University of Ferrara

PhD COURSE IN
"MOLECULAR MEDICINE AND PHARMACOLOGY"

CYCLE XXXII

DIRECTOR Prof. Francesco Di Virgilio

*G*ENETICS AND EPIGENETIC STUDY IN THE FETAL-MATERNAL DIADE IN RECURRENT PREGNANCY LOSS (RPL)

Scientific Disciplinary Sector (SDS) MED/03

Candidate:
Dott. Aleotti Valentina

Supervisor:
Prof. Michele Rubini

Years 2016/2020

"Ever tried. Ever failed. No matter. Try again. Fail again. Fail better."

Samuel Beckett

Preface

Conference and Lectures

Poster at European Society of Human Reproduction and Embriology 2020, Copenhagen, Denmark.
Identification of genomic variants responsible for pregnancy loss: a pilot study.
Silvia Buonaiuto∗, Imma Di Biase∗, **Aleotti Valentina**∗, Amin Ravaei, Gianluca Damaggio, Palmira D'Ambrosio, Oriana Catapano, Gabriella Esposito, Marco Chierici, Madhuri Pulijala, Qasim Ayub, Cesare Furlanello, Erik Garrison, Nicole Soranzo, Antonio Capalbo, Michele Rubini, Sebastiano Di Biase, Vincenza Colonna (∗ equal contribution).

Poster at Biology of Genomes,Cold Spring Harbor Laboratory, US,2020.
Deciphering genetic causes of idiopathic pregnancy loss from an embryonic perspective. Silvia Buonaiuto∗, Imma Di Biase∗, **Aleotti Valentina**∗, Gianluca Damaggio, Palmira D'Ambrosio, Oriana Catapano, Amin Ravaei, Gabriella Esposito, Marco Chierici, Madhuri Pulijala, Qasim Ayub, Cesare Furlanello, Erik Garrison, Nicole Soranzo, Antonio Capalbo, Michele Rubini, Sebastiano Di Biase, Vincenza Colonna (∗ equal contribution).

Poster NETTAB/BBCC 2019 Meeting November 11-13, 2019 Fisciano Campus, University of Salerno, Italy.
Identification of genomic variants responsible for pregnancy loss: a pilot study
Silvia Buonaiuto∗, Imma Di Biase∗, **Valentina Aleotti**∗, Gianluca Damaggio, Palmira D'Ambrosio, Oriana Catapano, Gabriella Esposito, Marco Chierici, Madhuri Pulijala, Qasim Ayub, Cesare Furlanello, Erik Garrison, Nicole Soranzo, Michele Rubini, Sebastiano Di Biase, and Vincenza Colonna. (∗ equal contribution)

Poster at Conference AGI-SIMAG 26-28 September 2019 Cortona, Arezzo, Italy.
Genomics of Pregnancy Loss. Silvia Buonaiuto∗, Imma Di Biase∗, **Aleotti Valentina**∗, Gianluca Damaggio, Palmira D'Ambrosio, Oriana Catapano, Gabriella Esposito, Marco Chierici, Madhuri Pulijala, Qasim Ayub, Cesare Furlanello, Erik Garrison, Nicole Soranzo, Michele Rubini, Sebastiano Di Biase, and Vincenza Colonna. (∗ equal contribution)

Poster on Departmental Day of Research Unife on 16/09/2020, Ferrara, Italy.
*Interplay beetween Embryo's and Mother's MTHFR genotype in the pathogenesis of Recurrent Miscarrige.*Capucci R., Brasile O., Graziano A., Poggi A., Dall'Olio L., **Aleotti V.**, Ravaei A., Khan M. F. J., Aiello V., Astolfi G., Rubini M.

Poster on Departmental Day of Research Unife on 16/09/2020, Ferrara, Italy.
*Epigenetics of Recurrent Miscarriage: Analysis of Global DNA Methylation in Embryo's Chorionic Villi and in Mother's Blood Cells.***Aleotti V.**, Capucci R., Brasile O., Khan M.

F. J., Aiello V., Astolfi G., Sirignano C., Dall'Olio L., Ravaei A., Rubini M.

Oral Presentation- Seminar for PhD student, 24/05/2019, University of Ferrara, Ferrara, Italy.
Epigenetics of recurrent miscarriage: analysis of global DNA methylation in embryo's chorionic villi and maternal blood. **Dr. Aleotti Valentina**.

Poster on conference Agorà SIMP, 22-24 March 2018, Catania, Italy.
Maternal and embryo's MTHFR genotypes influence the risk of spontaneous abortion. Capucci R., Brasile O., Graziano A., Poggi A., Bassi Andreasi R.,**Aleotti V.** , Ravaei A., Khan MFJ., Aiello V., Astolfi G., Rubini M.

Oral Presentation - 3rd International Day for Sjögren, 22-23 July 2017-Salerno, Italy.
Translational research on Sjögren syndrome - **Dr. Aleotti Valentina, Prof. Michele Rubini**.

Oral Presentation- Seminar for PhD student, 03 March 2017- University of Ferrara, Ferrara, Italy.
Folate genetics and clinical outcome- **Dr.Aleotti Valentina, Prof. Michele Rubini**.

Paper

Khan MFJ, Little J, **Aleotti V**, Mossey PA, Steegers-Theunissen RPM, Autelitano L, Meazzini MC, Ravaei A, Rubini M. "LINE-1 methylation in cleft lip tissues: influence of infant MTHFR c.677C>T genotype". Oral Diseases 2019 (IF 2.43) PMID: 31161688

Contents

List of Figures

List of Tables

Dedicated to Giuseppe and Lorenzo...

Chapter 1

Introduction

1.1 Miscarriage

Pregnancy loss (PL) is defined by *World Health Organization*(**WHO**) as the loss of pregnancy resulting in fetal death before 20-28 weeks and by Italian law as the natural interruption of pregnancy caused by pathologies. In particular, any expulsion or fetal/embryo death occurring within 180 days (25 weeks and 5 days) (**ISTAT**). Roughly 10-15 % of all clinically recognized pregnancies result in miscarriage. Recurrent pregnancy loss (RPL) denotes two or more consecutive spontaneous loss of pregnancy happening within 20-24 weeks of gestation. RPL affects at least 1-2% of women in reproductive age. Most of the miscarriages happens during the first trimester of gestation. The main known causes of RPL are chromosomal anomalies, anatomical abnormalities, autoimmune diseases, infection diseases, advanced maternal and paternal age and environmental agents. In 50% of cases, the causes of RPL are still unknown [1–5]. The diagnosis of miscarriages is based on embryo heart activity and gestational sac features revealed by ultrasonography. However, the diagnosis takes place only after the death of the embryo, and only a few cases are followed to understand the genetic causes with techniques that can discriminate aneuploidies (karyotyping, quantitative PCR) or large deleterious copy number variants (comparative genomic hybridization), while no information is available on small-size DNA changes incompatible with life. Thus, the current ability to inform prognosis and manage decisions in cases of perinatal lethality is limited, with important consequences in counselling for RPLs and in-vitro fertilization. The protocols to treat this pathology provide a committed and focused service to couples who already had experience of RPL, so it doesn't prevent from a new event of miscarriage. The state of affairs in Italy was reported by the last study of the *Italian National Institute of Statistics* (ISTAT). The amount of RPL increased from 1982 to 2015. From 2001 the number starts to increase, with a maximum in 2007 of 77.129 cases. In 2015 were registered 66.127 cases, about 14% of clinically recognized pregnancies. Figure 1.1 [6] shows the rate of miscarriage in function of the age of the mother: the maximum values are with class of ages 30-34 in 1995 and 2005. Instead for 2015 the peak is with the next class of mean age group. In later years from 1995 to 2015 the rate decreased slightly for women of younger age (up to 29 years old). Starting at 30 years old, the rate increases significantly with age: +5% between 30 and 34 years old, +54 % in the 35-39 class, double in the 40-44 class and +45% in 45-49 class.

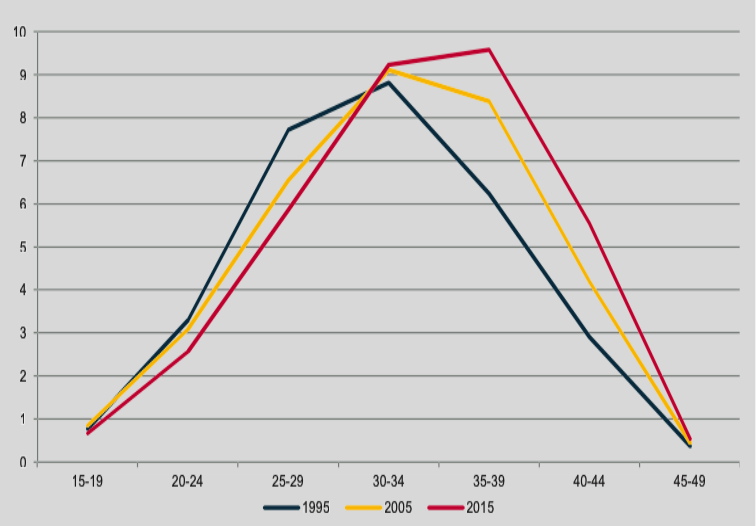

FIGURE 1.1: **Rate of RPL in Italy for class of age of 1995, 2005, 2015 (1.000 women)** (***ISTAT***, *La salute riproduttiva della donna*)

1.2 Anatomy of placenta and use of the villi to extract DNA from embryos

One week after fertilization, the blastocyst hatches from the pellucid zone and subsequently implants into the uterine wall. The placenta is a unique materno-fetal organ which begins to grow around the point of implantation and is delivered at the birth of the fetus. The placenta has several biological functions:

• **Metabolism.** The fetus takes the nutrients essential for its growth from the placenta. Some nutrients like glycogen, cholesterol, and fatty acids are synthesized in the placenta. It transport from and to the embryo of gases, nutrients, drugs, hormones, antibodies, wastes, and infectious agents.

• **Endocrine.** After the dismission of the Corpus luteum there is an increase of production by the placenta of human chorionic gonadotropin (HCG) that is also a marker for the begin of pregnancy. The placenta also produces progesterone that helps the embryo during the implant, estrogen for proliferation and the growth of the uterus and enlargement of the breasts. Human placental lactogen has a structure and function similar to growth hormone and modifies the metabolism of the mother to facilitate the energy supply to the fetus.

The placenta is composed of four layers that separate maternal and fetal blood [7]. A schematic of the villi can be seen in figure 1.2:

• **Syncitiotrophoblast** is the outer layer of the trophoblast and is essential for the first stage in the uterine wall invasion and homing. It also establishes an interface between maternal blood and embryo.

• **Cytotrophoblast** is the inner layer of the trophoblast between Syncitiotrophoblast and the external wall of the blastocyst. It contains trophoblastic stem cell.

• **Villi connective tissue** is involved in the invasion of the uterine decidua and at the same time is essential to absorb nutrients from the mother for the growth of the embryo.

• **Fetal capillary endothelium** is the last layer across which all exchange of gases, nutrients, hormones, and wastes occurs.

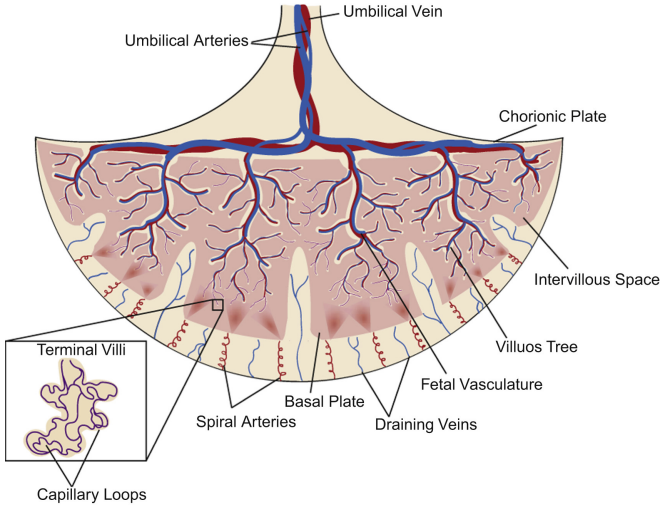

FIGURE 1.2: **Schematic representation of the anatomy of the villi**.
Mayo R.P et al., 2016

1.3 Causes of miscarriage

For 50-60% of RPL cases, the cause is recognized to be chromosomal abnormalities, endocrine defects, an anatomic malformation in utero, thrombophilic, autoimmune disease, infections, paternal causes and environmental factors. Nonetheless, for the remaining 40-50% the causes are unknown [8].

1.3.1 Genetic Factor of RPL

Among RPL caused by genetic faults, it is estimated that 50-70% is due to meiotic chromosome segregation errors, whose frequency increases with increasing maternal age [9, 10]. In fact a complex series of age-related changes in nuclear and cytoplasmic competence, affecting such fundamental processes as spindle formation and chromosome segregation, mitochondrial function and the integrity of the cytoskeleton. A poor-quality oocyte is less likely to fertilise and, if fertilised, will produce an embryo which is generally slow to divide and unlikely to implant. This happens most frequently in women over 35 years of age [11]. Sperm DNA fragmentation induced by oxidative stress also causes RPL through impairment of placentation [12, 13]. Most of the RPL were caused by chromosomal anomalies. The most frequent anomalies are the numerical (86%), followed by mosaicism (8%) and the structural anomalies (6%) [14]. The incidence of chromosomal anomalies in the embryo is around 50% of couples affected by recurrent abortion in the first trimester and around 25% in the second trimester. The most common chromosomal alterations are trisomies (50%), with prevalent reference to chromosomes 13,14,15,16,18,21,22 followed by monosomy of chromosome X, from karyotype 45, X (13-19%) and polyploids 22% (16% triploidies), 7% structural aberrations, 8 % mosaicism or autosomal monosomies. The most frequent trisomy is 16 with 30% of all trisomies [8].

1.3.2 Anatomic malformation

Uterine malformations, although rare, are more often found in the population of women with poliabortivity problems. The anomalies of the fusion /absorption process of the Mullerian ducts are in figure 1.3:

1.**Agenesis/hypoplasia** characterized by the absence of the bass horn and the uterus or the uterus alone. There are 5 different types: i) Vaginal, ii) Cervical, iii) Fundal, iv) Tubal and v) Combined.
2.**Unicornuate** has an incidence of spontaneous abortion of 27-59% and derives from the lack of development of one of Muller's ducts. There are four different types: i)communicating, ii) non-communicating, iii) no cavity, iv) no horn.
3.**Didelphous** has an abortion incidence of 7-43% and consists of the presence of two completely separate emitters that generally end with two necks. The vagina may also have a longitudinal septum, which frequently results in the formation of a hemivagina with complete obstruction of the vaginal outflow.
4.**Bicornuate** is the cause of abortion between 1-27% of cases, and in this case, the two ducts of Muller take on each other but do not merge: this determines the development of a uterus with a single neck and two endometrial cavities, partially separated. It can be i) complete or ii) partial.
5.**Septum** consists in the septal absorption deficit that divides the two cavities originating from the Muller ducts and is present in 10-26% of abortion cases. In RPL cases is observed cervical incontinence, with an incidence of 0.05-1% of all pregnancies and is thought to be responsible for approximately 16% of second-trimester abortions.
6.**Arcuate** where the uterine cavity displays a concave contour towards the fundus.
7.**DES drug related** where the uterine cavity has a "T-shape", as a result of fetal exposure to diethylstilbestrol [8].

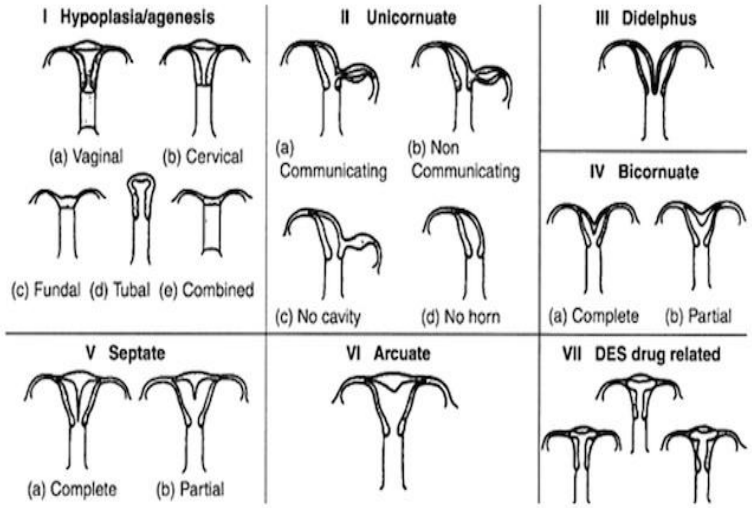

FIGURE 1.3: **Type of uterine malformation**

1.3.3 Endocrine defects

From the literature in 8-12% of cases of RPL, there is an association with endocrine diseases, like hypothyroidism, diabetes mellitus, poor secretion of progesterone by the corpus luteum, hyperprolactinemia and polycystic ovary syndrome (PCOS). Poor progesterone secretion during the luteal phase plays a significant role in RPL. It is recognized that the secretion of this hormone in pregnancy is very important for the survival of the embryo until the uteroplacental displacement occurs between the

7th and 9th week of gestation. It is found experimentally that progesterone exerts an immunomodulatory effect, contributing to the maintenance of pregnancy. The thyroid hormones are really important for the growth and metabolism of the developing fetus. Since the fetal thyroid gland reaches full training and functionality after the 12th week, the work before that week is done by the maternal thyroid. Hyperthyroidism does not seem to correlate with RPL, while hypothyroidism carries an increased risk of FPL and RPL. Although further studies are needed to understand why this happens. Polycystic ovary syndrome(PCOS) has been found in RPL and this risk could be cause by premature ovulation with the production of a ovum, which is not fertilizable or which once fertilized is lost prematurely. Women who suffer from PCOS are termed hypo-fertile because of their rare and irregular ovulation. Another endocrine disease that has been discovered as the cause of abortion is hyperprolactinemia. Prolactin is a hormone produced by the pituitary during pregnancy and is employed to prepare the breast for breastfeeding. Hyperprolactinemia occurs when prolactin levels exceed physiological ones, and this could interfere with normal gonadotropin secretion and with the activity of the corpus luteum, with consequent insufficiency of the luteal phase. Insulin-dependent diabetes mellitus also shows a correlation with abortions, in particular, there is a correlation between the percentage of blood glucose in the first trimester and the recurrent abortion rate. The destruction of the fetus is due to lethal malformations caused precisely by maternal disease [8, 15].

