Q1070X mutation. *Heart Rhythm* 2008;**5**:553–561.

Bondeson ML, Ericson K, Gudmundsson S, Ameur A, Ponten F, Wesstrom J, Frykholm C, Wilbe M. A nonsense mutation in CEP55 defines a new locus for a Meckel-like syndrome, an autosomal recessive lethal fetal ciliopathy. *Clin Genet* 2017;**92**:510–516.

Brown L, Mount S, Reddy R, Slim R, Wong C, Jobanputra V, Clifford P, Merrill L, Brown S. Recurrent pregnancy loss in a woman with NLRP7 mutation: not all molar pregnancies can be easily classified as either 'partial' or 'complete' hydatidiform moles. *Int J Gynecol Pathol* 2013;**32**:399–405.

Carey L, Nash BM, Wright DC. Molecular genetic studies of complete hydatidiform moles. *Transl Pediatr* 2015;**4**:181–188.

Carss KJ, Hillman SC, Parthiban V, McMullan DJ, Maher ER, Kilby MD, Hurles ME. Exome sequencing improves genetic diagnosis of structural fetal abnormalities revealed by ultrasound. *Hum Mol Genet* 2014;**23**:3269–3277.

Choi M, Scholl UI, Ji W, Liu T, Tikhonova IR, Zumbo P, Nayir A, Bakkaloglu A, Ozen S, Sanjad S *et al.* Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci U S A* 2009;**106**:19096–19101.

Cristofoli F, De Keersmaecker B, De Catte L, Vermeesch JR, Van Esch H. Novel STIL compound heterozygous mutations cause severe fetal microcephaly and Centriolar lengthening. *Mol Syndromol* 2017;**8**:282–293.

Deveault C, Qian JH, Chebaro W, Ao A, Gilbert L, Mehio A, Khan R, Tan SL, Wischmeijer A, Coullin P *et al.* NLRP7 mutations in women with diploid androgenetic and triploid moles: a proposed mechanism for mole formation. *Hum Mol Genet* 2009;**18**:888–897.

Docherty LE, Rezwan FI, Poole RL, Turner CL, Kivuva E, Maher ER, Smithson SF, Hamilton-Shield JP, Patalan M, Gizewska M *et al.* Mutations in NLRP5 are associated with reproductive wastage and multilocus imprinting disorders in humans. *Nat Commun* 2015;**6**:8086.

Dohrn N, Le VQ, Petersen A, Skovbo P, Pedersen IS, Ernst A, Krarup H, Petersen MB. ECEL1 mutation causes fetal arthrogryposis multiplex congenita. *Am J Med Genet A* 2015;**167a**:731–743.

Donaghue C, Davies N, Ahn JW, Thomas H, Ogilvie CM, Mann K. Efficient and cost-effective genetic analysis of products of conception and fetal tissues using a QF-PCR/array CGH strategy; five years of data. *Mol Cytogenet* 2017;**10**:12.

Ellard S, Kivuva E, Turnpenny P, Stals K, Johnson M, Xie W, Caswell R, Lango Allen H. An exome sequencing strategy to diagnose lethal autosomal recessive disorders. *Eur J Hum Genet* 2015;**23**:401–404.

Filges I, Nosova E, Bruder E, Tercanli S, Townsend K, Gibson WT, Rothlisberger B, Heinimann K, Hall JG, Gregory-Evans CY *et al.* Exome sequencing identifies mutations in KIF14 as a novel cause of an autosomal recessive lethal fetal ciliopathy phenotype. *Clin Genet* 2014;**86**:220–228.

Franssen MT, Musters AM, van der Veen F, Repping S, Leschot NJ, Bossuyt PM, Goddijn M, Korevaar JC. Reproductive outcome after PGD in couples with recurrent miscarriage carrying a structural chromosome abnormality: a systematic review. *Hum Reprod Update* 2011;**17**:467–475.

Frost JM, Moore GE. The importance of imprinting in the human placenta. *PLoS Genet* 2010;**6**:e1001015.

Goddijn M, Leschot NJ. Genetic aspects of miscarriage. *Baillieres Best Pract Res Clin Obstet Gynaecol* 2000;**14**:855–865.

Hassold T, Chen N, Funkhouser J, Jooss T, Manuel B, Matsuura J, Matsuyama A, Wilson C, Yamane JA, Jacobs PA. A cytogenetic study of 1000 spontaneous abortions. *Ann Hum Genet* 1980;**44**:151–178.

Hirshfeld-Cytron J, Sugiura-Ogasawara M, Stephenson MD. Management of recurrent pregnancy loss associated with a parental carrier of a reciprocal translocation: a systematic review. *Semin Reprod Med* 2011;**29**:470–481.

Huang JY, Su M, Lin SH, Kuo PL. A genetic association study of NLRP2 and NLRP7 genes in idiopathic recurrent miscarriage. *Hum Reprod* 2013;**28**:1127–1134.

Karim S, Jamal HS, Rouzi A, Ardawi MSM, Schulten HJ, Mirza Z, Alansari NA, Al-Quaiti MM, Abusamra H, Naseer MI *et al.* Genomic answers for recurrent spontaneous abortion in Saudi Arabia: an

dd

array comparative genomic hybridization approach. *Reprod Biol* 2017;**17**:133–143.

Larsen EC, Christiansen OB, Kolte AM, Macklon N. New insights into mechanisms behind miscarriage. *BMC Med* 2013;**11**:154.

Levy B, Sigurjonsson S, Pettersen B, Maisenbacher MK, Hall MP, Demko Z, Lathi RB, Tao R, Aggarwal V, Rabinowitz M. Genomic imbalance in products of conception: single-nucleotide polymorphism chromosomal microarray analysis. *Obstet Gynecol* 2014;**124**:202–209.

Lopez-Carrasco A, Oltra S, Monfort S, Mayo S, Rosello M, Martinez F, Orellana C. Mutation screening of AURKB and SYCP3 in patients with reproductive problems. *Mol Hum Reprod* 2013;**19**:102–108.

Maisenbacher MK, Merrion K, Pettersen B, Young M, Paik K, Iyengar S, Kareht S, Sigurjonsson S, Demko ZP, Martin KA. Incidence of the 22q11.2 deletion in a large cohort of miscarriage samples. *Mol Cytogenet* 2017;**10**:6.

Manokhina I, Hanna CW, Stephenson MD, McFadden DE, Robinson WP. Maternal NLRP7 and C6orf221 variants are not a common risk factor for androgenetic moles, triploidy and recurrent miscarriage. *Mol Hum Reprod* 2013;**19**:539–544.

McKie AB, Alsaedi A, Vogt J, Stuurman KE, Weiss MM, Shakeel H, Tee L, Morgan NV, Nikkels PG, van Haaften G *et al.* Germline mutations in RYR1 are associated with foetal akinesia deformation sequence/lethal multiple pterygium syndrome. *Acta Neuropathol Commun* 2014;**2**:148.

Messaed C, Chebaro W, Di Roberto RB, Rittore C, Cheung A, Arseneau J, Schneider A, Chen MF, Bernishke K, Surti U *et al.* NLRP7 in the spectrum of reproductive wastage: rare non-synonymous variants confer genetic susceptibility to recurrent reproductive wastage. *J Med Genet* 2011;**48**:540–548.

Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, Church DM, Crolla JA, Eichler EE, Epstein CJ *et al.* Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet* 2010;**86**:749–764.

Nagirnaja L, Palta P, Kasak L, Rull K, Christiansen OB, Nielsen HS, Steffensen R, Esko T, Remm M, Laan M. Structural genomic variation as risk factor for idiopathic recurrent miscarriage. *Hum Mutat* 2014;**35**:972–982.

Parry DA, Logan CV, Hayward BE, Shires M, Landolsi H, Diggle C, Carr I, Rittore C, Touitou I, Philibert L *et al.* Mutations causing familial biparental hydatidiform mole implicate c6orf221 as a possible regulator of genomic imprinting in the human oocyte. *Am J Hum Genet* 2011;**89**:451–458.

Pereza N, Ostojic S, Kapovic M, Peterlin B. Systematic review and meta-analysis of genetic association studies in idiopathic recurrent spontaneous abortion. *Fertil Steril* 2017;**107**:150–159.

Qian J, Cheng Q, Murdoch S, Xu C, Jin F, Chebaro W, Zhang X, Gao H, Zhu Y, Slim R *et al.* The genetics of recurrent hydatidiform moles in China: correlations between NLRP7 mutations, molar genotypes and reproductive outcomes. *Mol Hum Reprod* 2011;**17**:612–619.

Qiao Y, Wen J, Tang F, Martell S, Shomer N, Leung PCK, Stephenson MD, Rajcan-Separovic E. Whole exome sequencing in recurrent early pregnancy loss. *Mol Hum Reprod* 2016;**22**:364–372.

Quintero-Ronderos P, Mercier E, Fukuda M, Gonzalez R, Suarez CF, Patarroyo MA, Vaiman D, Gris JC, Laissue P. Novel genes and mutations in patients affected by recurrent pregnancy loss. *PLoS One* 2017;**12**:e0186149.

Rae W, Gao Y, Bunyan D, Holden S, Gilmour K, Patel S, Wellesley D, Williams A. A novel FOXP3 mutation causing fetal akinesia and recurrent male miscarriages. *Clin Immunol* 2015;**161**:284–285.

Rajcan-Separovic E, Diego-Alvarez D, Robinson WP, Tyson C, Qiao Y, Harvard C, Fawcett C, Kalousek D, Philipp T, Somerville MJ *et al.* Identification of copy number variants in miscarriages from couples with idiopathic recurrent pregnancy loss. *Hum Reprod* 2010a;**25**:2913–2922.

Rajcan-Separovic E, Qiao Y, Tyson C, Harvard C, Fawcett C, Kalousek D, Stephenson M, Philipp T. Genomic changes detected by array CGH in human embryos with developmental defects. *Mol Hum Reprod* 2010b;**16**:125–134.

Robberecht C, Pexsters A, Deprest J, Fryns JP, D'Hooghe T, Vermeesch JR. Cytogenetic and morphological analysis of early products of conception following hystero-embryoscopy from couples with recurrent pregnancy loss. *Prenat Diagn* 2012;**32**:933–942.

Royal College of Obstetricians and Gynaecologists. The investigation and treatment of couples with recurrent first trimester and second-trimester miscarriage. Green Top Guideline No. 17. London: RCOG, 2011.

Seyedhassani SM, Houshmand M, Kalantar SM, Aflatoonian A, Modabber G, Hadipour F, Fallahzadeh MH. The point mutations of mitochondrial tRNA threonine and proline in idiopathic repeated pregnancy loss. *Iran J Reprod Med* 2010a;**8**:45–50.

Seyedhassani SM, Houshmand M, Kalantar SM, Modabber G, Aflatoonian A. No mitochondrial DNA deletions but more D-loop point mutations in repeated pregnancy loss. *J Assist Reprod Genet* 2010b;**27**:641–648.

Shamseldin HE, Kurdi W, Almusafri F, Alnemer M, Alkaff A, Babay Z, Alhashem A, Tulbah M, Alsahan N, Khan R *et al.* Molecular autopsy in maternal-fetal medicine. *Genet Med* 2018;**20**:420–427.

Shamseldin HE, Swaid A, Alkuraya FS. Lifting the lid on unborn lethal Mendelian phenotypes through exome sequencing. *Genet Med* 2013;**15**:307–309.

Shamseldin HE, Tulbah M, Kurdi W, Nemer M, Alsahan N, Al Mardawi E, Khalifa O, Hashem A, Kurdi A, Babay Z *et al.* Identification of embryonic lethal genes in humans by autozygosity mapping and exome sequencing in consanguineous families. *Genome Biol* 2015;**16**:116.

Stouffs K, Vandermaelen D, Tournaye H, Liebaers I, Lissens W. Mutation analysis of three genes in patients with maturation arrest of spermatogenesis and couples with recurrent miscarriages. *Reprod Biomed Online* 2011;**22**:65–71.

Sui Y, Chen Q, Sun X. Association of skewed X chromosome inactivation and idiopathic recurrent spontaneous abortion: a systematic review and meta-analysis. *Reprod Biomed Online* 2015;**31**:140–148.

Suzuki T, Behnam M, Ronasian F, Salehi M, Shiina M, Koshimizu E, Fujita A, Sekiguchi F, Miyatake S, Mizuguchi T *et al.* A homozygous NOP14 variant is likely to cause recurrent pregnancy loss. *J Hum Genet* 2018;**63**:425–430.

Tsurusaki Y, Yonezawa R, Furuya M, Nishimura G, Pooh RK, Nakashima M, Saitsu H, Miyake N, Saito S, Matsumoto N. Whole exome sequencing revealed biallelic IFT122 mutations in a family with CED1 and recurrent pregnancy loss. *Clin Genet* 2014;**85**:592–594.

Ulker V, Gurkan H, Tozkir H, Karaman V, Ozgur H, Numanoglu C, Gedikbasi A, Akbayir O, Uyguner ZO. Novel NLRP7 mutations in familial recurrent hydatidiform mole: are NLRP7 mutations a risk for recurrent reproductive wastage? *Eur J Obstet Gynecol Reprod Biol* 2013;**170**:188–192.

Viaggi CD, Cavani S, Malacarne M, Floriddia F, Zerega G, Baldo C, Mogni M, Castagnetta M, Piombo G, Coviello DA *et al.* First-trimester euploid miscarriages analysed by array-CGH. *J Appl Genet* 2013;**54**:353–359.

Wang Y, Cheng Q, Meng L, Luo C, Hu H, Zhang J, Cheng J, Xu T, Jiang T, Liang D *et al.* Clinical application of SNP array analysis in first-trimester pregnancy loss: a prospective study. *Clin Genet* 2017;**91**:849–858.

Warren JE, Turok DK, Maxwell TM, Brothman AR, Silver RM. Array comparative genomic hybridization for genetic evaluation of fetal loss between 10 and 20 weeks of gestation. *Obstet Gynecol* 2009;**114**:1093–1102.

Wilbe M, Ekvall S, Eurenius K, Ericson K, Casar-Borota O, Klar J, Dahl N, Ameur A, Anneren G, Bondeson ML. MuSK: a new target for lethal fetal akinesia deformation sequence (FADS). *J Med Genet* 2015;**52**:195–202.

Wilson R, Geyer SH, Reissig L, Rose J, Szumska D, Hardman E, Prin F, McGuire C, Ramirez-Solis R, White J *et al.* Highly variable penetrance of abnormal phenotypes in embryonic lethal knockout mice. *Wellcome Open Res* 2016;**1**:1.

Wright CF, Fitzgerald TW, Jones WD, Clayton S, McRae JF, van Kogelenberg M, King DA, Ambridge K, Barrett DM, Bayzetinova T *et al.* Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet* 2015;**385**:1305–1314.

Zhang H, Liu W, Chen M, Li Z, Sun X, Wang C. Implementation of a high-resolution single-nucleotide polymorphism array in analyzing the products of conception. *Genet Test Mol Biomarkers* 2016;**20**:352–358.

Zhang Y, Li G, Fan Y, Cui Y, Huang S, Ma J, Yan J, Chen ZJ. Novel missense mutation in WNT6 in 100 couples with unexplained recurrent miscarriage. *Hum Reprod* 2015;**30**:994–999.

Zhang YX, Zhang YP, Gu Y, Guan FJ, Li SL, Xie JS, Shen Y, Wu BL, Ju W, Jenkins EC *et al.* Genetic analysis of first-trimester miscarriages with a combination of cytogenetic karyotyping, microsatellite genotyping and arrayCGH. *Clin Genet* 2009;**75**:133–140.

Zheng HY, Tang Y, Niu J, Li P, Ye DS, Chen X, Shi XY, Li L, Chen SL. Aberrant DNA methylation of imprinted loci in human spontaneous abortions after assisted reproduction techniques and natural conception. *Hum Reprod* 2013;**28**:265–273.

Zhou Q, Wu SY, Amato K, DiAdamo A, Li P. Spectrum of cytogenomic abnormalities revealed by array comparative genomic hybridization on products of conception culture failure and normal karyotype samples. *J Genet Genomics* 2016;**43**:121–131.