1.3.4 Autoimmune disease

Antiphospholipid syndrome (APS) is an autoimmune disease characterized by episodes of thrombosis and can cause RPL to pregnant women. This pathology is caused by a group of autoantibodies (antibodies directed against components of the tissue of the individual that produces it), called antiphospholipid antibodies. These seem to interfere with clotting causing various symptoms, a consequence of thrombosis in every part of the body. Another disease that can cause RPL in 30% of the cases is Systemic Lupus Erythematosus (SLE). An autoimmune disease can produce many autoantibodies that affecting all organs and district, caused a chronic inflammation. It prefers the female sex. As symptoms, there are polymorphic, cutaneous-visceral manifestations [8, 16].

1.3.5 Thromobofilic factors

Genetic predisposition to thrombophilia is a major cause of RPL. Thrombophilia is defined as a group of inherited or acquired conditions with an addition in the thrombo-embolic tendency. In most instances, these conditions remain asymptomatic. Instead, pregnancy is characterized by a state of hypercoagulation. In the presence of thrombophilia, thrombotic phenomena can take place in the vascular apparatus of the placenta causing loss of pregnancy. Thrombophilic factors are protein S deficiency, protein C deficiency, prothrombin gene mutation(PTG), antithrombin deficiency, pAPCR or Leiden factor V, factor XII deficiency, high plasma levels of factor VIII and factor XI, hyperfibrinogenomy, hyperhomocysteinemia [8, 17, 18].

1.3.6 Infections factors

There is a group of infections, which, if acquired during pregnancy, can be dangerous to the fetus. These infectious agents cross the placenta, infecting the fetus and

could cause its death. The most common infections during pregnancy are: *Toxoplasma*, *Rubella* , *Herpes simplex*, *Cytomegalovirus*, or untreated vaginal infections such as bacterial vaginosis, which can activate an inflammatory process with the production of mediators that trigger uterine contractions and lead to late abortions or preterm delivery. The risk change according to the infectious agent and the weeks of gestation; tends to decrease with the progression of the gestational age [8, 19].

1.3.7 Environmental factors

In literature, it seems that environmental factors contribute to the occurrence of RPL and FPL. Cigarette smoking hurts trophoblast function and is linked with a dose-dependent increase in the risk of miscarriage [20]. A survey conducted by *Miceli et al.in 2005* [21], highlighted the inhibiting effect that nicotine and its metabolites have on the production of progesterone and on the release of Prostaglandin E2 (PGE2) by human luteal cells. Likewise, carbon monoxide forms contribute to the production of high levels of fetal carboxyhemoglobin, which causes poor tissue oxygenation and a toxic effect on villus cells, reducing the passage of nutrients. Abuse of cocaine and assumption of alcohol increase negative effects on fecundity and fetal growth [22]. The use of FANS and antidepressant can increase risk of spontaneous abortion. The FANS must be suspended within 32 weeks of pregnancy due to the risk of premature closure of the Botallo duct and alterations in the renal flow of the fetus [23]. Frequently, the advanced maternal age can contribute with the RPL: it affects 12 % of cases between 20 and 24 years. The risk start to increase at age 35 [24] and in 26 % of cases over 40 years [8].This is related to the fact that advancing maternal age increase the probability that the fetus undergoes chromosomal alterations. This risk is linked to female gametogenesis. At birth, each woman has all the oocytes which she will have during her life. A result of this is that with advancing maternal age the sites are exposed to mutagens, with a consequent increase in the risk of aneuploidies, due to alterations in the meiotic division [8, 17, 25].

Folate

In the first trimester, deficiencies of folic acid or vitamin B9 increase the risk of fetal malformation, as neural tube defects (NTD) associated with spina bifida or anencephaly. In general, a folate deficiency can result in adverse effects such as intrauterine growth retardation, premature birth, placental lesions and in worst cases lead to abortion [20, 26]. Folates are indispensable nutrients that human body used for one-carbon biosynthesis and epigenetic processes. Folates are derived completely from dietary sources, mainly from the consumption of green vegetables or citrus fruit, as spinaches, broccoli, asparagus, lettuce, oranges, lemons, strawberries and kiwi, also from cereals, beans and liver. Chemically, a folates molecule consist of three distinct chemical part linked together (fig 1.4): i) a pterin (2-amino-4-hydroxy-pteridine) (in blue colour) heterocyclic ring is linked by a methylene bridge to ii) a p-aminobenzoyl group (colour red) that in turn is bonded through an amide linkage to either iii) glutamic acid or poly-glutamate (colour black). After intestinal absorption, folate metabolism requires reduction and methylation into the liver to form 5-methylte-trahydrofolate (5-methylTHF), release into the blood and cellular uptake; then it can be used for the synthesis of DNA and RNA precursors or for the conversion of homocysteine (Hcy) to methionine, which is then used to form the main DNA methylating agent S-adenosylmethionine (SAM). A deficiency in cellular folates results in aberrant DNA methylation, point mutations, chromosome breakage

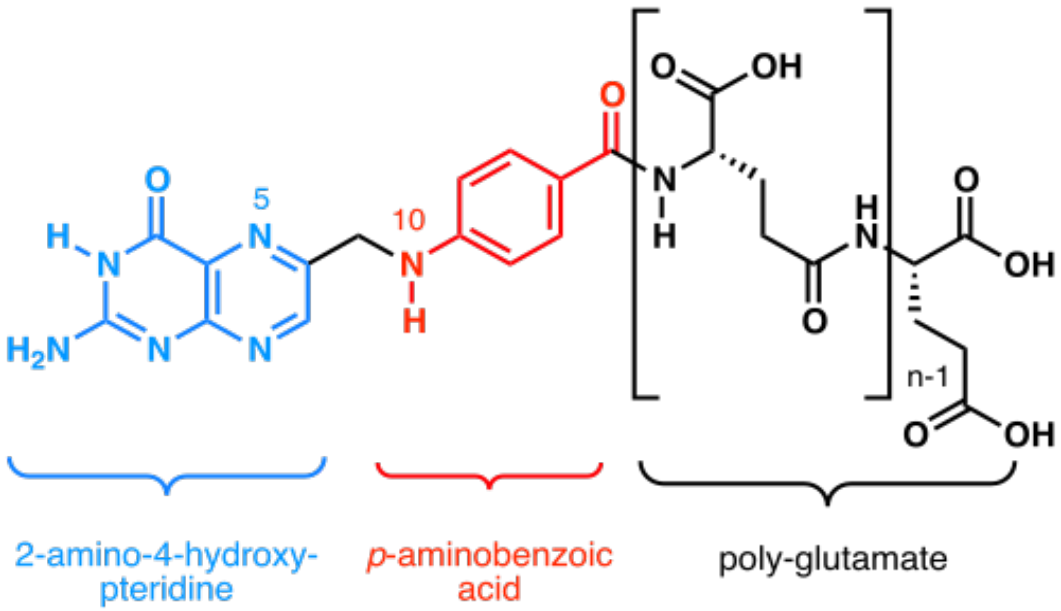

FIGURE 1.4: **Structure of folate molecule**

and increased frequency of micronuclei (MN), defective chromosome recombination and aneuploidy [27].

1.3.8 Paternal causes

Male gametes contribute one-half of the genomic content to the embryo. In particular, recent studies indicate that the male contributes to RPL with sperm quality and advanced paternal age [17, 28].

Sperm quality

Spermatozoa are highly susceptible to damage, which results in poor sperm quality. The exposure to *reactive oxygen species* (**ROS**), produced by leukocytes in sperm, damage the quality of the DNA. Also, structural chromosomal problems in sperm could contribute to lower concentrations of sperm cells and male infertility. In the work of *Allison et al., 2009* [29], it is shown that the elevated levels of circulating homocysteine can be toxic to spermatozoa DNA. A strict connection was found between paternal homozygous *MTHFR* C677T mutations and the increase of RPL. Heterozygosity for this mutation seems not to increase the risk. A higher miscarriage rates is seen in men with reciprocal translocations compared to men with Robertsonian translocations. This situation increases the rate of sperms aneuploidy for chromosomes 13,18,21, X, Y, detected by FISH in the spermatozoa of men with a history of unexplained RPL, especially with an increase of paternal age [17, 29, 30].

Advanced paternal age

Paternal age increases the chance of mutation or aneuploid during the maturation of male germ cells and decreases the sperm quality, in particular the motility. The age is also correlated with a decrease in conception rate, an increase of miscarriage rate and an increase in autosomal dominant diseases in the progeny. The advanced paternal age is associated with gain-of-function mutations within sperm that have damaging effects on the embryo. These gain-of-function mutations occur in three

genes, that encode for tyrosine kinase receptors: *FGFR2*, *FGFR3* and *RET*. Each of these mutations is associated with an autosomal dominant effect on the offspring. The effects of advanced age may have an impact on implantation, placental proliferation and placenta quality. These placenta effects may also play a role in RPL in old men, as the quantity of DNA damage in the sperms of men aged 36-57 is three times that of men 35 years old [17, 29–31].

1.4 Genetic variant with risk of RPL

In this study, I consider nine SNPs linked to the risk of miscarriage, summarized in table 1.1 with major and minor allele, ancestral allele and the chromosomes position. These SNPs were selected after extensively searching the literature. The information about SNPs were based on Human GRCh38.p13 assembly. In the table 1.2 were described a comparison of different frequency alleles and derived alleles for each gene in the general population and in the Italian population and its frequencies in both cases and control of the study. The information were based on 1000 Genome Browers.

Gene	*rsID*	*Major/minor allele*	*Ancestral allele*	*Derived allele*	*Chromosome position*	*Role*
MTHFR	rs1801133	C/T	C	C|C, C|T, T|T	1:11796321	Metabolism folic acid [32]
MTHFR	rs1801131	A/C	A	A|A, A|C, C|C	1:11794419	Metabolism folic acid[33]
HLA-G	rs371194629	14bp Del/Ins	NA	D|D, D|I, I|I	6:29830804	Immune tolerance in pregnancy [34][35][36]
ANXA5	rs1050606	A/C	C	A|A, A|C, C|C	4:121696891	Ca2+ dependent placenta anticoagulant protein [37] [38]
NKG2D	rs2617170	A/G	A	G|G, G|A, A|A	12:10408358	Detection and elimination of transformed and infected cells [39]
IL-10	rs1800871	A/G	G	G|G, G|A, A|A	1:206773289	Cytokines in the cascade of immune signalling [40] [41]
CTLA-4	rs231775	A/G	A	G|G, G|A, A|A	2:203867991	Member of immunoglobulin superfamily [42]
SMTH1	rs1979277	C/T	C	C|C, C|T, T|T	17:18328782	Metabolism folic acid [43]
PLK4	rs2305957	A/G	G	G|G, G|A, A|A	4:127811771	Regulates centriole duplication during cell cycle [44][45]

TABLE 1.1: Pannel of candidate variants genes analyzed

rsID	*Frequency of allele in general population*	*Frequency of allele in Italian Population (1000 Genome Project)*	*Frequency of derived allele in general population*	*Frequency of derived allele in Italian population*
rs1801133	C:0.755, T:0.245	C: 0.533, T: 0.467	C|C:0.591 T|T: 0.082, C|T: 0.327	C|C: 0.406, T|T: 0.135, C|T: 0.459
rs1801131	A:0.751, C:0.249	A: 0.687, C: 0.313	A|A: 0.573, C|C: 0.072, C|A: 0.355	A|A: 0.486, C|C: 0.112, C|A: 0.402
rs371194629	D:0.606, I:0.394	D: 0.612, I: 0.388	D|D: 0.367, D|I: 0.477, I|I: 0.156	D|D: 0.364, D|I: 0.495, I|I: 0.140
rs1050606	A:0.544, C:0.456	A: 0.439, C: 0.561	A|A: 0.315, A|C: 0.458, C|C: 0.227	A|A: 0.206, A|C: 0.467, C|C: 0.327
rs2617170	A:0.443, G:0.557	A: 0.374, G: 0.626	A|A: 0.203, G|G: 0.317, G|A: 0.480	A|A: 0.140, G|G: 0.393, G|A: 0.467
rs1800871	A:0.435, G:0.565	A:0.294, G:0.706	A|A: 0.216, A|G: 0.438, G|G: 0.346	A|A: 0.084, A|G: 0.421, G|G: 0.495
rs231775	A:0.573, G:0.427	A: 0.715, G: 0.285	A|A: 0.342, A|G: 0.461, G|G: 0.197	A|A: 0.486, A|G: 0.458, G|G: 0.056
rs1979277	C:0.772, A:0.228	C: 0.762, T: 0.238	T|T: 0.069, T|C: 0.318, C|C: 0.613	T|T: 0.589, T|C: 0.065, C|C: 0.346
rs2305957	A:0.303, G:0.697	A: 0.234, G: 0.766	A|A: 0.096, A|G: 0.413, G|G: 0.491	A|A: 0.075, A|G: 0.318, G|G: 0.607

TABLE 1.2: Description of frequencies allele in General's and Italian populations

1.4.1 *MTHFR* rs1801133 (c.677C>T) and rs1801131(c.1298A>C)

The enzyme *MTHFR* plays a decisive role in the folate metabolism pathway and regulates the intracellular folate pool for synthesis and methylation of DNA [46]. Previous studies have shown that the *MTHFR* gene is located on chromosome 1 at the end of the short arm (1p36.6) and is 2.2 KB in length with a total of 11 exons [32, 47]. Several SNPs in the *MTHFR* gene were reported, the c. 677C>T (rs1801133) and c.1298A>C (rs1801131) are the most studied functional ones. The c. 677C>T transition in exon 4 results in amino acid substitution of alanine to Valine (p. Ala222Val) which is linked to decreased thermal stability and reduced activity of *MTHFR* enzyme [48].The minor allele frequency (MAF) of rs1801133 for allele T was 0.31 (79177/251468, GnomAD-exome, dbSNP). Instead, *MTHFR* c.1298A>C, in exon 7, determines a glutamate to alanine substitution (p.Glu429Ala) that could increase the serum

folate levels [49].The minor allele frequency (MAF) of rs1801131 for allele C was 0.28 (72672/251462, GnomAD-exome, dbSNP). Numerous studies have investigated the association between c. 677C>T and c. 1298A>C polymorphisms in women and the risk of Recurrent Pregnancy Loss (RPL), but the effects remain controversial [33, 50, 51]

1.4.2 *HLA-G* 14bp INDEL

The *HLA-G* 14 base pair insertion/deletion is present within 3'UTR region of exon 8. In pregnancy, the KIR2DL4 receptor and its ligand HLA-G is considered important for fetal-maternal immune tolerance and successful pregnancy. In fact, the expression of *HLA-G* is determined by the kind of trophoblast and stage of pregnancy progression. *HLA-G* membrane-bound molecules (of maternal and paternal origin) are presented by all extravillous trophoblast subpopulations [34, 35].

1.4.3 *ANXA5* rs1050606

The gene *ANXA5* is located on chromosome 4 in long harms 4.q27. The minor allele frequency (MAF) of rs1050606 for allele C was 0.49 (15333/31242, GnomAD, dbSNP). The protein Annexin A5 is a 35 kDa protein encoded by *ANXA5* gene and it is a placental anticoagulant protein. It is abundantly expressed in placenta and its expression seems to be decreased in the presence of antiphospholipid antibodies (APL) and the encoding gene possesses a complex promoter region that is subject to intricate regulation. Annexin A5 functions as an inhibitor of coagulation binding the anionic phospholipids exposed on the surface of platelets. In the study of *Hayashi et al., 2013* [37], the SNP rs1050606 is a risk factor of RPL. This evidence was observed in a previous written report on the German population [38].

1.4.4 *NKG2D* rs2617170 (c.311A>G)

The gene *NKG2D* (c.311A>G) rs2617170 is located on chromosome 12, in position 12p13.2.The minor allele frequency (MAF) of rs2617170 for allele A was 0.37(87972/236222, GnomAD-exome, dbSNP). This gene is a member of group of genes. Their expression is mostly in natural killer (NK) cells. These types of cell have important function in the maintenance of pregnancy. They balance two effects: on the maternal side, the NK cell-mediated cytotoxicity limits the excessive trophoblast invasion. On the other hand, decidual NK cells produce a form of cytokines that induce angiogenesis such as vascular endothelial for contributing to the uterine vascular remodeling. A dysregulation of this cytotoxic activity can induce RPL. In the previous study of *Hizem et al., 2014*[39], it is shown that the SNP rs2617170 influences the result of pregnancy. In particular, the TT genotype in this SNP confers a high cytotoxic activity.

1.4.5 *IL-10* rs1800871(c.819C>T)

The gene *IL-10* (c. 819C>T) rs1800871 is located on chromosome 1 at position 1q32.1. The minor allele frequency (MAF) of rs1800871 for allele A was 0.32 (41249/125568, TOPMED, dbSNP). The interleukin-10 (IL-10) is a key cytokine, that plays a critical role in the maintenance of maternal–fetal tolerance of its pleiotropic activities. High levels of IL-10 were detected in normal pregnant women while reduced IL-10 levels were present in women with RPL. It, likewise, can cause preterm birth, and preeclampsia. In particular, the rs1800871T alleles were found to be related to low production of IL10. A meta-analysis of *Su et al. 2016* [52] and a study of *Cochery et*

al., 2009 [40] confirmed the correlation of this SNP with RPL but the result remains controversial. In fact, a study of *Qaddourah 2014* [41], and a meta-analysis of *Gu et all., 2016* [53] shows opposite effects.

1.4.6 *CTLA-4* rs231775 (c.49A>G)

The gene *CTLA-4* (c.49A>G) rs231775 is located on chromosome 2 at position 2q33.2, for 6.2 KB and consist of four exons and three introns. The minor allele frequency (MAF) of rs231775 for allele G was 0.40 (50382/125568, TOPMED, dbSNP). Cytotoxic T-lymphocyte antigen 4 (CTLA4) is expressed constitutively on regulatory T cells and promotes maternal fetal tolerance. Furthermore, in vivo studies have demonstrated that expression of regulatory T cells and *CTLA-4* in peripheral and decidual lymphocytes is down-regulated in human miscarriages. The *CTLA-4* rs231775 SNP causes substitution with Threonine (Thr) in position 17 with an Alanine (Ala), in the leading peptide of the *CTLA-4* receptor and his, due to the recombinant CTLA417Ala had a significantly stronger power to inhibit T- cell proliferation and activation compared with its counter- part CTLA417Thr. These findings show that the 17Thr>17Ala change in *CTLA-4* results in stronger CTLA4- triggered inhibition of T-cell proliferation and activation. In a study of *Zhang et al., 2018*[54] has found an association of the *CTLA-4* tag-SNPs rs231775 with an increased risk for RPL [42].

1.4.7 *SMTH1* rs1979277 (c.1420C>T)

The gene *SMTH1* (c.1420C>T) rs1979277 is located on chromosome 17 at position 17p11.2.The minor allele frequency (MAF) of rs197927 for allele G was 0.12 (16217/125568, TOPMED, dbSNP) This gene produces a serine hydroxyl-methyltran-ferase 1 involved in the cycle of folic acid: in particular, SMTH uses serine as the single-carbon donor for the conversion of THF into 5,10-methyleneTHF, that can be used for thymidylate synthesis in the reaction catalyzed by TYMS that produces dTMP and DHF, which is then reduced back to THF by DHFR. Polymorphisms of SMTH were investigated only in one case-control study of its relation with a neural tube defect (NTD). Especially, the *SMTH1* 1420C>T polymorphism (rs1979277), that results in Leu474Phe replacement and impairs SMTH nuclear transport and subsequent thymidilate synthesis was investigated. Genotype 1420CC and 1420CT resulted in a decreased maternal risk of birth of a child with DS with respect to the 1420TT genotype [55–57].

1.4.8 *PLK4* rs2305957

The gene *PLK4* rs2305957 is located on chromosome 4 at position 4q28.1. The minor allele frequency (MAF) of rs2305957 for allele A was 0.33 (41498/125568, TOPMED, dbSNP). In studies of *McCoy et al., 2015*[45] and *Zhang et al., 2017*[38], Polo-Line Kinase 4 (*PLK4*) was observed associated with miscarriage, in correlation with maternal genotype. In particular, this gene seems to play a role in the centriole duplication and can alter the result of mitosis through a dysregulation of centriole areas and contribute to mitotic-origin aneuploidy risk during human early embryo development. In a previous study, the genotype AA in mothers is correlated with higher risk of RPL.

1.5 Epigenetics

Epigenetic refers to the study of alterations in gene function that are mitotically and/or meiotically heritable and that do not entail a change in the DNA sequence. Epigenetic mechanisms are essential for development and differentiation, but can be disrupted by exogenous factors. The most investigated epigenetic mechanisms include DNA methylation, Histone modifications and RNA-mediated silencing [58].

1.5.1 DNA methylation

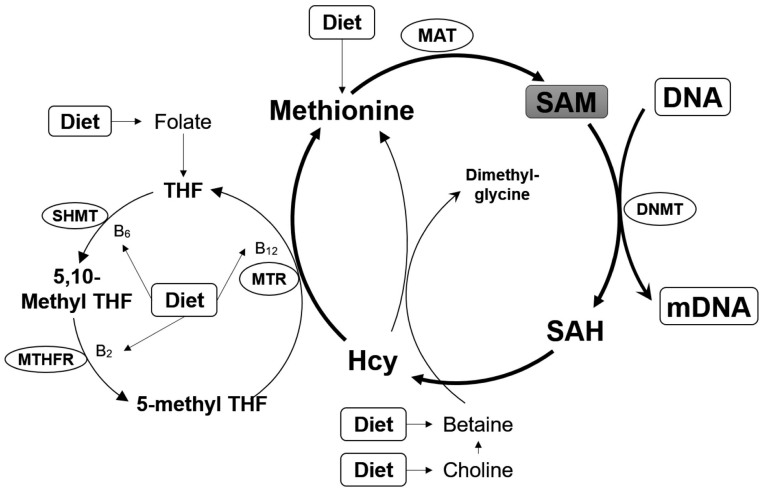

FIGURE 1.5: **Cycle of DNA methylation**. **MAT** = methionine adenosyltransferase; **SAM** = S-adenosyl methionine; **SHMT** = serine hydroxymethyltransferase; **THF** = tetrahydrofolate; **DNMT** = DNA methyltransferase; **MTR** = 5-methyltetragydrofolate-homocysteine methyltransferase; **MTHFR** = methylentetrahydrofolate reductase; **Hcy** = homocysteine; **SAH** = S-adenosylhomocysteine; **mDNA** = methylated DNA. Revaised from *Naifeng Zhang, 2015*

The most common epigenetic change is DNA methylation. DNA cytosine methylation (**5mC**) is chemically relatively stable, but it is a dynamic epigenetic modification regulated by *DNA methyltransferases* (**DNMTs**)(fig 1.5). These modifications involve the covalent addition of a methyl group to the 5-position of cytosine with *S-adenosyl methionine* (**SAM**) as the methyl donor. DNA methylation is almost entirely restricted to CpG dinucleotides clustered within the gene promoter and in repeated elements as *Long Interspersed Elements* (**LINE-1**). In general, hypermethylation of the DNA promoter inactivates the gene expression and hypomethylation activates the expression [58]. In somatic cells, DNA methylation patterns are stably inherited through mitosis. However, variations in the amount and distribution of methylated cytidines have been described in somatic cells during differentiation and ageing and a direct link between aberrant DNA methylation patterns and tumorigenic processes has been established. In humans, the fluctuation of rates of DNA methylation between normal tissues has also been described. The lowest level is found in placenta where 3,2% of the cytosines are methylated, and the highest level (4,26%) is found in the brain [59]. Scientific evidence increasingly suggests that exposures during the intrauterine period can increase the danger of developing disease in later life. Prenatal exposure to maternal tobacco smoking is linked with lower pulmonary function and increased asthmatic symptoms in childhood. Changes in

the epigenome are one mechanism by which prenatal exposures affect disease risk later in life. DNA methylation is the most studied type of epigenetic mark. In mammals, there are two developmental periods, in germ cells and in preimplantation embryos, in which methylation patterns are reprogrammed genome wide, generating cells with a broad developmental potential. Mostly, this period involves demethylation and later remethylation in a cell or a tissue specific manner. Therefore, prenatal tobacco smoke exposure in early pregnancy may have important and lasting effects on DNA methylation and consequently influence gene expression and disease phenotypes across the life course [60]. Global hypomethylation of genomic DNA and hypermethylation of gene promoter regions occur simultaneously in an extensive variety of malignancies. Folate is an important precursor of one-carbon units required for DNA methylation. Thus, folate metabolism has been suggested to influence epigenetic alterations in cancer. High folic acid might contribute to the maintenance of global methylation through an equal supply of one-carbon units for the methylation machinery, thereby stabilizing the genome [61].

Long Interspersed Nuclear Element 1 (LINE-1)

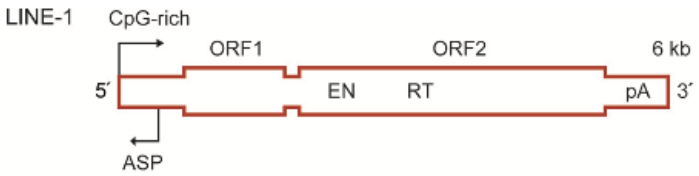

FIGURE 1.6: **LINE-1 structure**.
ASP=antisense promoter; **EN**=endonuclease; **RT**=reverse transcriptase; **pA**=poly(A) tail. Revaised from *Ardeljan et al., 2017*.

The LINE-1 are considered the most common repetitive elements of interspersed DNA repeats and are moderately CpG rich and most heavily methylated. Because of their high genome dissemination, LINE-1 methylation is a surrogate marker of global DNA methylation level [62]. In literature, hypomethylation of these repeated sequences has been related to pathological processes, including tumorigenesis, abnormal placental function, birth defects, aging and other chronic diseases. LINE-1 representing about 18-20% of the total human genome with up to 500 000 copies [63]. Although the majority of LINE-1s are 5'-truncated and transcriptionally inactive, about 150 full-length LINE-1s with two intact *open reading frames* (**ORF**), called ORF1 and ORF2, and an additional 100 LINE-1s with only intact ORF2, have been set up to be transcriptionally active (fig 1.6). These full-length L1 elements are approximately 6 kb in length and comprise a 5'UTR containing an RNA Pol II promoter, two open reading frames (ORFs) and 3'UTR that terminates with a *poly* (A) *tail* and contains a poly-adenylation signal. The ORFs encode activities required for transposition: ORF1 is a nucleic acid chaperone, and ORF2 has endonuclease and reverse transcriptase activities. The activities of LINE-1 and other transposable elements have effects on cellular function, genome stability and development. Also L1 transcription were previously reported to be active mainly in the germ line and transformed cells [62, 64]. ORF2 endonuclease activity can serve as a source of genome instability by generating *DNA double-strand breaks* (**DSBs**) uncoupled from successful transposition. L1 insertion can also bypass the need for the ORF2 endonuclease by hijacking endogenous DSBs. Although this endonuclease-independent insertion may facilitate DSB repair, it also holds the potential to interfere with endogenous repair pathways

as well as causing potentially deleterious insertions. Both endonuclease- dependent and -independent insertion has been linked up with local DNA rearrangements including deletions, duplications, inversions, and translocations. L1 itself is a source of unstable microsatellite sequences and insertion generates additional poly (A) microsatellites de novo. L1 insertion also influences transcription, both positively and negatively. An antisense promoter in the 5 UTR can drive transcription of flanking cellular genes and appears to be broadly utilized in tissue-specific gene regulation. Oppositely, as targets of heterochromatin formation, L1 elements can locally silence gene expression. Intragenic L1 elements can impede transcription and serve as sense and antisense promoters, alternative splice junctions, and termination sites. More generally, L1 expression is connected with cellular responses such as activation of the innate-immune response and the DNA-damage response and their downstream effects such as cell cycle arrest, senescence, and apoptosis [64]. Similar to locus-specific DNA methylation changes, global DNA methylation changes also represent a biodosimeter of lifelong environmental exposures [63, 65].

1.6 Influence of MTHFR Gene on Global Methylation in RPL

Smoking, nutrition, and age are known to influence methylation patterns throughout the life course. This evidence has been documented in adults, but in embryo during pregnancy and humans at birth, are poorly understood. Emerging evidence suggests that advanced parental age, maternal pregestational BMI can influence DNA methylation but it is probably that many more factors modulate the infant methylome including both environmental and genetic components. This factors can modulate the one-carbon metabolism (1-CM), modifying the provision of of methyl groups via S-adenosyl methionine (SAM) for DNA methylation. That factors also can modulate the folate cycle and methionine pathway, that has fundamental roles in a variety of essential biochemical processes pertinent to healthy development. *MTHFR* rs1801133 has an important functions in normal one-carbon metabolism. The consequences of perturbed 1-CM, include reduced re-methylation of total homocysteine (tHcy) into methionine, leading to elevated plasma tHcy levels, as well as a reduction in cellular methylation potential. This has as consequence to increase risk of adverse outcome, including cardiovascular and cerebrovascular disease, colorectal cancer. Also can cause adverse outcome during pregnancy. In fact, the DNA methylation plays an important role in regulating the growth of the trophoblast cell lines so any alterations at this level can cause RPL. Given the association between DNA methylation patterns and gene expression, it is probable that aberrant methylation patterns at birth may predispose individuals to higher disease risk later in life via developmental programming (as obesity). It is known that *MTHFR* rs1801133 influences global DNA methylation status through a direct interaction with folate status. In particular, the interaction between the T/T genotype in *MTHFR* and low folate status cause a reduction of 50% in the level of genomic DNA methylation in this group. The homozygosity of *MTHFR* in combination with low folate status, is also known to predispose people to high plasma levels of fasting tHcy. [57, 66, 67].

1.7 Project GREP

Part of the samples obtained for the project of my thesis, where contributed to another project, called "*Genomics of REcurrent Pregnancy loss (**GREP**)*", lead by the Institute of Genetics and Biophysics of the National Research Council (**CNR**) in collaboration with the University of Napoli Federico II, and several other partners in Italy, United Kingdom, Malaysia, Pakistan, and Australia. GREP aims to identify genetic variants likely to cause RPL not seen by current diagnostic tools (mainly comparative genomic hybridization), either because of the size or because they are located in non-coding regions not considered in medical diagnostics. The main objective of GREP is to build a predictive model that integrates genomic variation and functional annotations, based on the analysis of whole-genome sequences of miscarried embryos. The goal of this study is to improve prenatal diagnosis, in particular diagnosis of RPL and increase the rate of success of pre-implantation genetic testing.

The steps of the GREP pilot project are (Figure 1.7):

1. The collection of samples by the University of Ferrara during 2016-2018. I participated to the first step through the collection and processing of samples of CV. In particular, I built a biobank of DNA samples obtained after separation of chorionic villi from maternal material.

2. Screening for euploid samples to be sequenced.

3. Data anaysis of whole-genome sequence to determine genetic causes of PL.

4. Validation.

FIGURE 1.7: **Pilot project phases**

Chapter 2

Aims and Hypothese

2.1 Aim

My PhD project is a part of the Italian project called *"Silent Intrauterine infections and early pregnancy loss"* **G.R Reg. E/R (PRUA1GR-2013 00000220**. This dissertation aims to understand if the genetic and epigenetic profile of the women and the product of conception play a role in determining the risk of Recurrent Pregnancy Loss (RPL). For the first time, the analyses focus on gDNA in both the mother's blood and chorionic villi. As a control reference, voluntary termination of pregnancy (VTP) cases were used.

2.2 Hypotheses

The analyses focus on three hypotheses:

1. Understand if the genetic profile in the fertile women and the relative product of conception, entirely or in association with exposure to specific environmental factors during the periconceptional period, play a role in determining the risk of recurrent spontaneous abortion. To examine this I identified a multiple candidate's genetic variation at the level of the embryo-maternal interface to determine susceptibility to RPL in literature.

2. Clarify if there is a relationship between alterations in the levels of global genomic methylation and the risk of recurrent spontaneous abortion, focusing not only on the pregnant women, but also on the relative product of conception. In addition, I examined the association with embryo's sex and age, pregnancy parameters or exposure to environmental risk factors.

3. Detect association between global genomic methylation (LINE-1 sequences) and MTHFR gene variants in RPL cases and VTP.