Supplementary Table 1 Systematic review- Search terms with corresponding Mesh terms

	OVID Medline			**Embase**	

OR	• Abortion, Spontaneous
	  • Abortion habitual
	  • Abortion incomplete
	  • Abortion missed
	  • Abortion septic
	  • Abortion threatened
	  • Abortion veterinary
	  • Embryo loss
	• Pregnancy loss (free text)
	• Miscarriage (free text)
	• Lethal (key word)
		Limited to human	OR	• Spontaneous abortion
				  ➢ Abortion spontaneous
				  ➢ Miscarriage
				• Abortion
				  ➢ Recurrent abortion
				  ➢ Spontaneous abortion
				  ➢ Fetus wastage
				  ➢ Septic abortion
				• Recurrent abortion
				  ➢ Abortion habitual
				  ➢ Abortion recurrent
				  ➢ Habitual abortion
				  ➢ Repeated abortion
				  ➢ Successive abortion
				• Pregnancy loss (key word)
				• Miscarriage (keyword)
				• Lethal (key word
					Limited to human

	AND			AND	

OR	• Exome (focus and keyword)
	• Genome, Human
	• Sequencing adj (massively parallel or next generation or exome or genome)
	• Sequencing analysis
	  ➢ High throughput nucleotide sequencing
	  ➢ Molecular annotation
	  ➢ Oligonucleotide array sequence analysis
	  ➢ GWAS
	  ➢ Positon-specific scoring matrics
	  ➢ Sequence analysis, DNA
	  ➢ DNA barcoding, Taxonomic
	  ➢ DNA contamination
	  ➢ DNA mutational analysis
	  ➢ Multiocus sequence typing
			OR	• Exome (exp and keyword)
				• Human genome
				• Genome (keyword)
				• Sequencing adj (massively parallel or next generation or exome or genome)
				• High throughput nucleotide sequencing
				• Sequence analysis
				• Copy number Variation
				  ➢ CNV (copy number variation)
				  ➢ Copy number variations
				  ➢ DNA copy number variation
				  ➢ DNA copy number variations
				• Single nucleotide polymorphism
				  ➢ Polymorphism, single nucleotide
				  ➢ Single nucleotide variant	

gg

	➢ Sequence analysis ➢ Peptide mapping ➢ Sequence analysis, RNA • DNA copy number variations • Polymorphism, Genetic ➢ Genomic structural variation ➢ Pharmacogenomics variants ➢ Polymorphism, restriction fragment length ➢ Polymorphism, single nucleotide ➢ Polymorphism, single-stranded conformational • Microarray analysis ➢ Oligonucleotide array sequence analysis ➢ Protein array analysis ➢ Tissue array analysis • Comparative Genmoic Hybridization			➢ Single nucleotide variation ➢ DNA polymorphism • Microarray analysis • Comparative genomic hybridization ➢ CGH (comparative genomic hybridization) ➢ Comparative genome hybridisation ➢ Comparative genome hybridization ➢ Comparative genomic hybridisation ➢ Genomic hybridization, comparative ➢ Hybridization, comparative genomic	

hh

Supplementary Table 2- Systematic review- Full inclusion and exclusion criteria

Include	Include **human studies**
	Include **recurrent miscarriage studies/ pregnancy loss**, up to 20 weeks
	Include **genetic sequencing**
	Include **genes linked to miscarriage**
	Include any combination of **male, female or embryo genetics**
	Include **mitochondrial DNA**
	Include **molar pregnancies** where there is a heritable basis eg. Hydatidiform moles
	Include **meiotic disorders** where there is a heritable basis
Exclude	Exclude **non-human**
	Exclude **non-lethal genes** in utero
	Exclude **conference abstracts**
	Exclude **methodologies** where there is no clinical relevance / genetics discussed.
	Exclude **expression genetics** eg. RNA, protein, miRNAs, lncRNA
	Exclude **preimplantation genetics diagnosis** (PGD), **IVF** and **ICSI**
	Exclude **post birth** studies
	Exclude **infertility**
	Exclude **Associations** and **Complex Genetics**
	Exclude **Aneuploidy** where there is no heritability

Supplementary Table 3- Systematic review- Modified Newcastle-Ottawa scale

Criteria	Score
Sample Size	
>100	2
10-100	1
1-10	0
Inclusion-exclusion	
Described in full	2
Limited Description	1
Not included	0
Genotype Method	
Sequencing	2
Other	1
Not mentioned	0
Statistical Analysis	
Well described 'gold standard' statistical analysis	2
Poorly described statistics	1
No Statistics Mentioned	0
Case Definition	
Independent Validation	1
Self-reported	0
Controls	
Yes	1
No	0
Comparability	
Study controls for same gene	2
Study control for general screen	1
Not applicable/ no comparison	0

Supplementary Table 4- Systematic review - Individual scores for Newcastle-Ottawa scale.

Paper	Sample Size (2)	Inclusion-Exclusion (2)	Genotype Method (2)	Statistical analysis (2)	Case Definition (1)	Controls (1)	Comparability (2)	Total (12)
Abdalla et al., 2012	0	1	2	0	1	0	0	4
Aghajanova et al., 2015	1	2	2	0	1	0	0	6
Andreasen et al., 2013	1	2	2	1	1	0	0	7
Bagheri et al., 2015	2	1	1	2	1	0	0	7
Bagislar et al., 2006	1	2	1	2	1	1	0	8
Bendroth-Asmussen et al., 2015	0	0	1	0	1	0	0	2
Bhuiyan et al., 2008	0	0	2	0	1	0	0	3
Bondeson et al., 2017	0	2	2	0	1	0	0	5
Brown et al., 2013	0	2	1	0	1	0	0	4
Cristofoli et al., 2017	0	2	2	1	1	0	0	6
Deveault et al., 2009	1	2	2	0	1	0	0	6
Docherty et al., 2015	1	2	2	0	1	0	0	6
Dohrn et al., 2015	0	1	1	0	1	0	0	3
Donaghue et al., 2017	2	1	1	0	1	0	0	5
Ellard et al., 2015	0	2	2	0	1	1	2	8
Filges et al., 2014	0	1	2	0	1	0	0	4
Huang et al., 2013	2	2	1	2	1	1	1	10
Karim et al., 2017	1	2	1	2	1	0	0	7
Levy et al., 2014	2	1	1	1	1	0	0	6
Lopez-Carrasco et al., 2013	2	2	2	0	1	1	2	10
Maisenbacher et al., 2017	2	1	1	0	1	0	1	6
Manokhina et al., 2013	1	1	2	2	1	1	0	8
McKie et al., 2014	1	2	1	0	1	0	0	5

kk

Messaed et al., 2011	2	1	2	1	1	1	1	9
Nagirnaja et al., 2014	2	1	1	0	1	1	2	8
Parry et al., 2011	0	1	2	0	1	0	0	4
Qian et al., 2011	1	2	2	0	1	1	2	9
Qiao et al., 2016	0	2	2	0	1	0	0	5
Quintero-Renderos e al., 2017	1	2	2	1	1	0	0	7
Rae et al., 2015	0	0	2	0	1	0	0	3
Rajcan-Separovic et al., 2010 a	1	2	1	0	1	0	0	5
Rajcan-Separovic et al., 2010 b	1	2	1	0	0	1	2	7
Robberecht et al., 2012	1	1	1	1	1	0	0	5
Seyedhassani et al., 2010a	1	2	1	0	1	0	0	5
Seyedhassani et al., 2010b	1	2	1	2	1	1	2	10
Shamseldin et al., 2015	1	1	2	0	1	0	0	5
Shamseldin et al.,2013	0	1	2	0	1	1	1	6
Stouffs et at., 2011	1	1	1	0	1	1	2	7
Suzuki et al., 2018	0	0	2	0	1	0	0	3
Tsurusaki et al., 2014	0	0	2	0	1	0	0	3
Ulker et al., 2013	0	0	2	0	1	0	0	3
Viaggi et al., 2013	1	1	1	1	1	0	0	5
Wang et al., 2017	2	2	1	2	1	0	1	9
Warren et al., 2009	1	1	1	0	1	0	0	4
Wilbe et al., 2015	0	1	2	0	1	1	2	7
Zhang et al., 2009	1	1	1	0	1	0	0	4
Zhang et al., 2015	1	2	1	0	1	1	2	8
Zhang et al., 2016	1	2	1	1	1	0	0	6
Zheng et al., 2013	2	2	1	2	1	1	2	11
Zhou et al., 2016	2	2	1	0	1	1	1	9