2.3 Other contribute

My PhD project also contributed to the *Genomics of REcurrent Pregnancy loss (GREP)* project lead by the CNR in Naples, whose aim is:

[1] Identify genetic variants likely to cause RPL that are not seen by current diagnostic tools either because of their very small size or because they lie in non-coding

regulatory regions of the genome that are not routinely sequenced. To achieve this, GREP aims to build a predictive model to integrate high-coverage whole-genome sequence data with information on functional annotations and gene networks relevant to embryonic development.

Chapter 3

Material and methods

3.1 Data and samples collection

In this study were analyzed data from 156 women that underwent Voluntary Termination of Pregnancy (VTP) and 91 recurrent pregnancy loss **(RPL)**, 40 first pregnancy loss **(FPL)** in a total of 131 women experiencing Pregnancy Loss (PL). The samples were provided by the Unit of Obstetrics and Gynecology at The Sant'Anna University Hospital of Ferrara, Italy, from 2016 to 2020. The inclusion criteria were: patient's age in the range 18–42 years; gestational age within the first 12 weeks; and the cohort of VTP pregnancy was composed of females that have a volutary miscarriage until 90 days according to the Italian law, named *Bill 194, Article 4*. The exclusion criteria were: patients positive for infectious agents/diseases, such as HIV, hepatitis B virus, hepatitis C virus, syphilis, during the year before the sample collection; the presence of congenital or acquired immune deficiency syndrome/diseases, or immunosuppressive therapies during the year before the sample collection. And causes of recurrent pregnancy loss as genetics, severe uterine or hormonal dysregulation, and use of teratogenic drugs. The samples were composed by the products of the conception and the relative peripheral blood of mothers. The project **G.R Reg. E/R (PRUA1GR-2013 00000220)** *"Silent Intrauterine infections and early pregnancy loss"* was approved by Local Ethical Committee **(CE/FE 170475)** and was carried out in compliance with the Helsinki Declaration. All participants provided written informed consents before recruitment. Data set consists of medical data from interviews on 283 cases. In the questionnaire, there is information about mother's age, Body Mass Index (BMI), gestational age, menarche, geographical origin, educational level, lifestyle factors (smoking, alcohol consumption, drug consumption), chronic disease, folic acid intake, obstetric and gynecologic history of mother and sister of the patient. All the data were anonymized right after data and sample collection.

3.1.1 Chorionic Villi

The material of conception with relative media arrives in the laboratory in sterilized Falcon tubes of 15 ml. The selection of chorionic villi from the maternal material must be done under a laminar flow hood, as shows in figure 3.1, to preserve the sterility of the sample. The most challenging step of the protocol consists of distinguishing and separating exclusively the villi from the rest of the maternal material like placenta, and decidua. As shows in figure 3.2, on the left the petri dish contains maternal material and on the right the petri dish with the selected chorionic villi (CV). The dissection takes place under the hood using a stereomicroscope (e.g. Leica Microsystems Srl, All Microscopy and Histology, Milan, I-20142 Italy). The operator should recognize chorionic villi at the stereomicroscope and separate them from the decidua using scalpels. The villi can be stored at -20 °C in vials, for a few

months or at 4°C in RPMI media for not more than a week before proceeding to DNA extraction.

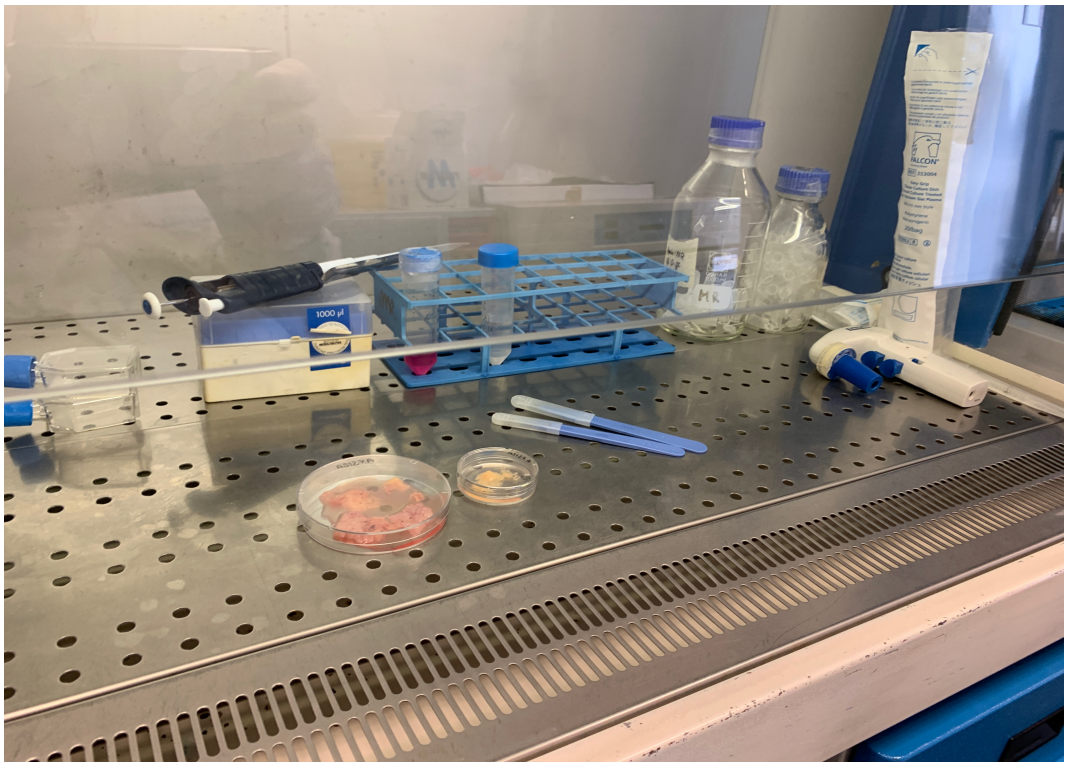

FIGURE 3.1: **Laminar flow hood** in laboratory of medical genetic, to preserve the sterility of the sample

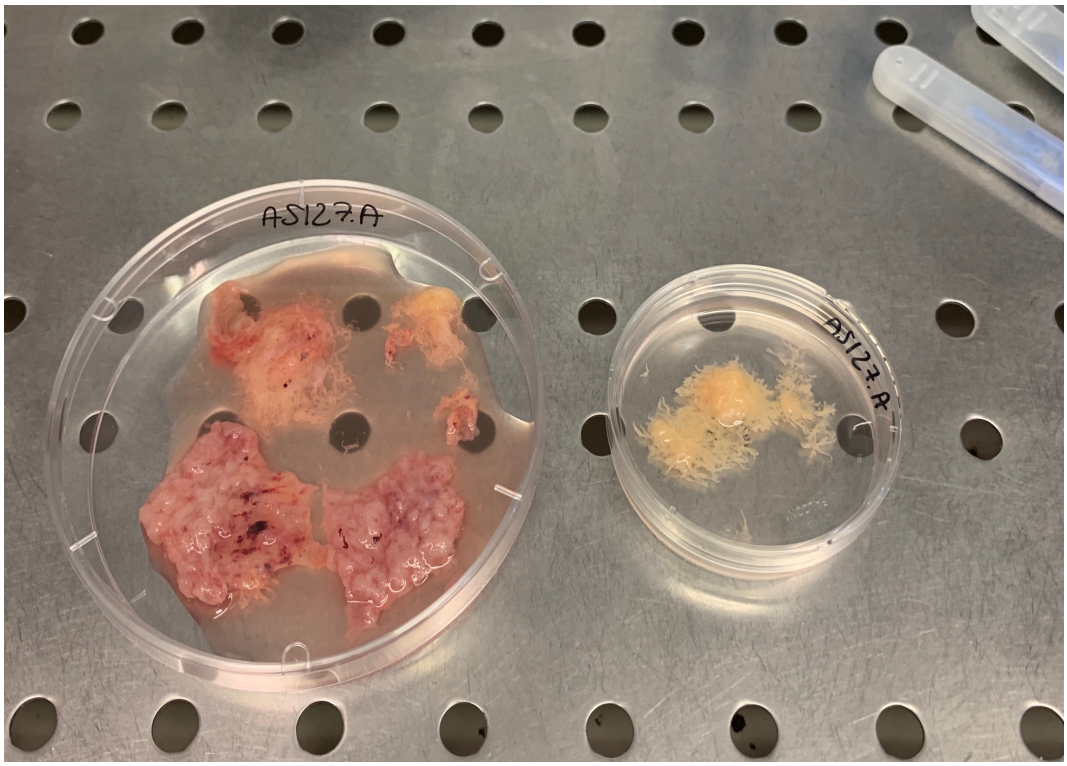

FIGURE 3.2: **Petri dish for separation of miscarriage**. On the left petri dish with maternal material mix with chorionic villi and on the right chorionic villi

3.1.2 Peripheral blood samples from the mothers

The peripheral blood from mothers was collected in tubes with EDTA at the same time of collection of the miscarriage material.

3.2 Genomic DNA extraction, quantification and storage

Genomic DNA (gDNA) was extracted from chorionic villi dissected from abortion tissue specimens using three different kit, in particular two types of resin called Instagene TM Matrix (Bio-Rad) and Nucleon BACC1 (GE Healthcare UK). One membrane was used, the QIAamp DNA Blood Mini Kit (Qiagen) in according to manufacturer protocols *(QIAamp DNA Mini and Blood Mini Handbook 05/2016. Instruction Manual, InstaGene^{TM} Matrix, LIT544 Rev G, RPN8501-PL Rev C 04/2008)*.

The gDNA obtained from white blood cell (WBC) of peripheral blood samples from mothers was instead extracted using the Nucleon BACC1 (GE Healthcare UK), according to manufacturer protocols (*RPN8501-PL Rev C 04/2008*).

All genomic DNA was quantified using Qubit® dsDNA BR Assay Kit (Life technologies Oregon, USA). After each gDNA was inserted in Matrix 2D-Barcoded (Thermo Fisher Scientific) and they made possible set up a DNA-Biobank located in -80°C freezer equipped with Access Key and constant monitoring of use and function conditions. Working conditions took place using genomic DNA at concentrations of 10 ng/μL or 1ng/μL, depending on methodology, in order to normalize results.

3.3 Detection of chromosomal anomalies

The revelation of chromosomal anomalies was done using four techniques: classical cytogenetic, quantitative PCR (qfPCR), comparative genomic hybridization (array-CGH), and shallow sequencing of randomly amplified genomic fragments.

3.3.1 Classical cytogenetic techniques

Samples for which was possible to obtain cell cultures were analyzed with metaphase spread slides and Q-banded to obtain the karyotype. The basic cytogenetic method involves chromosome harvest, slide preparation, staining and banding of the chromosomes, analysis of chromosome numbers in the population, and analysis of the banding patterns. Two different techniques were used to obtain the karyotypes of the embryos: the protocol of direct technique, after 48 hours, described in the paper of *Shulman L.P. et all.,1990* [68] and the indirect technique representing by conventional tissue culture techniques [69] were used to assess the genetic stability of chorionic villi cultures. The nomenclature of band assignment and chromosome aberrations is summarized by the International System for Human Cytogenetic Nomenclature (ISCN).

3.3.2 qfPCR, arrayCGH, low-coverage sequencing

This part was performed by our collaborators in Naples and in Vicenza. In brief, after extraction of DNA from chorionic villi of the product of conception using standard protocols, a rapid screening of sex and numerical anomalies for chromosomes 13, 15, 16, 18, 21, 22 and X was carried out with the miscarriage DNA samples performing quantitative PCR assays. Samples that resulted euploid for these chromosomes were further analyzed by Comparative Genomic Hybridization (Agilent SurePrint G3) and low-coverage sequencing of randomly amplified genomic regions.

3.4 Genotyping

The genomic DNA obtained from WBC and abortion tissue specimens were used for the genotyping of a panel of candidate gene variants with 5'- nucleus Real-Time PCR assay using allele-specific TaqMan probes. PCR conditions for all reaction were as follows: 50°C for 2 min, 95°C for 10 min and (95°C for 15 s, 60°C for 1 min) x 50 cycles. The instrument used for reading plate was Applied BioSystems ABI PRISM 7300 (Applied BioSystems, Foster City, CA). All the panel of candidate genes is listed in the table 1.1. The gene HLA-G 14 bp was analyzed by a polymerase chain reaction (PCR) sequence-specific primer method (PCR-PAGE) [70]. The amplification was performed by PCR with a GeneAmp PCR System 2700 thermal cycler (Applied Biosystems, Foster City, CA) in a 25 μL reaction mixture containing 100ng of genomic DNA, 10XPCR buffer, 50mM MgCl2, 10mM dNTPs, 20pmol of each primer and 1U of Taq polymerase (Invitrogen Co., Carlsbad, Ca). The PCR conditions comprised initial denaturation at 94°C for 2 min, followed by 10 cycles of denaturation at 94°C for 15 s, annealing at 64°C for 30 s and extension at 72°C for 30 s, then 25 cycles of denaturation at 94°C for 15 s, annealing at 63°C for 30 s and extension at 72°C for 30 s, and final extension at 72°C for 5 min. The purified PCR products size were analyzed using an 8% polyacrylamide gel. The product size was 224 bp for

Ins/Ins (I/I) and 210 bp for Del/Del (D/D) and both 224 bp and 210 bp for Del/Ins (D/I) genotypes. The PCR products were visualized using silver staining.

3.5 DNA methylation

In this study, I examined the DNA methylation in my samples of gDNA, both in WBC and CV (RPL and VTP). The analysis of methylation in LINE-1 sequences were done following the steps described below.

3.5.1 Pyrosequencing

This protocol is in accord to study of Khan MFJ. et al., 2017 [71]. The genomic DNA samples were next bisulfite-converted using EZ DNA Methylation-Lightning Kit (Zymo Research, Irvine, CA, USA), according to manufacturer's instructions (Instruction manual CAT. No. D5030T, ver 1.0.4). After that, the CpG islands of LINE-1 sequences were amplified by PCR. The PCR reactions were performed with a total volume of 25 μL containing: 10X PCR Buffer, 50mM MgCl2, 2.5 mM dNTPs, 10 pM Reverse Primer LINE-1, 10 pM Forward Primer LINE-1, 5U Taq-Polymerase and 2.5 μL of Bisulfite DNA. The cycling profile was composed of 27 cycles of 94 °C for 15 sec, 60°C for 30 sec and 72°C for 30 sec, followed by 72°C for 2 min. The amplicon of 147 bp was analyzed on 8% polyacrylamide gel using silver staining (parameter by running: 150V for 45 min). The products of PCR were sequenced by pyrosequencing using the PyroMark Q96 system (Qiagen), according to manufacturer's instructions (PyroMark Q96 ID User Manual 01/2016).

The sequence of LINE-1 was the follows (figure 3.3):

During pyrosequencing run the dispensation order was (colored in red):

ACTCAGTGTTCAGTCAGTTAGTCTG

It is necessary insert the sequence to analyze:

TTYGTGGTGYGTYGTTTTTTAAGTYG.

The average of LINE-1 methylation level was calculated as of the mean of the proportions of C (%) at the 4 CpG sites analyzed, which were located at positions +306, +318, +321 and +328 (positions of the corresponding Guanine in the forward DNA strand, in relation to the first nucleotide base of the consensus promoter sequence in Genebank sequence no. X58075, position 305-331), represented in the red rectangles.

The sequence of the primers used for this analysis (IDT, Tema ricerca, Italy) and was obtained from study of *Delaney et al.*, 2015 [72]:

Bio-Reverse LINE-1
5'-/5Biosg/AAAATCAAAAAATTCCCTTTC-3'

Forward LINE-1
5'-TTTTGAGTTAGGTGTGGGATATA-3'

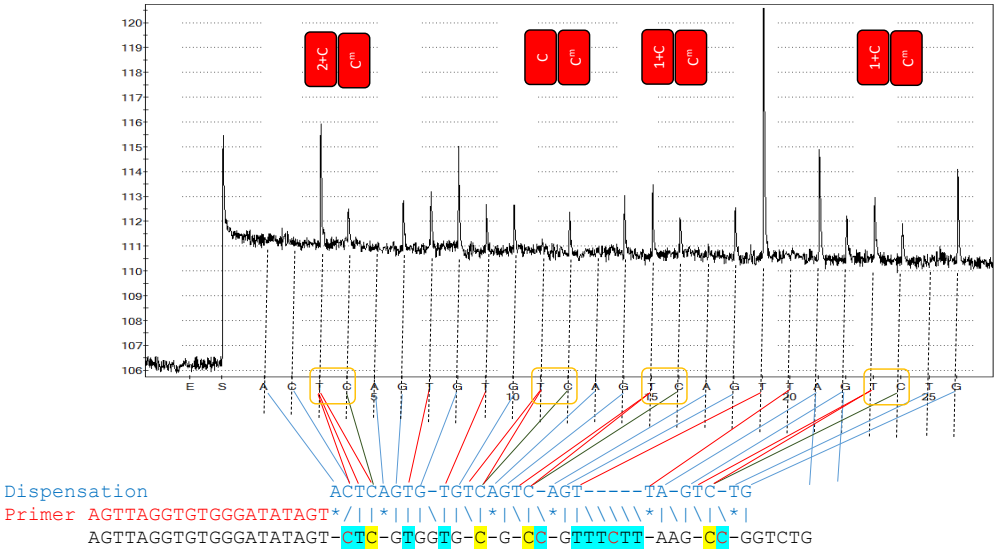

FIGURE 3.3: **LINE-1 methylation sequence**

Pyrosequencing primer LINE-1
5′-AGTTAGGTGTGGGATATAGT-3′

3.6 Analysis of embryo's sex

Analysis of embryo's sex was carried out by qPCR 5′-exonuclease test using Y-chromosome specific TaqMan probe, *TPSY1*, using the Custom Assay PrimeTime Standard qPCR (IDT, Tema Ricerca, Italy) following the manufacture's protocol reported in PrimeTime Gene Expression Master Mix protocol (PCR-10018-PR, v3; 1072639, IDT) [73].

The sequence of *TPSY1* was:

Probe (2.5 nmoles)
5′-/56-FAM/CGAAGCCGA/ZEN/GCTGCCCATCA/3IABkFQ/-3′

Primer 1 (5.0 nmoles)
5′-CATCCAGAGCGTCCCTGG-3′

Primer 2 (5.0 nmoles)
5′-TTCCCCTTTGTTCCCCAAA-3′

3.7 Statistical analysis

The analyses done in this thesis follows guideline for statistical analysis for case-control study of *Clarke GM. et al.,2011* [74]. All the statistical analyses were performed using the R studio software (R studio version 3.6.2). The Hardy-Weindberg equilbrium (HWE) was tested with "Hardy- Weinberg" package in R studio [75]. Odds ratios (ORs) and 95% confidence intervals (95% CIs) were estimated on nine SNPs with MAF > 5-10% using the formulas described calculation in figure 3.4 for allele analysis. The formula for calculation of ORs for genotype is showed in figure 3.5.

```
WBC_or_gen= myd_conv %>% filter(WBC != "NA") %>% group_by( type, rs, WBCgeno)  %>%tally()  %>% spread (WBCgeno, n
)

or_data=WBC_or_gen %>% mutate(totGenotypes=sum(RR, DR, DD, na.rm = T), totAllele=totGenotypes*2, cR=sum(RR*2, DR,
na.rm = T), fR=cR/totAllele, cD=sum(DR, DD*2, na.rm = T), fD=cD/totAllele )


or_table = or_data %>%    gather(variable, value, -(type:rs)) %>% unite(temp, type,  variable) %>% spread(temp,
value )
```

```
Dallele=or_table %>% mutate (type='dominant allele', or= (spo_cD*vol_cR)/ (spo_cR* vol_cD),
                             log_or=log(or),
                             se_log_or=sqrt(1/(spo_cD) + 1/(vol_cR) + 1/(spo_cR)+ 1/(vol_cD)),
                             ciUp=exp(log_or +1.96*se_log_or),
                             ciLow=exp(log_or -1.96*se_log_or)) %>% mutate(type="z-test", z= log_or /se_log_or) %
>% mutate (p_valuez= 2*pnorm(-abs(z)))

Rallele=or_table %>% mutate (type= 'recessive allele', or= (spo_cR* vol_cD)/(spo_cD*vol_cR),
                             log_or=log(or),
                             se_log_or=sqrt(1/(spo_cD) + 1/(vol_cR) + 1/(spo_cR)+ 1/(vol_cD)),
                             ciUp=exp(log_or +1.96*se_log_or),
                             ciLow=exp(log_or -1.96*se_log_or)) %>% mutate(type="z-test", z= log_or /se_log_or) %
>% mutate (p_valuez= 2*pnorm(-abs(z)))
```

FIGURE 3.4: **Formula in script of R studio of Odds Ratio for alleles** and 95% confidence intervals.

```
codomDR=or_table %>% mutate (type="codominantDR", or= (spo_DR*vol_DD) / (spo_DD* vol_DR),
                             log_or=log(or),
                             se_log_or=sqrt(1/spo_DD + 1/spo_DR + 1/vol_DD + 1/vol_DR),
                             ciUp= exp(log_or +1.96*se_log_or), ciLow=exp(log_or -1.96*se_log_or)) %>% mutate(typ
e="z-test", z= log_or /se_log_or) %>% mutate (p_valuez= 2*pnorm(-abs(z)))

codomRR=or_table %>% mutate (type="codominantRR", or= (spo_RR*vol_DD) / (spo_DD* vol_RR),
                             log_or=log(or),
                             se_log_or=sqrt(1/spo_RR + 1/vol_DD + 1/spo_DD + 1/vol_RR),
                             ciUp=exp(log_or +1.96*se_log_or),
                             ciLow=exp(log_or -1.96*se_log_or)) %>% mutate(type="z-test", z= log_or /se_log_or) %
>% mutate (p_valuez= 2*pnorm(-abs(z)))
```

FIGURE 3.5: **Formula in script of R studio of Odds Ratio for genotypes** and 95% confidence intervals.

A z-test was applied, and a two-tailed p-value less than 0.05 was considered statistically significant. ORs were calculated for genotypes and alleles, assuming a co-dominant genetic model and with a stratification for several parameters included in the clinical database: sex of the embryo, Body Mass Index (BMI), folic acid intake, and smoking exposure of the mothers.

Methylation scores of the LINE-1 sequence was calculated as the average over all 4 CpG islands in a sample. The distribution of LINE-1 methylation levels were

checked for normality using the Shapiro-Wilk test. For WBC samples that not departed from normality, as parametric test, an unpaired t-student test was used to test the significance. For CV samples, I used a non parametric test called "*Mann-Whitney*". I used the t-student for WBC and Mann-Whitney for CV, to evaluate the difference of average of methylation with a stratification for several parameters included in the clinical database, the same for OR. All p-values were two sided, with a threshold for declaring statistical significance of $p < 0.05$. The formula used for Mann-Whitney test in R is wilcox.test().

Formula for calculation t-student test for average of LINE-1 methylation in case RPL (Fig 3.6):

```
general_meth<- myd_meth %>% select(ID_DNA, type_ID, WBCmean_meth,sd_WBC_meth, age_mother, bmi,embryo_age, CV_sex,
Vitamin_mother, Smoke_quantity) %>% mutate(smoke_mother=ifelse (str_detect(Smoke_quantity, "Never"), "no", "ye
s")) %>% filter(WBCmean_meth <=0.99, sd_WBC_meth >=0.09) %>% distinct()

as_meth <- general_meth %>% filter(type_ID=='Spontaneous')
av_meth <- general_meth %>% filter(type_ID=='Volunteer')

#calculation of t student

t.test(as_meth$WBCmean_meth,av_meth$WBCmean_meth)
```

FIGURE 3.6: **Formula in script of R studio of t-student test** for average of LINE-1 methylation in RPL using R studio

Chapter 4

Results

4.1 Description of data set

Data consists of a database of medical records from interviews of 91 recurrent pregnancy loss **(RPL)**, 40 first pregnancy loss **(FPL)**, and 156 Voluntary Termination of pregnancy **(VTP)** with no previous miscarriages. The large majority of cases is of European origin (82.7%, African 9.6%, Asian 7.6%). The first exploratory analysis consist on graphical representation of data. The gestational age at pregnancy termination, calculated as the interval between the pregnancy termination date and the last menstruation date, ranges from 4 to 19.4 weeks (median is 10.2 weeks for FPL and RPL and 9.5 weeks for VTP) and it is not significantly different among the FPL and RPL group (Figure 4.1B, Mann-Whitney p-value FPL-RPL= 0.250). Because the legal term for voluntary termination of pregnancy is 90 days (*Italian law, named Bill 194, Article 6, and Comma B.*), a fair comparison between VPT and FPL and RPL is not possible, however it is notable that in our data FPL and RPL take place at latest within 19.4 weeks, confirming that miscarriages tend to occur early in pregnancy [76]. I observed a trend of higher incidence of RPL in women with high education, (Figure 4.1A), most likely because the age of pregnancy tend to be higher than other women for social reasons. Consistently, the median mother age at pregnancy termination is significantly higher in RPL compared to FPL (39.0 years and 28.9 years respectively), suggesting that recurrent miscarriages happen more often at advanced age (Figure 4.1C, Mann-Whitney p-value FPL-RPL = 2.84*10-6).

Moving to medical parameters, I see no significant differences between RPL, FPL and VTP in terms of Body Mass Index ((Figure 4.2B, Mann-Whitney p-value FPL-RPL= 0.675, p-value VTP-FPL= 0.485, p-value VTP-RPL= 0.620), while I observe a wider range of the age at menarche in RPL (range: 9-16 years) compared to FPL (range: 8-17 years) and VTP(range: 11-15 years) (4.2A, F-test p-value VTP-RPL= 0.0016,p-value FPL-RPL= 0.0288).

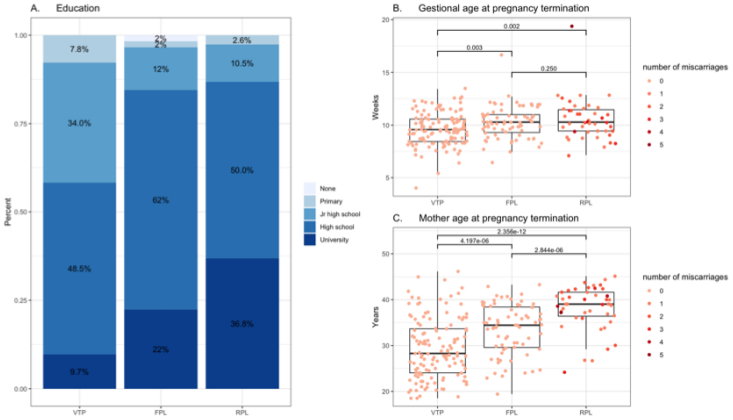

FIGURE 4.1: **Fig A**: Education. RPL samples tend to have higher educational level compared to VTP, probably due to high age at first pregnancy of educated women. **Fig B**: Gestational age at pregnancy termination. The age of the embryo in RPLs range from 4 to 19.4 weeks and there are no significant differences between the three groups. **Fig C**: Age of the mother. Median age of the mother at the event is significantly higher in RPL compared to FPL.

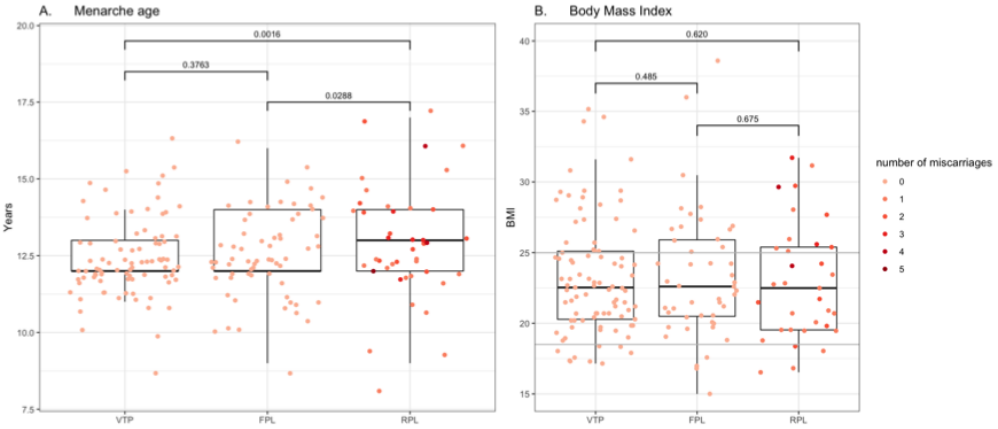

FIGURE 4.2: **Fig A**: Age of the menarche is significantly wider in RPL compared to FPL and VTP. **Fig B**: BMI does not influence pregnancy loss compared to VTP.

4.1.1 Quality of DNA extracted from CV and maternal contamination

Samples of tissue from CV are often in a state that is not ideal for DNA analyses, therefore I explored a range of possibilities for DNA preparation from CV that includes two methods of tissue homogenization and three methods of DNA isolation, as explained in Chapter 3. I do not observe significant difference between homogenization techniques, therefore I favour dry ones that is technically less challenging (Figure 4.3). Similarly, in the case of DNA isolation I considered two types of resin and one membrane and I do not observe significant differences in yield neither among the techniques or among samples from maternal blood, VTP and RPLs but only a slightly higher yield for one type of resin (Figure 4.3).

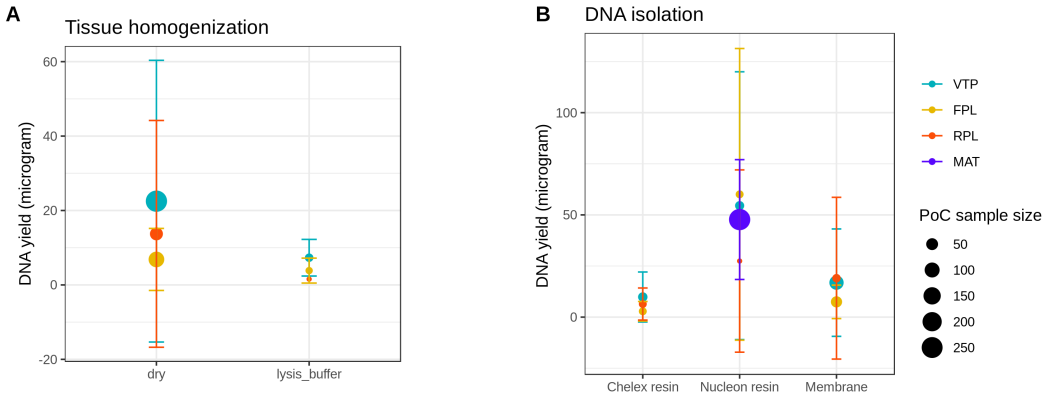

FIGURE 4.3: **Method for DNA extraction**. Test of different kit of extraction available in commerce

4.1.2 Screening for chromosomal aneuploidies for genetic analyses in the GREP project

Forty-four samples were contributed to the GREP project, where they were screened for chromosomal aneuploidies. In 40-50 % of all the pregnancies in the first trimester, RPL are mostly caused by chromosomal aneuploidies, such as trisomies or deletions of large chromosomal chunks [17, 77]. The GREP project focuses on cases in which the embryo is euploid when analyzed with current diagnostic techniques i.e. qfPCR and comparative genomic hybridization (arrayCGH). In 44 samples when screened for maternal contamination 11.4% drops off the analysis, due to technical challenges during sample collection. The first round of detection of aneuploidies on chromosomes 13, 15, 16, 18, 21, 22, X, and Y through Short Tandem Repeats analysis discarded 47.7% of samples. These types of repeats (tetra- or penta-nucleotide) are often expected to be found in heterozygosis, therefore triploidy is assumed when three alleles are found at several markers along a chromosome (complete) or part of it (partial). Similarly, uniparental disomy for a targeted region or chromosome is assumed when only one parental allele is amplified. Subsequent analysis through array-CGH and copy number variation detection form low-coverage sequencing discarded another 11.4% of samples (Figure 4.4). The most common aneuploidy in the analyzed cases is the trisomy of chromosome 22 (15.9%), followed by trisomy of chromosome 16 (9%) and 18 (4.5%) (Figure 4.4). Overall 29.5% of samples were euploid (Figure 4.5). These analyses were done by MeriGen Research S.r.l, IGENOMIX Italy and the CNR in Naples to select 12 euploid samples for whole-genome sequencing and further analyses that will be not described here.

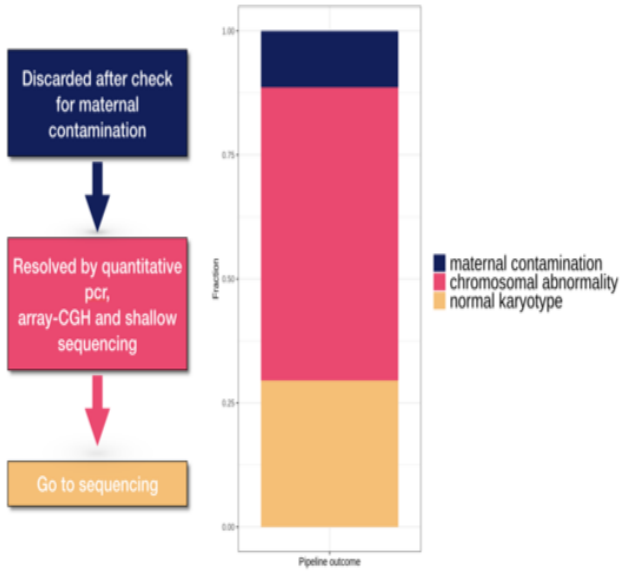

FIGURE 4.4: **Pre-sequencing screening outcome.** Approximately 10% of samples are discarded due to maternal contamination, and 60% for aneuploidies. Finally 30% of cases are euploid and proceed to sequencing.