Appendix 3- Genes in the FETALCES-001: Exome Comparison gene panel

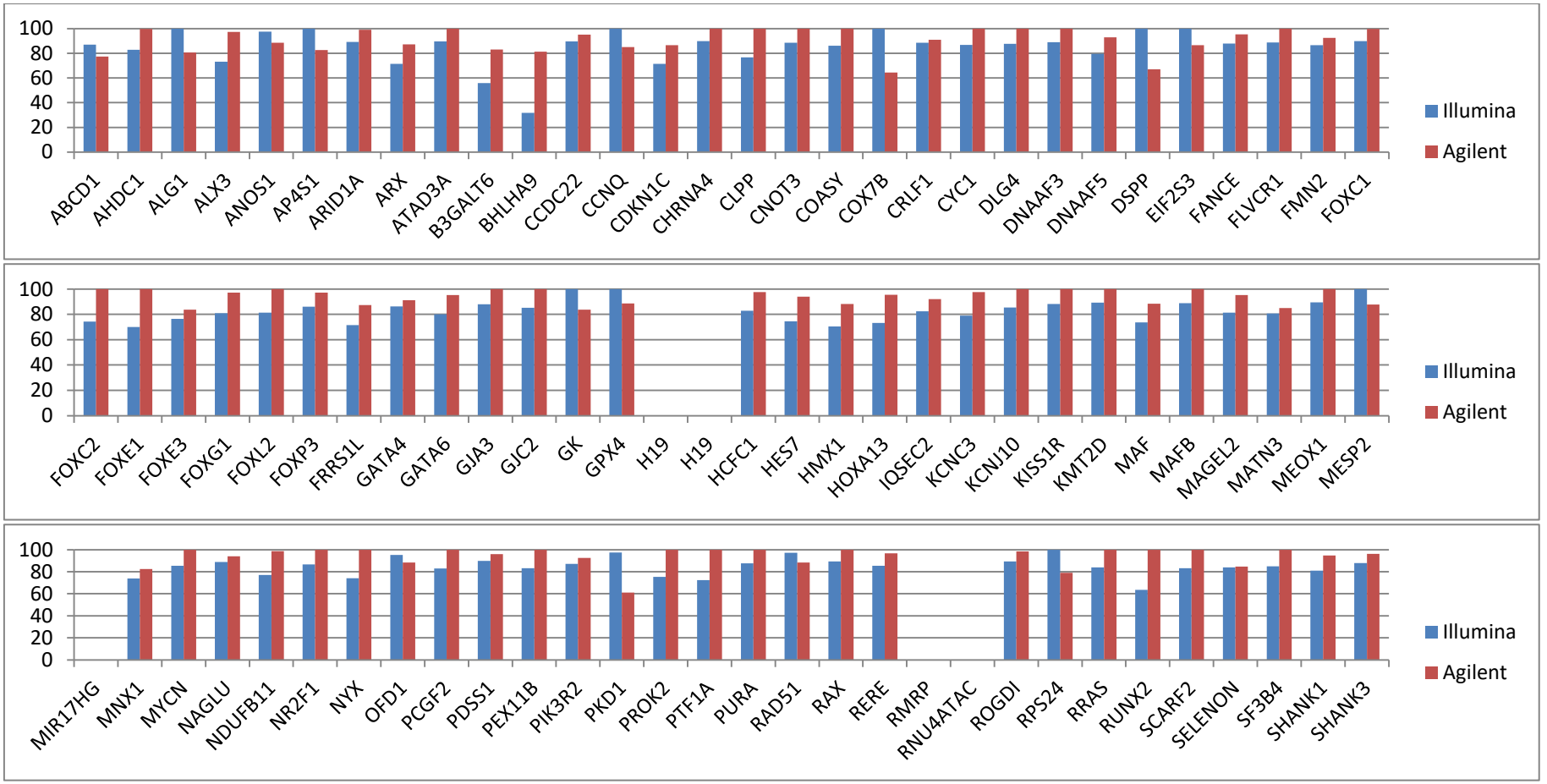
AAAS	AIRE	ARL6	BLOC1S6	CDKN1C	COG7	CTDP1	DMPK	EPHB4
AARS	AK2	ARMC4	BMP1	CDON	COG8	CTNNA2	DNAAF1	EPHX1
AASS	AKR1D1	ARSA	BMP2	CDT1	COL10A1	CTNNB1	DNAAF3	ERCC1
ABCA12	AKT1	ARSB	BMP4	CENPJ	COL11A1	CTNND1	DNAAF4	ERCC2
ABCB11	AKT3	ARSE	BMPER	CEP135	COL11A2	CTNS	DNAAF5	ERCC3
ABCB7	ALAD	ARX	BMPR1B	CEP152	COL18A1	CTSA	DNAH11	ERCC4
ABCC6	ALDH18A1	ASAH1	BOLA3	CEP164	COL1A1	CTSD	DNAH5	ERCC5
ABCC8	ALDH1A3	ASL	BRAF	CEP290	COL1A2	CTSK	DNAI1	ERCC6
ABCC9	ALDH3A2	ASNS	BRAT1	CEP41	COL25A1	CUL4B	DNAJC19	ERCC6L2
ABCD1	ALDH4A1	ASPA	BRCA1	CEP57	COL2A1	CUL7	DNM1	ERCC8
ABCD4	ALDH5A1	ASPH	BRCA2	CEP63	COL4A1	CUX2	DNMT3A	ERF
ABHD5	ALDH7A1	ASPM	BRIP1	CEP83	COL4A2	CYB5R3	DNMT3B	ESCO2
ABL1	ALDOA	ASS1	BRWD3	CFC1	COL4A3	CYC1	DOCK6	ETFA
ACAD9	ALDOB	ASXL1	BSND	CFL2	COL4A3BP	CYP11A1	DOCK7	ETFB
ACADM	ALG1	ASXL3	BTD	CFTR	COL4A4	CYP11B1	DOCK8	ETFDH
ACADS	ALG11	ATAD3A	BUB1B	CHAMP1	COL5A1	CYP17A1	DOLK	ETHE1
ACADVL	ALG12	ATIC	C12orf57	CHAT	COL5A2	CYP19A1	DPAGT1	EVC
ACAN	ALG13	ATM	C12orf65	CHD2	COL6A1	CYP1B1	DPM1	EVC2
ACAT1	ALG2	ATP13A2	C2CD3	CHD3	COL6A2	CYP21A2	DPM3	EXOSC3
ACE	ALG3	ATP1A3	C8orf37	CHD4	COL6A3	CYP2U1	DRC1	EXPH5
ACO2	ALG6	ATP6V0A2	CA2	CHD7	COL9A1	DAG1	DSG1	EXT1
ACOX1	ALG8	ATP6V1B1	CA5A	CHD8	COL9A2	DARS	DSP	EXT2
ACP5	ALG9	ATP6V1B2	CA8	CHKB	COL9A3	DARS2	DSPP	EYA1
ACSL4	ALMS1	ATP7A	CACNA1A	CHMP1A	COLEC11	DBT	DSTYK	EZH2
ACTA1	ALPL	ATP8B1	CACNA1C	CHRDL1	COMP	DCC	DUSP6	FAH
ACTA2	ALS2	ATR	CACNA1D	CHRNA1	COQ2	DCDC2	DVL1	FAM111A
ACTB	ALX1	ATRX	CAMTA1	CHRNA4	COQ4	DCHS1	DYM	FAM126A
ACTC1	ALX3	AUH	CARS2	CHRNB2	COQ8A	DCX	DYNC1H1	FAM161A
ACTG1	ALX4	AUTS2	CASK	CHRND	COQ9	DDB2	DYNC2H1	FAM20A
ACTG2	AMER1	B3GALNT2	CAVIN1	CHRNG	COX10	DDC	DYRK1A	FAM20C
ACVR1	AMPD2	B3GALT6	CBL	CHST14	COX15	DDHD1	EBP	FANCA
ACVRL1	AMT	B3GAT3	CBS	CHST3	COX6B1	DDHD2	ECEL1	FANCB
ACY1	ANKH	B3GLCT	CC2D1A	CHSY1	COX7B	DDOST	EDA	FANCC
ADA	ANKRD11	B4GALT7	CC2D2A	CHUK	CPS1	DDR2	EDAR	FANCD2
ADAMTS10	ANKRD26	B9D1	CCBE1	CIB2	CPT2	DDX11	EDN1	FANCE
ADAMTS17	ANO5	BANF1	CCDC103	CISD2	CRADD	DDX3X	EDNRA	FANCF
ADAMTSL2	ANOS1	BBS1	CCDC114	CKAP2L	CRB1	DDX59	EDNRB	FANCG
ADAR	ANTXR1	BBS10	CCDC151	CLCN7	CRB2	DEAF1	EEF1A2	FANCI
ADGRG1	ANTXR2	BBS12	CCDC22	CLCNKB	CREBBP	DENND5A	EFNB1	FANCL
ADNP	AP1S2	BBS2	CCDC39	CLDN19	CRELD1	DEPDC5	EFTUD2	FANCM
ADSL	AP3B1	BBS4	CCDC40	CLMP	CRLF1	DHCR24	EGR2	FAR1
AFF2	AP3B2	BBS5	CCDC65	CLN3	CRTAP	DHCR7	EHMT1	FAT4
AFF3	AP4B1	BBS7	CCDC78	CLN5	CRX	DHDDS	EIF2AK3	FBLN5
AFF4	AP4E1	BBS9	CCDC8	CLN6	CRYAA	DHFR	EIF2B3	FBN1
AGA	AP4M1	BCAP31	CCDC88C	CLN8	CRYBA1	DHH	EIF2S3	FBN2
AGK	AP4S1	BCKDHA	CCND2	CLP1	CRYBA4	DHODH	EIF4A3	FBP1
AGL	APOPT1	BCKDHB	CCNO	CLPB	CRYBB1	DHTKD1	ELAC2	FBXL4
AGPAT2	APTX	BCL11A	CCNQ	CLPP	CRYBB2	DIS3L2	ELN	FEZF1
AGPS	AR	BCOR	CD151	CNKSR2	CRYBB3	DKC1	ELOVL4	FGD1
AGRN	ARFGEF2	BCS1L	CD96	CNOT3	CRYGC	DLAT	EMC1	FGD4
AGXT	ARG1	BFSP2	CDAN1	CNTNAP1	CRYGD	DLD	EMD	FGF10
AHDC1	ARHGAP31	BGN	CDC6	CNTNAP2	CSPP1	DLG3	EMG1	FGF17
AHI1	ARID1A	BHLHA9	CDH1	COASY	CSTA	DLG4	ENPP1	FGF3
AIFM1	ARID1B	BICD2	CDH3	COG1	CSTB	DLL3	EOGT	FGF8
AIMP1	ARID2	BIN1	CDK5RAP2	COG4	CTC1	DLL4	EP300	FGF9
AIPL1	ARL13B	BLM	CDKL5	COG5	CTCF	DMP1	EPG5	FGFR1

FGFR2	GATM	GSPT2	IFITM5	KCNQ5	LMNA	MID1	NDUFA10	NYX
FGFR3	GBA	GTF2H5	IFT122	KCNT1	LMOD3	MIR17HG	NDUFAF2	OBSL1
FH	GBA2	GTPBP3	IFT140	KCTD1	LMX1B	MITF	NDUFB11	OCLN
FHL1	GBE1	GUCY2C	IFT172	KCTD7	LONP1	MKKS	NDUFS1	OCRL
FIG4	GCDH	GUSB	IFT43	KDM1A	LRAT	MKS1	NDUFS4	ODAPH
FKBP14	GCH1	H19	IFT80	KDM5C	LRBA	MLC1	NDUFS7	OFD1
FKRP	GDF5	H3F3A	IGF1	KDM6A	LRIG2	MLH1	NDUFS8	OPHN1
FKTN	GDF6	HACE1	IGF1R	KIAA1109	LRIT3	MLYCD	NDUFV1	ORC1
FLNA	GDI1	HADH	IGF2	KIF11	LRP2	MMAA	NEB	ORC4
FLNB	GFAP	HADHA	IGFBP7	KIF1A	LRP4	MMAB	NECTIN4	ORC6
FLRT3	GFM1	HAX1	IGHMBP2	KIF1BP	LRP5	MMACHC	NEDD4L	OSTM1
FLT4	GHR	HCCS	IGSF1	KIF22	LRPPRC	MMADHC	NEK1	OTC
FLVCR1	GJA1	HCFC1	IHH	KIF2A	LRRC6	MMP13	NEK8	OTOGL
FLVCR2	GJA3	HCN1	IKBKG	KIF5C	LTBP2	MNX1	NEK8	OTX2
FMN2	GJA8	HDAC4	IL11RA	KIF7	LTBP3	MOCS1	NEU1	OXCT1
FMR1	GJB2	HDAC8	IL17RD	KISS1R	LTBP4	MOCS2	NEU1	P3H1
FN1	GJC2	HES7	IL1RAPL1	KIT	LYST	MOGS	NEXMIF	PACS1
FOLR1	GK	HESX1	IMPAD1	KLF1	LZTFL1	MPDU1	NF1	PAFAH1B1
FOXC1	GLB1	HEXA	INPP5K	KLHL40	MAB21L2	MPI	NFIX	PAH
FOXC2	GLDC	HEXB	INPPL1	KLHL41	MAF	MPLKIP	NFU1	PAK3
FOXE1	GLE1	HGSNAT	INSR	KLHL7	MAFB	MPV17	NGLY1	PALB2
FOXE3	GLI2	HIBCH	INVS	KMT2A	MAGEL2	MPZ	NHEJ1	PAPSS2
FOXF1	GLI3	HINT1	IQCB1	KMT2C	MAMLD1	MRE11	NHP2	PARN
FOXG1	GLIS2	HIVEP2	IQSEC2	KMT2D	MAN1B1	MRPS22	NHS	PAX2
FOXL2	GLIS3	HLCS	IRF6	KMT5B	MAN2B1	MSH2	NIPBL	PAX3
FOXN1	GLMN	HMGCL	IRX5	KPTN	MANBA	MSH6	NKX2-1	PAX6
FOXP1	GLUD1	HMGCS2	ISPD	KRAS	MAP2K1	MSX1	NKX2-5	PAX8
FOXP2	GLUL	HMX1	ITCH	KRIT1	MAP2K2	MSX2	NKX3-2	PAX9
FOXP3	GM2A	HNF1B	ITGA3	KRT74	MAP3K1	MTHFR	NMNAT1	PC
FOXRED1	GMPPA	HNF4A	ITGA6	KYNU	MAPRE2	MTM1	NODAL	PCBD1
FRAS1	GMPPB	HNRNPU	ITGA7	L1CAM	MASP1	MTO1	NOG	PCCA
FREM1	GNA11	HOXA1	ITGA8	L2HGDH	MAT1A	MTOR	NOTCH1	PCCB
FREM2	GNA14	HOXA13	ITGB4	LAMA1	MATN3	MTR	NOTCH2	PCDH19
FRMD7	GNAI3	HOXB1	ITPR1	LAMA2	MBOAT7	MTRR	NPC1	PCGF2
FRMPD4	GNAO1	HOXC13	IVD	LAMA3	MBTPS2	MUSK	NPC2	PCNT
FRRS1L	GNAQ	HOXD13	JAG1	LAMB1	MC2R	MYBPC1	NPHP1	PCYT1A
FTCD	GNAS	HPD	JAGN1	LAMB3	MCCC1	MYCN	NPHP3	PDCD10
FTL	GNB5	HPGD	JAK3	LAMC2	MCCC2	MYH3	NPHP4	PDE10A
FTSJ1	GNPAT	HPRT1	JAM3	LAMC3	MCEE	MYH6	NPHS1	PDE4D
FUCA1	GNPTAB	HPS1	KANSL1	LAMP2	MCOLN1	MYH8	NPHS2	PDE6G
FUZ	GNPTG	HPSE2	KARS	LARGE1	MCPH1	MYH9	NPR2	PDE6H
FYCO1	GNS	HR	KAT6A	LARP7	MDH2	MYLK	NR0B1	PDGFRB
FZD6	GORAB	HRAS	KAT6B	LARS2	MECOM	MYO5A	NR2F1	PDHA1
G6PC3	GPC3	HSD17B10	KBTBD13	LBR	MECP2	MYO5B	NR2F2	PDHX
GAA	GPC6	HSD17B3	KCNA2	LDB3	MED12	MYO7A	NR5A1	PDSS1
GABRA1	GPI	HSD17B4	KCNB1	LEMD3	MED13L	MYT1	NRAS	PDSS2
GABRB3	GPSM2	HSD3B7	KCNC1	LFNG	MED17	MYT1L	NRXN2	PEPD
GABRG2	GPX4	HSF4	KCNC3	LGI4	MEF2C	NAA10	NSD1	PET100
GALC	GRHL2	HSPD1	KCNE1	LHX3	MEGF10	NAA15	NSDHL	PEX1
GALE	GRHL3	HSPG2	KCNH1	LHX4	MEGF8	NAGA	NSMF	PEX10
GALK1	GRIA3	HUWE1	KCNJ1	LIAS	MEOX1	NAGLU	NSUN2	PEX11B
GALNS	GRIK2	HYAL1	KCNJ10	LIFR	MESP2	NAGS	NT5C2	PEX12
GALT	GRIN1	HYDIN	KCNJ11	LIG4	MFRP	NALCN	NT5C3A	PEX13
GAMT	GRIN2A	HYLS1	KCNJ2	LINS1	MFSD2A	NBAS	NTRK1	PEX14
GATA2	GRIN2B	IDS	KCNJ6	LIPN	MFSD8	NBN	NTRK2	PEX16
GATA4	GRIP1	IDUA	KCNQ1	LIPT1	MGAT2	NDE1	NUBPL	PEX19
GATA6	GRM1	IER3IP1	KCNQ2	LMBR1	MGP	NDP	NUP107	PEX2
GATAD2B	GRM6	IFIH1	KCNQ3	LMBRD1	MICU1	NDUFA1	NUP62	PEX26