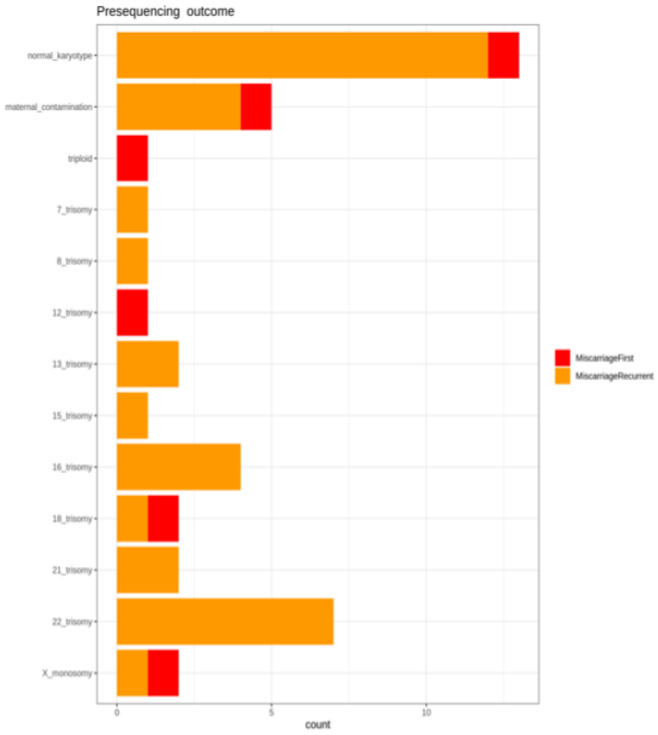

FIGURE 4.5: **Chromosomal aneuploidies detected in the embryos.** The most common aneuploidies are trisomy on chromosome 22 followed by trisomy on chromosome 16 and 18.

4.2 Genetic association between polymorphism in the genes under study and RPL

In this genetic association case-control study, I compared the frequency of allele and genotypes at specific genetic marker loci, called single-nucleotide polymorphisms (SNPs), in individuals from mothers and chorionic villi (CV) with recurrent pregnancy loss (RPL) and mothers and CV with volunteer termination pregnancy (VTP) in order to determine whether a statistical association exists beetween the RPL and the SNPs.

In table 1.1 of the Chapter 1, I describe the SNPs linked to the risk of miscarriage, selected after extensively searching in the literature. In brief, these genes are correlated with embryo's develope. The analyses done in this chapter follows guideline for statistical analysis for case-control study of *Clarke GM. et al.,2011* [74].

In 15% of samples, genotyping was repeated for quality testing. The indeterminate samples and the maternal and fetous genotypes dis-cordant were excluded by the analysis.

4.2.1 Hardy-Weindberg equilibrium (HWE)

The first step of the association analysis was to evaluate if the SNPs were in HWE. When stratified by sample type, I see two exceptions (Table 4.1): the gene *PLK4* rs2305957 in VTP of WBC (p=0.01). For CV the gene *ANXA5* rs1050606 in RPL of CV (p=0.012) and in VTP of CV, SNP rs1050606 (p=0.034), rs231775 (p=0.040) and rs2305957(p=0.030) are not in HWE equilibrium.

Nevertheless, when considering a threshold for significance of 0.05 in global evaluation all the loci are in HWE (Table 4.1) since their observed genotypic frequencies are not significantly different from the expected one, therefore all the SNP were used for subsequent analyses. All the SNPs used for this studied have a MAF> 5-10%, as shown in Chapter 1.

	RPL WBC				*VTP WBC*				*WBC*	*RPL CV*				*VTP CV*				*CV*
SNP	*DD (n)*	*DR (n)*	*RR (n)*	*HWE p-value*	*DD (n)*	*DR (n)*	*RR (n)*	*HWE p-value*	*Total HWE*	*DD (n)*	*DR (n)*	*RR (n)*	*HWE p-value*	*DD (n)*	*DR (n)*	*RR (n)*	*HWE p-value*	*Total HWE*
MTHFR rs1801133	42	54	30	*p= 0.16*	74	59	20	*p=0.18*	*p=0.05*	34	53	14	*p=0.43*	60	66	19	*p=0.96*	*p=0.70*
MTHFR rs1801131	65	43	17	*p=0.050*	69	63	20	*p=0.43*	*p=0.05*	48	42	10	*p=0.95*	71	63	11	*p=0.66*	*p=0.83*
HLA-G 14 bp	33	66	27	*p=0.66*	43	75	31	*p=0.95*	*p=0.68*	39	43	19	*p=0.32*	46	65	30	*p=0.51*	*p=0.20*
ANXA5 rs1050606	31	62	32	*p=0.94*	39	70	40	*p=0.53*	*p=0.60*	35	37	29	*p=0.012*	44	59	42	*p=0.034*	*p=0.050*
NKG2D rs2617170	54	57	13	*p=0.83*	61	70	21	*p=0.96*	*p=0.97*	46	49	6	*p=0.18*	55	64	26	*p=0.40*	*p=0.94*
IL-10 rs1800871	70	51	4	*p=0.20*	71	68	10	*p=0.30*	*p=0.10*	54	35	11	*p=0.22*	65	64	12	*p=0.59*	*p=0.80*
CTLA-4 rs231775	55	57	12	*p=0.72*	70	71	11	*p=0.27*	*p=0.26*	45	49	6	*p=0.16*	58	76	10	*p=0.040*	*p=0.10*
SMTH1 rs1979277	72	50	3	*p=0.14*	85	53	9	*p=0.95*	*p=0.41*	61	33	5	*p=0.93*	79	53	11	*p=0.73*	p=0.67
PLK4 rs2305957	68	51	5	*p=0.30*	92	42	11	*p=0.010*	*p=0.71*	56	41	3	*p=0.23*	74	69	2	*p=0.030*	p=0.055

TABLE 4.1: **HWE in WBC and CV**. **DD**= dominant homozygous, **DR**= heterozygous, **RR**= recessive homozygous

4.2.2 Odds ratio of the alleles under co-dominant model

All the SNP considered have two alleles, of which the most frequent in the general population according to Ensembl Release 99 (January 2020)[www.ensembl.org] was considered dominant. I calculated the ratio of the odds of disease in RPL carrying the recessive allele compared with RPL carrying the dominant one under a model of co-dominance. The results are shown in table 4.2 for both WBC and CV.
In WBC my results indicates that c.677T allele of the gene *MTHFR* rs1801133 associates in mothers with significant increased risk of RPL (OR= 1.72 p=0.0018) following a co-dominant model. In the rest of SNP in WBC no other p-value is significant. For CV, there is significant result for the gene *NKG2D* rs2617170, where the recessive allele seems to protect embryo from RPL (OR=0.64, IC 95%=0.44-0.95, p=0.025). No other SNP is significant.

	Allele analysis in WBC				Allele analysis in CV			
SNP	***RPL***	***VTP***	***OR (IC 95%)***	***p-value***	***RPL***	***VTP***	***OR (IC 95%)***	***p-value***
MTHFR rs1801133								
D	138	207	1.00 (Reference)		121	186	1.00 (Reference)	
R	114	99	1.72(1.22-2.43)	**0.0018**	81	104	1.19(0.82-1.73)	0.33
MTHFR rs1801131								
D	173	201	1.00 (Reference)		138	205	1.00 (Reference)	
R	77	103	0.86(0.60-1.24)	0.44	62	85	1.083(0.73-1.60)	0.68
HLA-G 14bp								
D	132	161	1.00 (Reference)		121	157	1.00 (Reference)	
R	120	137	1.068(0.76-1.49)	0.69	81	125	0.84(0.58-1.21)	0.35
ANXA5 rs1050606								
D	124	148	1.00(Reference)		107	147	1.00 (Reference)	
R	126	150	1.002(0.71-1.39)	0.98	95	143	0.91(0.63-1.30)	0.61
NKG2D rs2617170								
D	165	192	1.00(Reference)		141	174	1.00 (Reference)	
R	83	112	0.86(0.60-1.22)	0.40	61	116	0.64(0.44-0.95)	**0.025**
IL-10 rs1800871								
D	191	210	1.00(Reference)		143	194	1.00 (Reference)	
R	59	88	0.73(0.50-1.081)	0.11	57	88	0.87(0.59-1.30)	0.52
CTLA-4 rs231775								
D	167	211	1.00(Reference)		139	192	1.00 (Reference)	
R	81	93	1.10(0.76-1.57)	0.60	61	96	0.87(0.60-1.40)	0.50
SMTH1 rs1979277								
D	194	223	1.00(Reference)		155	211	1.00 (Reference)	
R	56	71	0.90(0.60-1.35)	0.63	43	75	0.78(0.50-1.19)	0.25
PLK4 rs2305957								
D	187	226	1.00(Reference)		153	217	1.00 (Reference)	
R	61	64	1.15(0.77-1.71)	0.48	47	73	0.91(0.44-0.95)	0.67

TABLE 4.2: **Odds ratio of disease in RPL** carrying the **R** recessive allele compared with RPL carrying the dominant one **D**, under a model of co-dominance in WBC and CV

4.2.3 Odds ratio of the genotype under co-dominant model

I calculated the ratio of the odds of disease in RPL carrying homozygous and heterozygous genotypes for the recessive allele compared with RPL carrying the homozygous dominant genotype under a model of co-dominance. The results for both WBC and CV are shown in table 4.3 and Figures 4.6. Of all the SNP considered, two show significant association with risk or protection in mothers (WBC) and embryos (CV), while the other SNPs instead seems to have no association with RPL.
Details of the ORs for the rs1801133 SNP in the *MTHFR* gene WBC shown in Figure 4.7, indicates that c.677T allele in mothers associates with significant increased risk of RPL (OR= 1.72, IC 95%=1.22-2.43, p=0.0018) following a co-dominant model, with an almost 3-fold increased risk in T/T homozygote women (OR=2.64,IC 95%=1.33-5.22, p=0.004), and a lower risk in heterozygotes (OR=1.61,IC 95%=0.95-1.77, p=0.075). Similarly, the genotype GG in gene *NKG2D*, seems protect embryo from RPL (OR=0.27, IC 95%=0.10-0.72, p=0.0067) as shown in Figure 4.7.

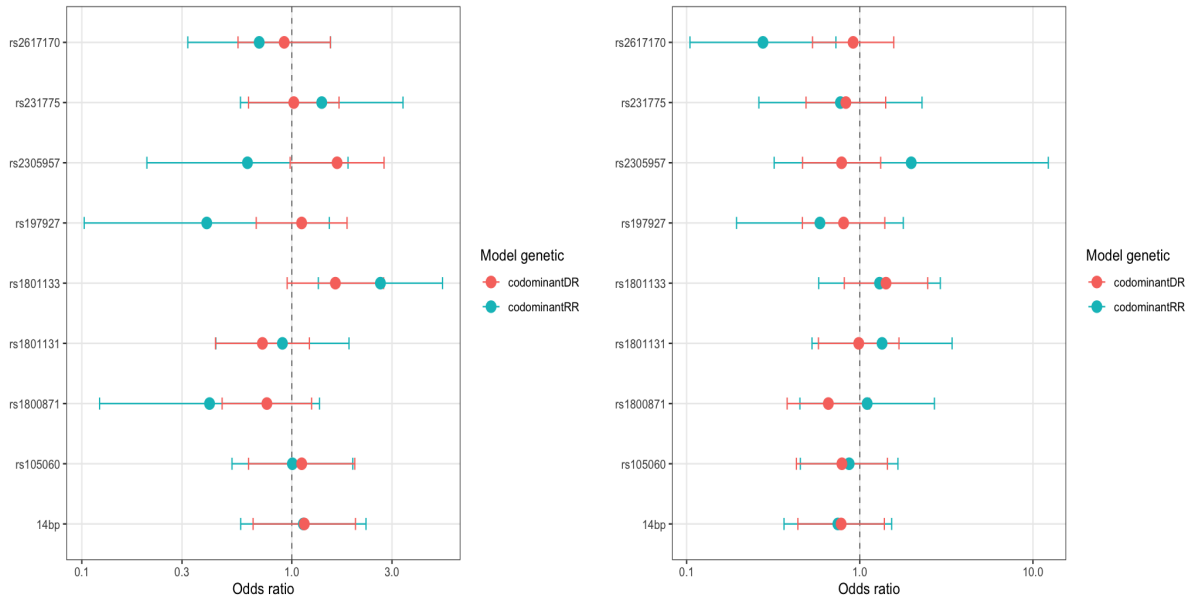

FIGURE 4.6: **Odds ratio and confidence intervals** calculated under a model of co-dominance for the heterozygous (**DR**) and homozygous recessive (**RR**) genotypes at the considered genetic loci. Data relative to WBC on the left and to CV on the right.

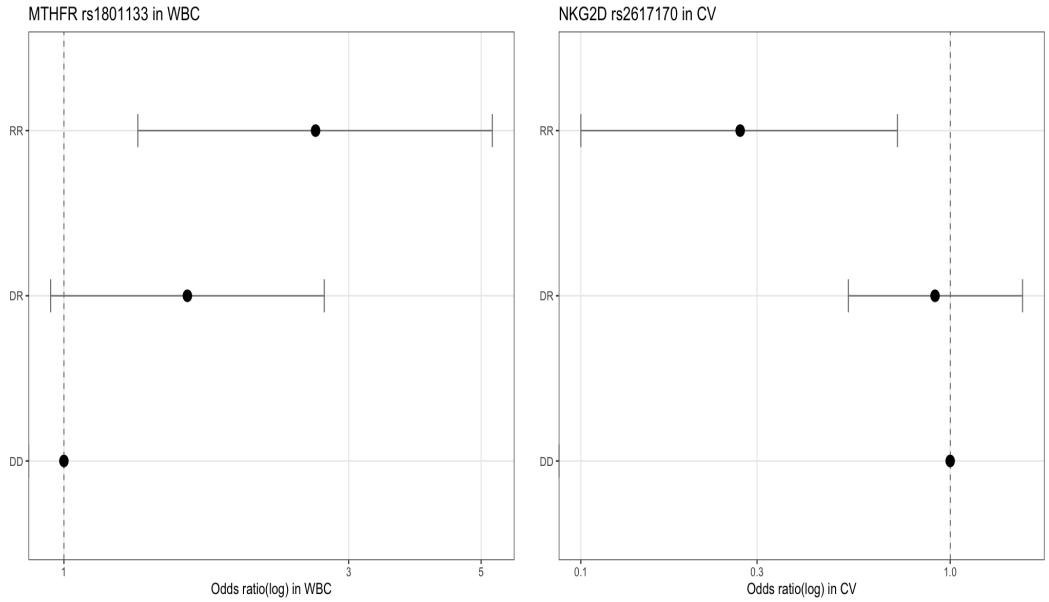

FIGURE 4.7: **Odds ratios and confidence intervals** calculated under a model of co-dominance for the heterozygous (**DR**) and homozygous recessive (**RR**) genotypes at the considered genetic loci. Data relative to *MTHFR* rs1801133 in WBC on the left and to *NKG2D* rs2617170 in CV on the right.

Co-dominant model in WBC					Co-dominant model in CV			
SNP	*RPL*	*VTP*	*OR (IC 95%)*	*p-value*	*RPL*	*VTP*	*OR (IC 95%)*	*p-value*
MTHFR rs1801133								
DD	42	74	1.00 (Reference)		34	60	1.00 (Reference)	
DR	54	59	1.61(0.95-2.73)	0.075	53	66	1.41(0.81-2.46	0.21
RR	30	20	2.64(1.33-5.22)	**0.004**	14	19	1.30(0.57-2.91)	0.52
MTHFR rs1801131								
DD	65	69	1.00 (Reference)		48	71	1.00 (Reference)	
DR	43	63	0.72(0.43-1.21)	0.21	42	63	0.98(0.57-1.68)	0.95
RR	17	20	0.90(0.43-1.87)	0.78	10	11	1.34(0.52-3.41)	0.53
HLA-G 14bp								
DD	33	43	1.00 (Reference)		39	46	1.00 (Reference)	
DR	66	75	1.14(0.65-2.01)	0.63	43	65	0.78(0.43-1.38)	0.39
RR	27	31	1.13(0.57-2.25)	0.71	19	30	0.74 (0.36-1.52)	0.42
ANXA5 rs1050606								
DD	31	39	1.00(Reference)		35	44	1.00 (Reference)	
DR	62	70	1.14(0.65-2.010)	0.71	37	59	0.78(0.43-1.38)	0.44
RR	32	40	1.0064(0.51-1.95)	0.98	29	42	0.86(0.45-1.66)	0.66
NKG2D rs2617170								
DD	54	61	1.00(Reference)		46	55	1.00 (Reference)	
DR	57	70	0.91(0.55-1.52)	0.74	49	64	0.91(0.53-1.57)	0.74
RR	13	21	0.69(0.31-1.52)	0.36	6	26	0.27(0.10-0.72)	**0.0067**
IL-10 rs1800871								
DD	70	71	1.00(Reference)		54	65	1.00 (Reference)	
DR	51	68	0.76(0.46-1.24)	0.27	35	64	0.65(0.38-1.13)	0.13
RR	4	10	0.40(0.61-1.35)	0.13	11	12	1.10(0.45-2.69)	0.82
CTLA-4 rs231775								
DD	55	70	1.00(Reference)		45	58	1.00 (Reference)	
DR	57	71	1.021(0.62-1.67)	0.93	49	76	0.83(0.48-1.41)	0.49
RR	12	11	1.38(0.56-3.38)	0.46	6	10	1.27(0.43-4.090)	0.64
SMTH1 rs1979277								
DD	72	85	1.00(Reference)		61	79	1.00 (Reference)	
DR	50	53	1.13(0.67-1.83)	0.67	33	53	0.80(0.46-1.39)	0.44
RR	3	9	0.39(0.10-1.50)	0.16	5	11	0.58(0.19-1.78)	0.34
PLK4 rs2305957								
DD	68	92	1.00(Reference)		56	74	1.00 (Reference)	
DR	51	42	1.64(0.98-2.74)	0.060	41	69	0.78(0.46-1.32)	0.36
RR	5	11	0.61(0.20-1.80)	0.38	3	2	1.98(0.32-12.26)	0.45

TABLE 4.3: **Odds ratio and confidence intervals** with p-value calculated under a model of co-dominance for the heterozygous (**DR**) and homozygous recessive (**RR**) genotypes at the considered genetic loci. Data relative to *MTHFR* rs1801133 in WBC on the left and to *NKG2D* rs2617170 in CV on the right.