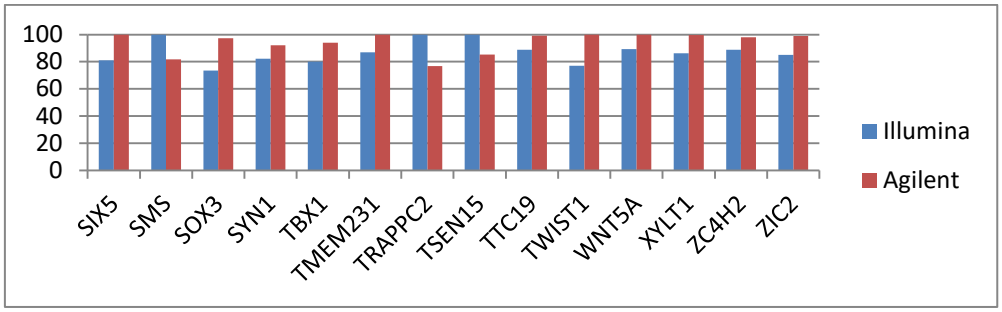
PEX3	POMK	RAF1	SAMHD1	SLC26A3	SP110	TBX6	TRIP11	VDR
PEX5	POMT1	RAI1	SASS6	SLC27A4	SPAG1	TBXAS1	TRIP12	VIPAS39
PEX6	POMT2	RAPSN	SATB2	SLC2A1	SPECC1L	TCF12	TRIP13	VLDLR
PEX7	POR	RARB	SBDS	SLC2A10	SPEG	TCF4	TRPM1	VPS13B
PGAP1	PORCN	RARS2	SC5D	SLC2A2	SPG11	TCIRG1	TRPS1	VPS33B
PGAP2	POU1F1	RASA1	SCARF2	SLC33A1	SPR	TCN2	TRPV3	VPS53
PGAP3	PPIB	RAX	SCN11A	SLC35A1	SPRED1	TCOF1	TRPV4	VRK1
PGK1	PPM1D	RBM10	SCN1A	SLC35A2	SPRY4	TCTN1	TSC1	VSX2
PGM1	PPP2R1A	RBM8A	SCN1B	SLC35C1	SPTAN1	TCTN2	TSC2	WAC
PGM3	PPP2R5D	RBPJ	SCN2A	SLC35D1	SPTLC2	TCTN3	TSEN15	WASHC5
PHF21A	PPP3CA	RECQL4	SCN3A	SLC37A4	SRCAP	TECPR2	TSEN2	WDPCP
PHF6	PPT1	RELN	SCN4A	SLC39A13	SRD5A2	TEK	TSEN34	WDR11
PHF8	PQBP1	REN	SCN8A	SLC46A1	SRD5A3	TELO2	TSEN54	WDR19
PHGDH	PREPL	RERE	SCO1	SLC4A1	SRY	TERT	TSHB	WDR34
PHIP	PRG4	RET	SCO2	SLC4A11	ST14	TFAP2A	TSHR	WDR35
PHOX2B	PRKAR1A	RETREG1	SDCCAG8	SLC4A4	ST3GAL3	TFAP2B	TSPAN7	WDR45
PIEZO1	PRKD1	RFT1	SDHA	SLC52A2	ST3GAL5	TGDS	TTC19	WDR60
PIEZO2	PRMT7	RFX6	SDHAF1	SLC52A3	STAG1	TGFB1	TTC21B	WDR62
PIGA	PROK2	RIN2	SEC23B	SLC5A5	STAG2	TGFB2	TTC37	WDR73
PIGL	PROKR2	RIPK4	SEC24D	SLC5A7	STAMBP	TGFB3	TTC7A	WNT1
PIGN	PROP1	RIT1	SECISBP2	SLC6A1	STAR	TGFBR1	TTC8	WNT10B
PIGO	PRPS1	RLIM	SELENON	SLC6A17	STAT5B	TGFBR2	TTI2	WNT3
PIGT	PRRT2	RMND1	SETBP1	SLC6A3	STIL	TGIF1	TUBA1A	WNT4
PIGV	PRSS12	RMRP	SETD2	SLC6A5	STRA6	TGM1	TUBA8	WNT5A
PIK3CA	PRSS56	RNASEH2A	SETD5	SLC6A8	STS	TH	TUBB	WNT7A
PIK3R1	PRX	RNASEH2B	SF3B4	SLC6A9	STX1B	THAP1	TUBB2A	WRAP53
PIK3R2	PSAP	RNASEH2C	SGCA	SLC9A6	STXBP1	THOC2	TUBB2B	WT1
PITX1	PSAT1	RNASET2	SGSH	SLX4	SUCLG1	THOC6	TUBB3	WWOX
PITX2	PSMB8	RNU4ATAC	SH3PXD2B	SMAD3	SUFU	THRA	TUBB4A	XPA
PITX3	PSPH	ROBO1	SHANK1	SMAD4	SUMF1	TIMM8A	TUBG1	XPC
PKD2	PTCH1	ROGDI	SHANK2	SMARCA2	SURF1	TINF2	TUBGCP4	XRCC4
PKHD1	PTCHD1	ROR2	SHANK3	SMARCA4	SYN1	TK2	TUBGCP6	XYLT1
PKLR	PTDSS1	RORA	SHH	SMARCAL1	SYNE1	TMCO1	TUFM	XYLT2
PLA2G6	PTEN	RPE65	SHOC2	SMARCB1	SYNGAP1	TMEM138	TUSC3	YAP1
PLCB1	PTF1A	RPGRIP1	SHOX	SMARCE1	SYP	TMEM165	TWIST1	YY1
PLCB4	PTH	RPGRIP1L	SHROOM3	SMC1A	SZT2	TMEM216	TWIST2	ZBTB18
PLCE1	PTH1R	RPL11	SIL1	SMC3	TAB2	TMEM231	TXNL4A	ZBTB20
PLK4	PTHLH	RPL35A	SIX1	SMCHD1	TAC3	TMEM237	TYR	ZC4H2
PLOD1	PTPN11	RPL5	SIX3	SMN1	TACO1	TMEM67	TYRP1	ZDHHC9
PLOD2	PTPN14	RPS10	SIX5	SMO	TACR3	TMEM70	UBA1	ZEB2
PLP1	PTS	RPS17	SKI	SMOC1	TAF1	TMPRSS6	UBE2A	ZFP57
PMM2	PUF60	RPS19	SKIV2L	SMOC2	TAT	TMTC3	UBE2T	ZFYVE26
PMP22	PURA	RPS24	SLC12A1	SMPD1	TAZ	TNFRSF11B	UBE3A	ZIC1
PMS2	PXDN	RPS26	SLC12A6	SMS	TBC1D20	TNFRSF13B	UBE3B	ZIC2
PNKP	PYCR1	RPS6KA3	SLC13A5	SNAP25	TBC1D23	TNNI2	UBR1	ZIC3
PNPLA1	PYCR2	RRAS	SLC16A2	SNAP29	TBC1D24	TNNT1	UBTF	ZMPSTE24
PNPT1	PYGL	RRM2B	SLC17A5	SNRPB	TBCE	TNXB	UGT1A1	ZMYND10
POC1A	QARS	RSPH1	SLC19A3	SNRPE	TBCK	TP63	UMPS	ZMYND11
POC1B	QDPR	RSPH4A	SLC1A2	SNX14	TBL1XR1	TPM2	UNC80	ZNF423
POGZ	RAB18	RSPH9	SLC22A5	SOS1	TBR1	TPM3	UPF3B	ZNF711
POLD1	RAB23	RSPO4	SLC24A4	SOST	TBX1	TPP1	UQCRB	ZNF750
POLG	RAB39B	RTEL1	SLC25A15	SOX10	TBX15	TRAPPC11	UQCRQ	ZSWIM6
POLR1C	RAB3GAP1	RTTN	SLC25A19	SOX11	TBX18	TRAPPC2	UROC1	
POLR1D	RAB3GAP2	RUNX2	SLC25A20	SOX17	TBX20	TRAPPC9	UROS	
POLR3A	RAC1	RYR1	SLC25A22	SOX2	TBX22	TREX1	USB1	
POLR3B	RAD21	SACS	SLC25A38	SOX3	TBX3	TRIM32	USP27X	
POMGNT1	RAD51	SALL1	SLC25A4	SOX5	TBX4	TRIM37	USP9X	
POMGNT2	RAD51C	SALL4	SLC26A2	SOX9	TBX5	TRIO	UVSSA	

Appendix 4- Coverage for individual genes in Family 82

Where at least one sequencing method was fewer than 90% coverage at ≥20x coverage