4.3 Stratification ORs based on information from the clinical database

After calculation of OR for all SNP, the analyses focus on the stratification of the OR by the parameters present in the clinical database. In particular, the parameters analyzed here are: BMI, smoking exposure, and folic acid intake of mothers; and sex of the embryo.

4.3.1 BMI

Body Mass Index (BMI) is a calculated using a person's height and weight. This values is take during the pregnancy. The formula is BMI = kg/m^2 where kg is a person's weight in kilograms and m^2 is their height in metres squared. A BMI of 25.0 or more is considered "overweight", while the healthy range is 18.5 to 24.9, "normal weight". I used these threshold to subdivide mothers in two categories: normal weight and overweight. I found no significant association when considering WBC (Table 4.4) and one significant association when considering CV (Table 4.5). For the SNP of gene *NKG2D* rs2617170, the genotype A/A seems decrease the risk of RPL in CV, with a significant result (OR=0.18, IC 95%=0.045-0.72, p=0.015).

Co-dominant model analysis in WBC								
	BMI's mother "Normal weight"				*BMI's mother "Overweight"*			
SNP	***PL***	***VTP***	***OR (IC 95%)***	***p-value***	***PL***	***VTP***	***OR (IC 95%)***	***p-value***
MTHFR rs1801133								
DD	22	38	1.00 (Reference)		15	16	1.00 (Reference)	
DR	25	24	1.56(0.71-3.40)	0.26	11	9	1.30(0.42-4.027)	0.64
RR	13	12	2.23(0.95-5.26)	0.064	7	3	2.48(0.18-11.43)	0.24
MTHFR rs1801131								
DD	30	31	1.00 (Reference)		16	10	1.00 (Reference)	
DR	21	27	0.80(0.37-1.71)	0.57	11	13	0.52(0.17-1.63)	0.27
RR	8	11	0.73(0.28-1.89)	0.53	6	5	0.75(0.18-3.12)	0.69
HLA-G 14bp								
DD	19	26	1.00 (Reference)		9	7	1.00 (Reference)	
DR	30	29	1.41(0.64-3.092)	0.38	15	14	0.83 (0.24-2.84)	0.77
RR	11	11	1.35(0.54-3.37)	0.53	9	7	1(0.24-4.042)	1.00
ANXA5 rs1050606								
DD	13	18	1.00(Reference)		9	8	1.00 (Reference)	
DR	32	34	1.30(0.64-3.092)	0.38	14	16	0.77(0.23-2.56)	0.67
RR	14	14	0.84(0.33-2.13)	0.73	10	4	2.22(0.49-9.96)	0.30
NKG2D rs2617170								
DD	26	25	1.00(Reference)		13	13	1.00 (Reference)	
DR	28	34	0.79(0.37-1.66)	0.53	16	11	1.45(0.49-4.31)	0.50
RR	5	9	0.56(0.19-1.62)	0.29	3	4	0.75(0.13-4.035)	0.73
IL-10 rs1800871								
DD	37	30	1.00(Reference)		15	11	1.00 (Reference)	
DR	19	28	0.70(0.36-1.37)	0.31	17	17	0.73(0.26-2.050)	0.55
RR	3	7	0.55(0.25-1.17)	0.12	1	NA	NA	NA
CTLA-4 rs231775								
DD	26	32	1.00 (Reference)		13	14	1.00 (Reference)	
DR	29	33	1.081(0.36-2.21)	0.83	16	13	1.3(0.46-3.79)	0.60
RR	3	3	1.8(0.48-6.69)	0.38	4	1	4.30(0.42-43.73)	0.21
SMTH1 rs1979277								
DD	34	40	1.00(Reference)		23	15	1.00 (Reference)	
DR	22	23	1.12(0.53-2.36)	0.75	10	11	0.60(0.20-1.73)	0.34
RR	3	2	1(0.19-5.22)	1.00	NA	1	NA	NA
PLK4 rs2305957								
DD	31	38	1.00(Reference)		18	17	1.00 (Reference)	
DR	26	15	2.12(0.96-4.69)	0.062	13	10	1.22(0.42-3.53)	0.70
RR	2	9	0.23(0.047-1.141)	0.072	1	1	0.94(0.054-16.32)	0.97

TABLE 4.4: **Odd ratios and confidence intervals** with p-value calculated under a model of co-dominance for the heterozygous (**DR**) and homozygous recessive (**RR**) genotypes at the considered genetic loci in relation of the mother's bmi normal weight and overweight in WBC.

Co-dominant model analysis in CV								
	BMI's mother "Normal weight"				*BMI's mother "Overweight"*			
SNP	***PL***	***VTP***	***OR (IC 95%)***	***p-value***	***PL***	***VTP***	***OR (IC 95%)***	***p-value***
MTHFR rs1801133								
DD	18	26	1.00 (Reference)		9	15	1.00 (Reference)	
DR	27	32	1.21(0.55-2.68)	0.62	13	10	2.16(0.67-4.6.96)	0.19
RR	5	7	1.031(0.28-3.76)	0.96	4	2	4.16(0.66-26.13)	0.12
MTHFR rs1801131								
DD	22	36	1.00 (Reference)		15	12	1.00 (Reference)	
DR	21	24	1.43(0.64-3.15)	0.37	11	13	0.67(0.22-2.043)	0.48
RR	7	5	2.29(0.64-8.11)	0.19	1	2	0.40(0.21-9.37)	0.47
HLA-G 14bp								
DD	18	22	1.00 (Reference)		13	10	1.00 (Reference)	
DR	22	29	0.92(0.40-2.13)	0.85	11	13	0.65 (0.20-2.057)	0.46
RR	10	14	0.87(0.31-2.42)	0.79	3	4	0.57(0.10-3.18)	0.52
ANXA5 rs1050606								
DD	18	26	1.00(Reference)		7	9	1.00 (Reference)	
DR	15	24	0.90(0.37-2.18)	0.82	10	16	1.92(0.57-6.87)	0.31
RR	17	15	1.63(0.65-4.10)	0.29	5	8	0.80(0.18-3.57)	0.77
NKG2D rs2617170								
DD	22	20	1.00(Reference)		12	13	1.00 (Reference)	
DR	25	30	0.75(0.33-1.69)	0.49	13	8	1.76(0.54-5.72)	0.34
RR	3	15	0.18(0.045-0.72)	**0.015**	2	6	0.36(0.060-2.14)	0.26
IL-10 rs1800871								
DD	26	27	1.00(Reference)		14	10	1.00 (Reference)	
DR	18	28	0.66(0.29-1.48)	0.32	9	15	0.42(0.13-1.36)	0.15
RR	5	9	0.57(0.17-1.95)	0.37	4	2	1.42(0.21-9.37)	0.71
CTLA-4 rs231775								
DD	20	28	1.00 (Reference)		12	11	1.00 (Reference)	
DR	24	36	0.93(0.43-2.019)	0.86	14	16	0.80(0.27-2.38)	0.69
RR	5	1	7.00(0.75-64.60)	0.086	1	NA	NA	NA
SMTH1 rs1979277								
DD	31	41	1.00(Reference)		20	13	1.00 (Reference)	
DR	14	19	0.97(0.42-2.24)	0.95	6	13	0.30(0.090-0.98)	0.050
RR	3	5	0.79(0.17-3.57)	0.76	1	1	0.65(0.037-11.33)	0.76
PLK4 rs2305957								
DD	28	33	1.00(Reference)		15	13	1.00 (Reference)	
DR	18	32	0.66(0.30-1.42)	0.29	12	13	0.80(0.27-2.35)	0.68
RR	3	NA	NA	NA	NA	1	NA	NA

TABLE 4.5: **Odd ratios and confidence intervals** with p-value calculated under a model of co-dominance for the heterozygous (**DR**) and homozygous recessive (**RR**) genotypes at the considered genetic loci in relation of the mother's bmi normal weight and overweight in CV.

4.3.2 Smoking

There are a lot of observation that suggest that environmental factors contribute to the occurrence of RPL and FPL. Cigarette smoking hurts trophoblast function and is linked with a dose-dependent increase in the risk of miscarriage [20]. A survey conducted by Miceli et all.,in 2005[21], highlighted the inhibiting effect that nicotine and its metabolites have on the production of progesterone and on the release of Prostaglandin E2 (PGE2) by human luteal cells. Likewise, carbon monoxide forms contribute to the produc-ion of high levels of fetal carboxyhemoglobin, which causes poor tissue oxygenation and a toxic effect on villus cells, reducing the passage of nutrients. Smoking habits were classified using five options: "Never smoke", "Previous smoke", "From 1 to 5 cigarettes/day", "More of 5 cigarettes/day". I grouped mothers in two classes: the first includes the 'Never Smoke' and the second all the other options and I calculated the OR on this parameter. In table 4.6 in WBC, it is shown a significant result for SNP rs2305957 of the gene *PLK4* in G/A genotype, where increase the risk of RPL (OR=4.66, IC95%=1.54-14.14, p=0.0064). For SNP rs181133 of the gene *MTHFR*, the risk increase in T/T genotype (OR=4.2, IC95%=1.19-14.88, p=0.026) in women with no exposure to smoke during pregnancy. No other values are significative. In table 4.7 shows CV results. For SNP, rs2617170 of the gene *NKG2D* the recessive genotype A/A mitigate the effect of smoking that drives it to protect from RPL, in CV where mothers are not smokers (OR=0.079, IC95%=0.014-0.42, p=0.0029).

Co-dominant model analysis in WBC								
	Smoking exposure "Yes"				*Smoking exposure "No"*			
SNP	*RPL*	*VTP*	*OR (IC 95%)*	*p-value*	*RPL*	*VTP*	*OR (IC 95%)*	*p-value*
MTHFR rs1801133								
DD	14	18	1.00 (Reference)		19	20	1.00 (Reference)	
DR	18	17	1.36(0.51-3.56)	0.53	19	14	1.42(0.56-3.63)	0.45
RR	4	8	0.64(0.16-2.57)	0.53	16	4	4.2(1.19-14.88)	**0.026**
MTHFR rs1801131								
DD	16	22	1.00 (Reference)		31	15	1.00 (Reference)	
DR	14	12	1.60(0.58-4.38)	0.36	17	18	0.45(0.18-1.12)	0.089
RR	6	8	1.031(0.29-3.56)	0.96	6	5	0.58(0.15-2.21)	0.42
HLA-G 14bp								
DD	15	13	1.00 (Reference)		14	12	1.00 (Reference)	
DR	12	20	0.52(0.18-1.45)	0.21	34	17	1.71(0.65-4.50)	0.27
RR	9	8	0.97(0.29-3.26)	0.97	6	7	0.73 (0.19-2.79)	0.65
ANXA5 rs1050606								
DD	8	10	1.00(Reference)		16	12	1.00 (Reference)	
DR	18	29	0.77(0.18-1.45)	0.65	26	13	1.50(0.55-4.084)	0.42
RR	10	3	4.16(0.61-3.26)	0.078	12	10	0.90(0.29-2.77)	0.85
NKG2D rs2617170								
DD	20	17	1.00(Reference)		21	13	1.00 (Reference)	
DR	13	22	0.50(0.19-1.28)	0.15	24	17	0.87(0.34-2.21)	0.77
RR	3	4	0.63(0.12-3.25)	0.60	8	7	0.70(0.20-2.41)	0.58
IL-10 rs1800871								
DD	24	20	1.00(Reference)		26	16	1.00 (Reference)	
DR	11	19	0.48(0.18-1.24)	0.13	26	19	0.83(0.34-1.99)	0.68
RR	1	3	0.27(0.027-2.88)	0.28	2	3	0.38(0.057-2.56)	0.32
CTLA-4 rs231775								
DD	15	22	1.00(Reference)		23	14	1.00 (Reference)	
DR	20	18	1.62(0.65-4.066)	0.29	23	22	0.63(0.26-1.54)	0.31
RR	1	2	0.73(0.061-8.83)	0.81	8	1	4.86(0.54-143.18)	0.15
SMTH1 rs1979277								
DD	26	26	1.00(Reference)		23	14	1.00 (Reference)	
DR	10	13	0.76(0.28-2.065)	0.60	28	17	0.70(0.28-1.75)	0.21
RR	NA	1	NA	NA	3	31	2.60(0.25-27.11)	0.42
PLK4 rs2305957								
DD	17	28	1.00(Reference)		27	14	1.00 (Reference)	
DR	17	6	4.66(1.54-14.14)	**0.010**	23	17	0.70(0.28-1.72)	0.44
RR	2	6	0.54(0.099-3.035)	0.50	3	3	0.51(0.092-2.91)	0.45

TABLE 4.6: **Odds ratio and confidence intervals** with p-value calculated under a model of co-dominance for the heterozygous (**DR**) and homozygous recessive (**RR**) genotypes at the considered genetic loci in relation of smoking exposure in WBC.

Co-dominant model analysis in CV								
	Smoking exposure "Yes"				*Smoking exposure "No"*			
SNP	***RPL***	***VTP***	***OR (IC 95%)***	***p-value***	***RPL***	***VTP***	***OR (IC 95%)***	***p-value***
MTHFR rs1801133								
DD	7	9	1.00 (Reference)		15	22	1.00 (Reference)	
DR	15	21	0.91(0.27-3.016)	0.88	23	12	2.81(1.078-7.32)	**0.034**
RR	2	6	0.42(0.065-2.81)	0.37	5	4	1.83(0.42-7.96)	0.41
MTHFR rs1801131								
DD	13	19	1.00 (Reference)		17	18	1.00 (Reference)	
DR	9	15	0.87(0.29-2.59)	0.81	22	18	1.29(0.52-3.21)	0.57
RR	2	2	1.46(0.18-11.73)	0.72	4	2	2.11(0.34-13.099)	0.41
HLA-G 14bp								
DD	7	12	1.00 (Reference)		17	13	1.00 (Reference)	
DR	12	15	1.37(0.41-4.56)	0.60	21	19	0.84(0.32-2.19)	0.72
RR	5	9	0.95(0.22-4.006)	0.94	5	6	0.63 (0.15-2.55)	0.52
ANXA5 rs1050606								
DD	11	15	1.00(Reference)		16	17	1.00 (Reference)	
DR	10	15	0.90(0.29-2.77)	0.86	15	10	1.59 (0.55-4.56)	0.38
RR	3	6	0.68(0.13-3.34)	0.63	12	11	1.15(0.39-3.36)	0.78
NKG2D rs2617170								
DD	12	15	1.00(Reference)		18	10	1.00 (Reference)	
DR	11	17	0.80(0.27-2.36)	0.69	23	14	0.91(0.32-2.52)	0.86
RR	1	4	0.31(0.030-3.17)	0.32	2	14	0.079(0.014-0.42)	**0.0029**
IL-10 rs1800871								
DD	12	14	1.00(Reference)		22	16	1.00 (Reference)	
DR	9	16	0.65(0.21-2.017)	0.46	15	18	0.60(0.23-1.55)	0.29
RR	2	5	0.46(0.076-2.85)	0.40	6	3	1.45(0.31-6.70)	0.63
CTLA-4 rs231775								
DD	9	17	1.00(Reference)		20	16	1.00 (Reference)	
DR	13	17	1.44(0.48-4.26)	0.50	20	21	0.76(0.31-1.87)	0.55
RR	1	2	0.94(0.075-11.88)	0.96	3	1	2.40(0.22-25.33)	0.46
SMTH1 rs1979277								
DD	17	19	1.00(Reference)		23	25	1.00 (Reference)	
DR	5	15	0.37(0.11-1.24)	0.10	17	12	1.53(0.60-3.90)	0.36
RR	1	2	0.55(0.046-6.72)	0.64	2	1	2.17(0.18-25.60)	0.53
PLK4 rs2305957								
DD	12	22	1.00(Reference)		22	19	1.00 (Reference)	
DR	11	14	1.44(0.48-4.26)	0.50	. 18	19	0.81(0.33-1.99)	0.65
RR	1	NA	NA	NA	2	NA	NA	NA

TABLE 4.7: **Odds ratio and confidence intervals** with p-value calculated under a model of co-dominance for the heterozygous (**DR**) and homozygous recessive (**RR**) genotypes at the considered genetic loci in relation of smoking exposure in CV.