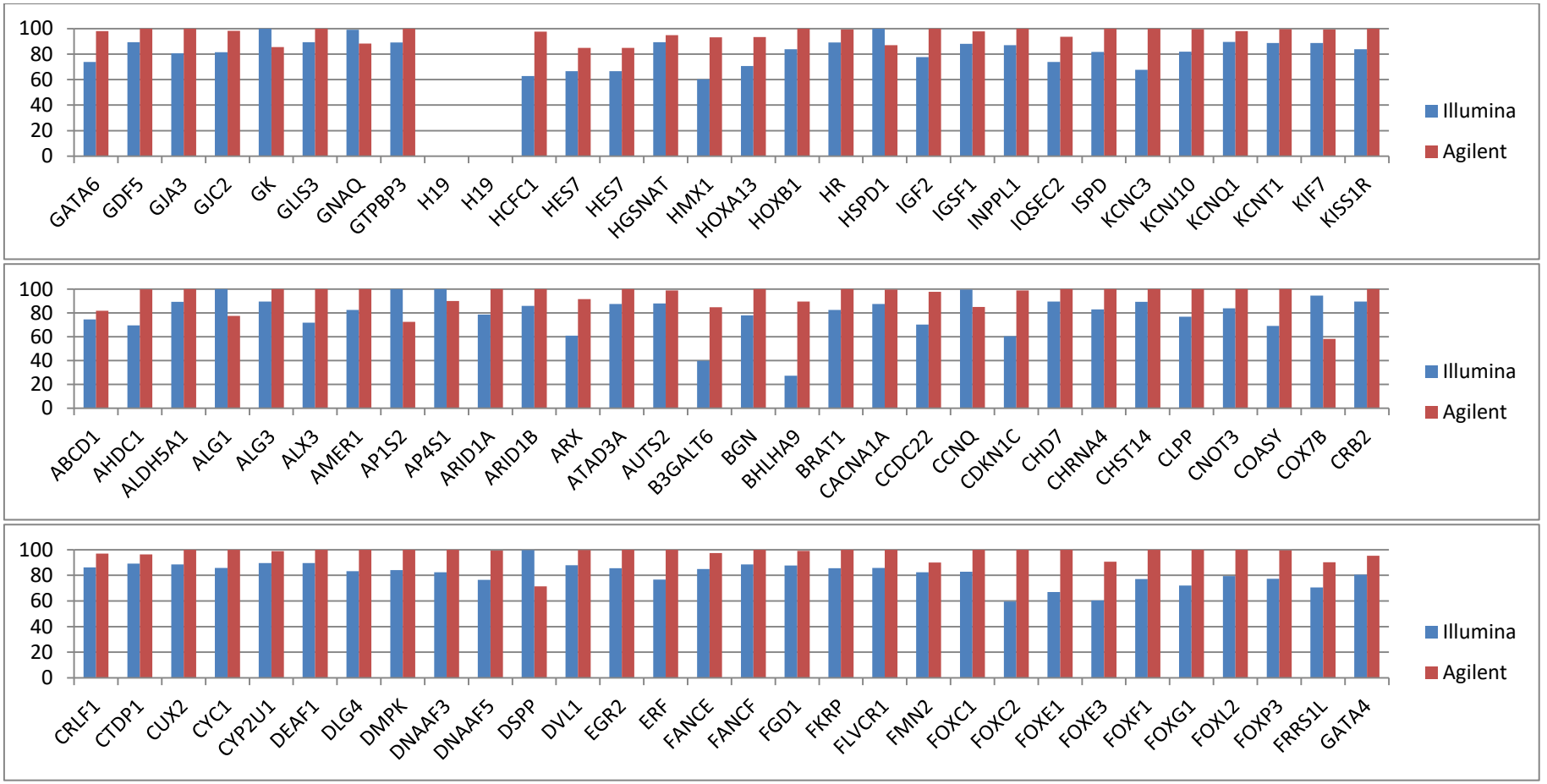
pp


100
80
60
40
20
0
SIX5
SMS
SOX3
SYN1
TBX1
TMEM231
TRAPPC2
TSEN15
TTC19
TWIST1
WNT5A
XYLT1
ZC4H2
ZIC2
Illumina
Agilent

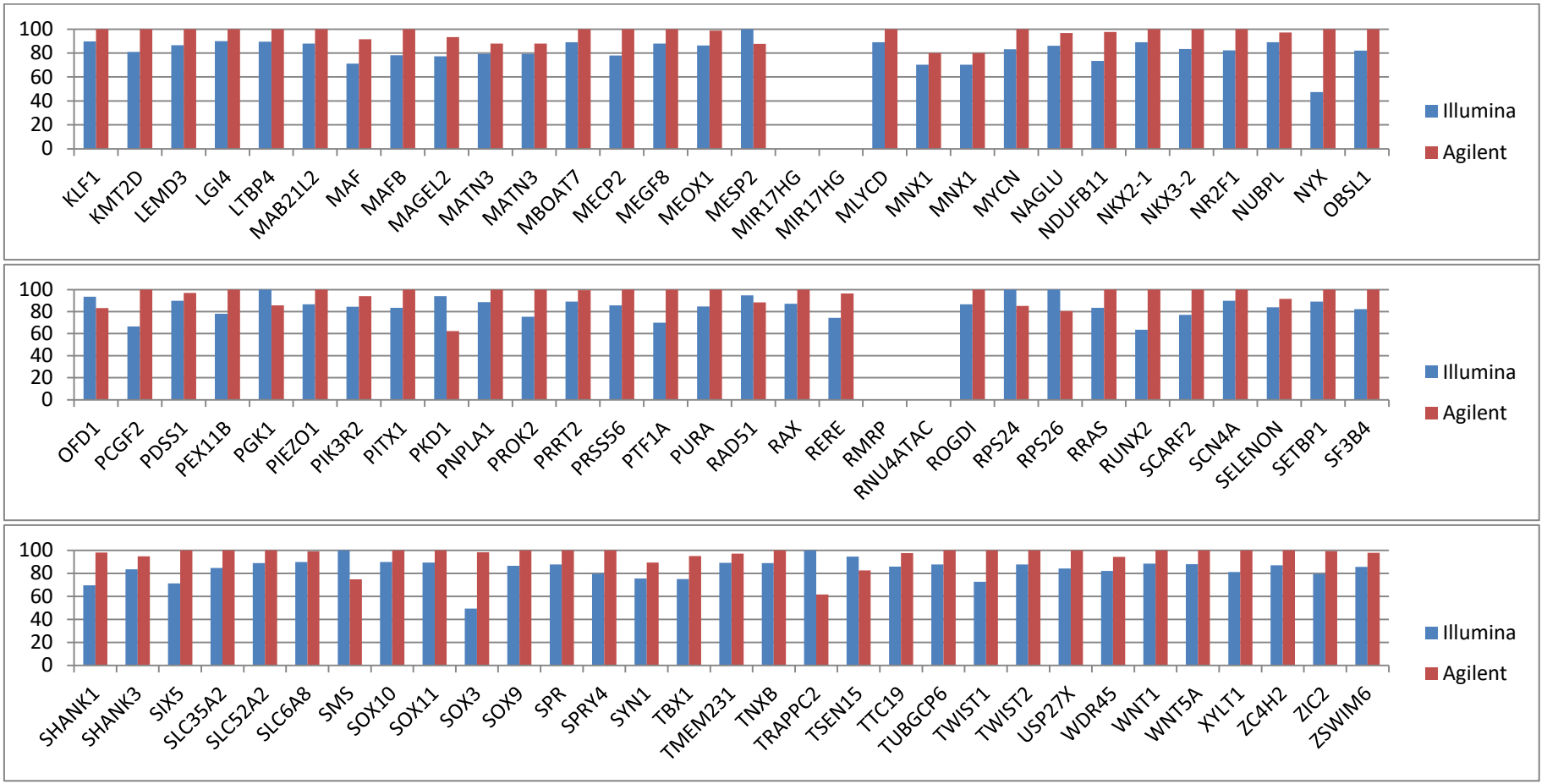
qq

Appendix 5- Coverage for individual genes in Family 141

Where at least one sequencing method was fewer than 90% coverage at ≥20x coverage

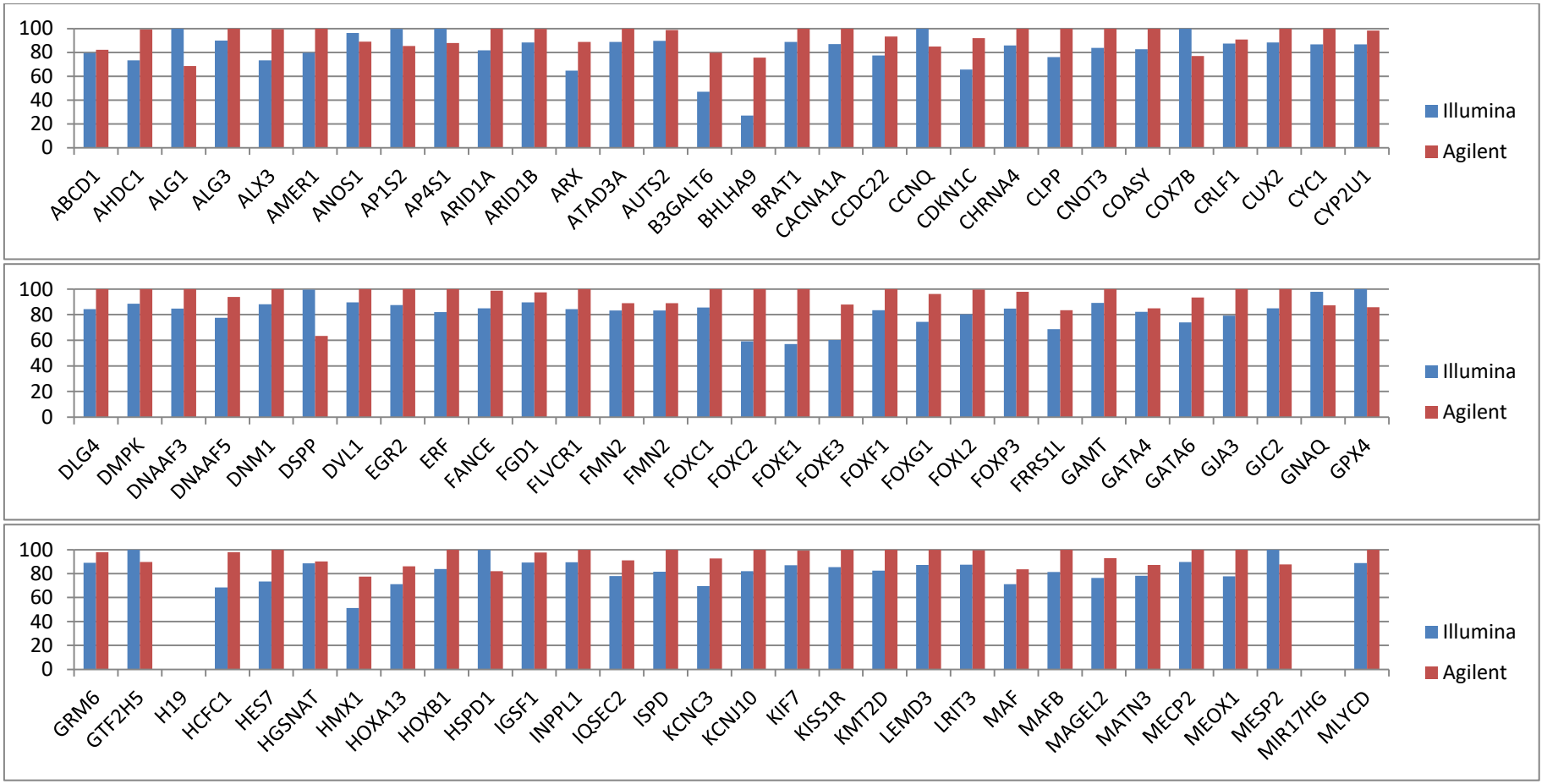


rr

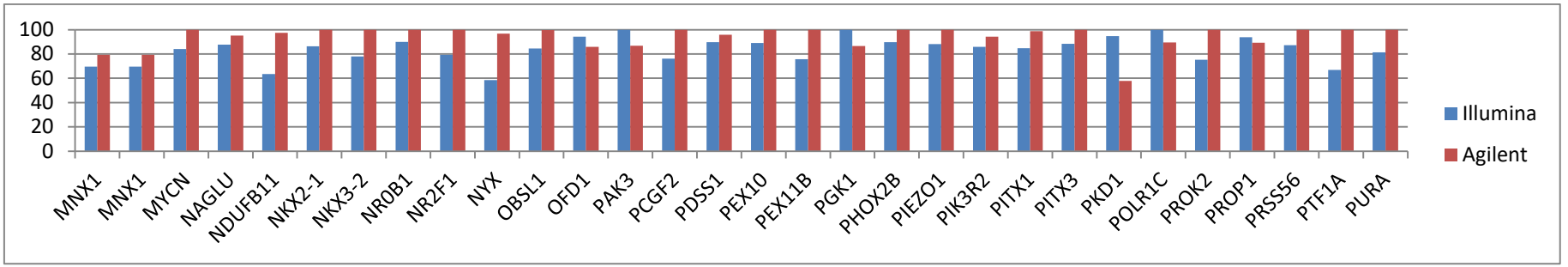

100
80
60
40
20
0
KLF1 KMT2D LEMD3 LGI4 LTBP4 MAB21L2 MAF MAFB MAGEL2 MATN3 MATN3 MBOAT7 MECP2 MEGF8 MEOX1 MESP2 MIR17HG MIR17HG MLYCD MNX1 MNX1 MYCN NAGLU NDUFB11 NKX2-1 NKX3-2 NR2F1 NUBPL NYX OBSL1
Illumina
Agilent
OFD1 PCGF2 PDSS1 PEX11B PGK1 PIEZO1 PIK3R2 PITX1 PKD1 PNPLA1 PROK2 PRRT2 PRSS56 PTF1A PURA RAD51 RAX RERE RMRP RNU4ATAC ROGDI RPS24 RPS26 RRAS RUNX2 SCARF2 SCN4A SELENON SETBP1 SF3B4
SHANK1 SHANK3 SIX5 SLC35A2 SLC52A2 SLC6A8 SMS SOX10 SOX11 SOX3 SOX9 SPR SPRY4 SYN1 TBX1 TMEM231 TNXB TRAPPC2 TSEN15 TTC19 TUBGCP6 TWIST1 TWIST2 USP27X WDR45 WNT1 WNT5A XYLT1 ZC4H2 ZIC2 ZSWIM6

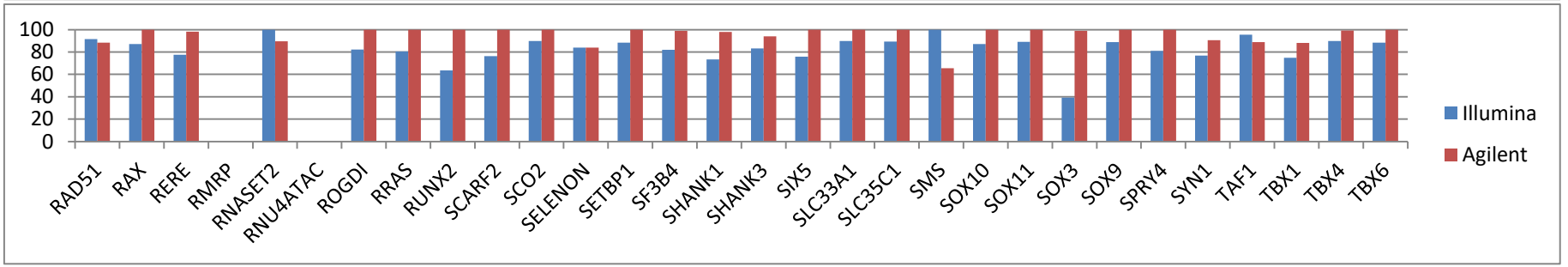
Appendix 6- Coverage for individual genes in Family 154

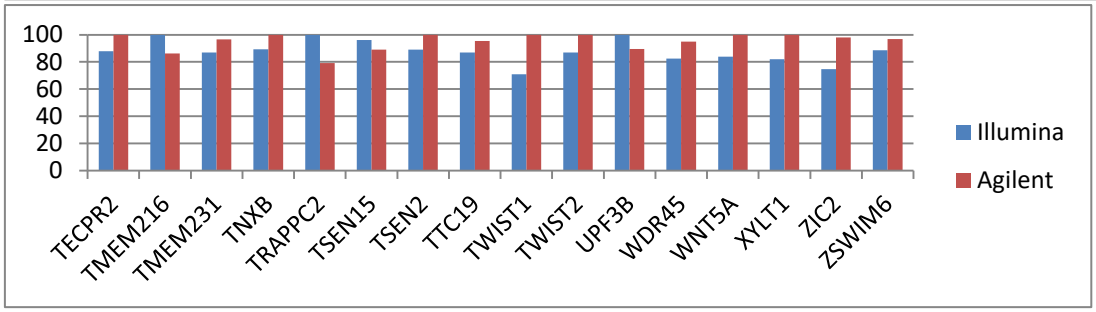
Where at least one sequencing method was fewer than 90% coverage at ≥20x coverage



tt


100
80
60
40
20
0
MNX1
MNX1
MYCN
NAGLU
NDUFB11
NKX2-1
NKX3-2
NR0B1
NR2F1
NYX
OBSL1
OFD1
PAK3
PCGF2
PDSS1
PEX10
PEX11B
PGK1
PHOX2B
PIEZO1
PIK3R2
PITX1
PITX3
PKD1
POLR1C
PROK2
PROP1
PRSS56
PTF1A
PURA
Illumina
Agilent


100
80
60
40
20
0
RAD51
RAX
RERE
RMRP
RNASET2
RNU4ATAC
ROGDI
RRAS
RUNX2
SCARF2
SCO2
SELENON
SETBP1
SF3B4
SHANK1
SHANK3
SIX5
SLC33A1
SLC35C1
SMS
SOX10
SOX11
SOX3
SOX9
SPRY4
SYN1
TAF1
TBX1
TBX4
TBX6
Illumina
Agilent


100
80
60
40
20
0
TECPR2
TMEM216
TMEM231
TNXB
TRAPPC2
TSEN15
TSEN2
TTC19
TWIST1
TWIST2
UPF3B
WDR45
WNT5A
XYLT1
ZIC2
ZSWIM6
Illumina
Agilent

uu