4.3.3 Folic acid intake

The folic acid intake measures if women used folic acid during pregnancy or not. In VTP, no women intakes folic acid, because the pregnancy probably was not planned, so I can't calculate OR because I don't have any case of VTP with folic acid supplementation. I calculated the OR based on this parameter: Table 4.8 shows that, for WBC, the recessive genotype T/T of the SNP rs1801133 of the gene *MTHFR* is associated with an increase of risk of RPL in women that do not take folic acid during pregnancy (OR=2.47, IC95%=1.00-6.07, p=0.050). Also, an increase of this risk seems to be associated to the heterozygous genotype G/A in rs2305957 of gene *PLK4*(OR=2.090, IC95%=1.099-3.97, p=0.024) in the same group of women. Table 4.9 relative to CV samples, shows that for the SNP rs2617170 of the gene *NKG2D* the recessive genotype A/A protects the embryo from RPL where mothers did not take folic acid during pregnancy (OR=0.31, IC95%=0.098-0.99, p=0.050) .

Co-dominant model analysis in WBC								
	Folic acid intake "No"				*Folic acid intake "Yes" VS "No"*			
SNP	*RPL*	*VTP*	*OR (IC 95%)*	*p-value*	*RPL*	*RPL*	*OR (IC 95%)*	*p-value*
MTHFR rs1801133								
DD	22	68	1.00 (Reference)		18	22	1.00 (Reference)	
DR	28	53	1.63(0.84-3.17)	0.15	25	28	1.091(0.47-2.48)	0.83
RR	12	15	2.47(1.00-6.07)	**0.050**	16	12	1.62(0.61-4.31)	0.32
MTHFR rs1801131								
DD	32	59	1.00 (Reference)		28	32	1.00 (Reference)	
DR	22	59	0.68(0.35-1.31)	0.25	21	22	1.090(0.49-2.38)	0.82
RR	8	17	0.86(0.33-2.23)	0.76	9	8	1.28(0.43-3.78)	0.64
HLA-G 14bp								
DD	15	40	1.00 (Reference)		18	25	1.00 (Reference)	
DR	34	66	1.37(0.66-2.83)	0.38	28	34	0.68(0.29-1.60)	0.38
RR	13	26	1.33(0.54-3.25)	0.52	13	13	0.83(0.29-2.33)	0.72
ANXA5 rs1050606								
DD	13	34	1.00(Reference)		16	13	1.00 (Reference)	
DR	36	65	1.45(0.67-3.090)	0.33	25	36	0.56(0.23-1.37)	0.20
RR	13	33	1.030(0.411-2.54)	0.94	17	13	1.062(0.38-2.97)	0.90
NKG2D rs2617170								
DD	26	56	1.00(Reference)		26	26	1.00 (Reference)	
DR	31	61	1.094(0.58-2.065)	0.78	23	31	0.74(0.34-1.59)	0.44
RR	5	18	0.59(0.20-1.78)	0.35	8	5	1.60(0.46-5.54)	0.56
IL-10 rs1800871								
DD	34	63	1.00(Reference)		32	34	1.00 (Reference)	
DR	25	62	0.74(0.40-1.39)	0.36	25	25	1.062(0.50-2.21)	0.87
RR	3	7	0.79(0.19-3.27)	0.74	1	3	0.35(0.035-3.58)	0.37
CTLA-4 rs231775								
DD	25	64	1.00(Reference)		26	25	1.00 (Reference)	
DR	29	62	1.068(0.58-1.96)	0.84	27	29	0.89(0.41-1.91)	0.77
RR	7	9	1.99(0.66-45.95)	0.21	5	7	0.68(0.19-2.45)	0.56
SMTH1 rs1979277								
DD	35	74	1.00(Reference)		33	35	1.00 (Reference)	
DR	25	48	1.10(0.58-2.065)	0.76	24	25	1.018(0.48-2.12)	0.96
RR	2	8	0.52(0.10-2.62)	0.43	1	2	0.53(0.045-6.12)	0.61
PLK4 rs2305957								
DD	30	80	1.00(Reference)		36	30	1.00 (Reference)	
DR	29	37	2.090(1.099-3.97)	**0.024**	19	29	0.54(0.25-1.16)	0.11
RR	3	11	0.72(0.18-2.78)	0.64	2	3	0.55(0.087-3.54)	0.53

TABLE 4.8: **Odds ratio and confidence intervals** with p-value calculated under a model of co-dominance for the heterozygous (**DR**) and homozygous recessive (**RR**) genotypes at the considered genetic loci in relation of folic acid during pregnancy in WBC.

Co-dominant model analysis in CV								
	Folic acid intake "No"				*Folic acid intake "Yes" VS "No"*			
SNP	*RPL*	*VTP*	*OR (IC 95%)*	*p-value*	*RPL*	*RPL*	*OR (IC 95%)*	*p-value*
MTHFR rs1801133								
DD	16	58	1.00 (Reference)		16	16	1.00 (Reference)	
DR	29	56	1.87(0.92-3.82)	0.083	24	29	0.82(0.34-1.99)	0.67
RR	6	16	1.35(0.45-4.040)	0.58	7	6	1.16(0.32-4.24)	0.81
MTHFR rs1801131								
DD	25	63	1.00 (Reference)		20	25	1.00 (Reference)	
DR	22	56	0.99(0.50-1.94)	0.97	20	22	1.13(0.48-2.64)	0.76
RR	4	11	0.91(0.26-3.14)	0.88	6	4	1.87(0.46-7.56)	0.37
HLA-G 14bp								
DD	18	44	1.00 (Reference)		20	18	1.00 (Reference)	
DR	23	54	1.041(0.49-2.16)	0.91	19	23	0.74(0.30-1.79)	0.50
RR	10	28	0.87(0.35-2.16)	0.76	8	10	0.72(0.23-2.22)	0.56
ANXA5 rs1050606								
DD	21	41	1.00(Reference)		13	21	1.00 (Reference)	
DR	16	50	0.62(0.28-1.34)	0.23	19	16	1.91(0.73-5.0076)	0.18
RR	14	39	0.70(0.31-1.56)	0.38	15	14	1.73(0.63-4.72)	0.28
NKG2D rs2617170								
DD	25	51	1.00(Reference)		19	25	1.00 (Reference)	
DR	22	53	0.84(0.42-1.68)	0.63	26	22	1.55(0.68-3.54)	0.29
RR	4	26	0.31(0.098-0.99)	**0.050**	2	4	0.65(0.10-3.97)	0.64
IL-10 rs1800871								
DD	26	58	1.00(Reference)		26	26	1.00 (Reference)	
DR	18	58	0.69(0.34-1.39)	0.30	17	18	0.94(0.40-2.22)	0.89
RR	7	10	1.56(0.53-4.55)	0.41	3	7	0.42(0.099-1.84)	0.25
CTLA-4 rs231775								
DD	23	55	1.00(Reference)		20	23	1.00 (Reference)	
DR	26	65	0.95(0.49-1.86)	0.89	22	26	0.97(0.42-2.22)	0.94
RR	2	9	0.53(0.10-2.65)	0.44	4	2	2.30(0.38-13.91)	0.36
rs1979277								
DD	32	70	1.00(Reference)		27	32	1.00 (Reference)	
DR	17	48	0.77(0.38-1.55)	0.47	15	17	1.045(0.44-2.47)	0.91
RR	1	10	0.21(0.026-1.78)	0.15	4	1	4.74(0.49-44.99)	0.17
PLK4 rs2305957								
DD	29	65	1.00(Reference)		26	29	1.00 (Reference)	
DR	20	63	0.71(0.36-1.38)	0.31	19	20	1.059(0.46-2.40)	0.89
RR	1	2	1.12(0.097-12.85)	0.92	2	1	2.23(0.19-26.063)	0.52

TABLE 4.9: **Odds ratio and confidence intervals** with p-value calculated under a model of co-dominance for the heterozygous (**DR**) and homozygous recessive (**RR**) genotypes at the considered genetic loci in relation of folic acid during pregnancy in CV.

4.3.4 Sex of the embryo

The availability of gDNA from CV gave me the opportunity to determine the sex of the embryo also in VTP samples and carry out the genetic association analyses according to sex stratification, using a OR analysis. In WBC, the SNP of gene *MTHFR* rs1801133 shows that the risk of RPL in T/T women turns out particularly high when embryo has a male sex (OR=3.64, IC95%=1.32-10.029, p=0.012), compared to female embryos(OR=1.66, IC95%=0.73-3.76, p=0.22)(Table 4.10). As for CV (Table 4.11), the heterozygous genotype C/T of the SNP rs1801133 shows increase of risk for RPL in female embryos (OR=2.28, IC95%=1.015-5.14, p=0.045).

Co-dominant model analysis in WBC								
	Embryo sex "Female"				*Embryo sex "Male"*			
SNP	***RPL***	***VTP***	***OR (IC 95%)***	***p-value***	***RPL***	***VTP***	***OR (IC 95%)***	***p-value***
MTHFR rs1801133								
DD	18	27	1.00 (Reference)		15	42	1.00 (Reference)	
DR	27	22	1.84(0.81-4.18)	0.14	19	32	1.66(0.73-3.76)	0.22
RR	12	9	2.00(0.69-5.71)	0.19	13	10	3.64(1.32-10.029)	**0.012**
MTHFR rs1801131								
DD	29	25	1.00 (Reference)		21	39	1.00 (Reference)	
DR	23	24	0.82(0.37-1.80)	0.63	16	34	0.87(0.39-1.93)	0.74
RR	5	9	0.47(0.14-1.61)	0.23	9	11	1.52(0.54-4.24)	0.42
HLA-G 14bp								
DD	13	15	1.00 (Reference)		12	25	1.00 (Reference)	
DR	34	30	1.30(0.53-3.18)	0.55	25	39	1.33(0.56-3.13)	0.50
RR	10	11	1.048(0.33-3.25)	0.93	10	19	1.096(0.39-3.070)	0.86
ANXA5 rs1050606								
DD	13	13	1.00(Reference)		14	24	1.00 (Reference)	
DR	29	25	1.16(0.45-2.95)	0.76	20	40	0.85(0.36-2.00)	0.72
RR	15	18	0.83(0.29-2.33)	0.72	12	19	1.082(0.40-2.87)	0.87
NKG2D rs2617170								
DD	24	25	1.00(Reference)		18	33	1.00 (Reference)	
DR	28	24	1.21(0.55-2.65)	0.62	22	40	1.0083(0.46-2.18)	0.98
RR	5	8	0.65(0.18-2.72)	0.50	5	12	0.76(0.23-2.51)	0.65
IL-10 rs1800871								
DD	34	31	1.00(Reference)		22	36	1.00 (Reference)	
DR	22	25	0.80(0.37-1.70)	0.56	22	40	0.90(0.42-1.89)	0.78
RR	1	2	0.45(0.039-5.27)	0.52	2	5	0.65(0.11-3.66)	0.62
CTLA-4 rs231775								
DD	26	26	1.00(Reference)		22	41	1.00 (Reference)	
DR	25	25	1.00(0.46-2.17)	1.00	18	39	0.86(0.40-1.84)	0.69
RR	5	7	0.71(0.20-2.54)	0.60	6	3	3.72(0.84-16.36)	0.081
SMTH1 rs1979277								
DD	29	32	1.00(Reference)		28	45	1.00 (Reference)	
DR	27	20	1.48(0.69-3.20)	0.30	17	31	0.88(0.41-1.87)	0.74
RR	1	4	0.27(0.029-2.61)	0.26	1	5	0.32(0.035-2.89)	0.31
PLK4 rs2305957								
DD	32	40	1.00(Reference)		25	47	1.00(Reference)	
DR	20	12	2.083(0.88-4.89)	0.091	21	28	1.41(0.66-2.97)	0.36
RR	4	3	1.66(0.34-7.99)	0.52	NA	6	NA	NA

TABLE 4.10: **Odds ratio and confidence intervals** with p-value calculated under a model of co-dominance for the heterozygous (**DR**) and homozygous recessive (**RR**) genotypes at the considered genetic loci in relation on embryo's sex in WBC.

Co-dominant model analysis in CV								
	Embryo sex "Female"				*Embryo sex "Male"*			
SNP	***RPL***	***VTP***	***OR (IC 95%)***	***p-value***	***RPL***	***VTP***	***OR (IC 95%)***	***p-value***
MTHFR rs1801133								
DD	21	32	1.00 (Reference)		13	28	1.00 (Reference)	
DR	27	18	2.28(1.015-5.14)	**0.045**	26	48	1.16(0.51-2.62)	0.71
RR	7	10	1.066(0.35-3.24)	0.90	7	9	1.67(0.51-5.48)	0.39
MTHFR rs1801131								
DD	26	29	1.00 (Reference)		22	42	1.00 (Reference)	
DR	22	24	1.022(0.46-2.23)	0.95	20	39	0.97(0.46-2.064)	0.95
RR	6	7	0.95(0.28-3.21)	0.94	4	4	1.90(0.43-8.37)	0.39
HLA-G 14bp								
DD	20	19	1.00 (Reference)		19	27	1.00 (Reference)	
DR	21	28	0.71(0.30-1.65)	0.43	22	37	0.84(0.38-1.86)	0.67
RR	14	11	1.20(0.44-3.31)	0.71	5	19	0.37(0.11-1.17)	0.092
ANXA5 rs1050606								
DD	19	14	1.00(Reference)		16	30	1.00 (Reference)	
DR	21	24	0.64(0.26-1.59)	0.34	16	35	0.85(0.36-2.00011)	0.72
RR	15	22	0.50(0.19-1.30)	0.15	14	20	1.31(0.52-3.27)	0.55
NKG2D rs2617170								
DD	29	27	1.00(Reference)		17	28	1.00 (Reference)	
DR	22	21	0.97(0.44-2.16)	0.95	27	43	1.034(0.47-2.23)	0.93
RR	4	12	0.31(0.089-1.080)	0.065	2	14	0.23(0.042-1.16)	0.076
IL-10 rs1800871								
DD	28	27	1.00(Reference)		26	38	1.00 (Reference)	
DR	20	27	0.71(0.32-1.56)	0.39	15	37	0.59(0.27-1.29)	0.18
RR	6	6	0.96(0.27-3.36)	0.95	5	6	1.21(0.33-4.41)	0.76
CTLA-4 rs231775								
DD	23	24	1.00(Reference)		22	34	1.00 (Reference)	
DR	28	32	0.91(0.42-1.96)	0.81	21	44	0.73(0.34-1.55)	0.42
RR	3	4	0.78(0.15-3.88)	0.76	3	6	0.77(0.17-3.41)	0.73
SMTH1 rs1979277								
DD	32	35	1.00(Reference)		29	44	1.00 (Reference)	
DR	19	19	1.093(0.49-2.42)	0.82	14	34	0.62(0.28-1.36)	0.23
RR	3	6	0.54(0.12-2.37)	0.41	2	5	0.60(0.11-3.34)	0.56
PLK4 rs2305957								
DD	28	28	1.00(Reference)		28	46	1.00 (Reference)	
DR	25	31	0.80(0.38-1.69)	0.57	16	38	0.69(0.32-1.46)	0.33
RR	2	1	2.00(0.17-23.33)	0.58	1	1	1.64(0.098-27.32)	0.72

TABLE 4.11: **Odds ratio and confidence intervals** with p-value calculated under a model of co-dominance for the heterozygous (**DR**) and homozygous recessive (**RR**) genotypes at the considered genetic loci in relation on embryo's sex in CV.

4.4 Analysis of methylation pattern of the LINE-1 sequence

In this chapter, I used a normality test called *"Shapiro Wilk test"* for check if my dates followed a normal distribution. The Shapiro-Wilk test showed that WBC samples in RPL (W=0.98, p-value= 0.64) and VTP (W=0.97, p-value= 0.080) followed a normal distribution of LINE-1 methylation. Instead, for CV samples, the RPL (W=0.89, p=3.96*10-5) and VTP (W=0.93, p=0.0006) not followed a normal distribution, as describe in table 4.12. The samples are aged matched for exclude a bias in the analyses: the age limit for embryo in RPL and VTP is 90 days.

I calculated the mean of RPL in WBC and the 77% (sd=0.41%) of CpG island is metilated. For CV of RPL, the mean is 60% (sd=1.0%). As control, in VTP I see that the 78% (sd=0.62%) of CpG island is metilated and in CV, the mean is 55% (sd=0.83%), as describe in table 4.13. The difference from two group in WBC is not significant (T student test p-value= 0.70). In CV, appears significant the fact that in the RPL, it seems more ipermethylated than control (Mann-Whitney p-value= 0.0010). The Figure 4.8 shows a box plot of the difference on methylation. On the left, there is the global methylation in WBC, in red for RPL and in blue for VTP. On the right, are represented the CV, with is visible more difference in mean of methylation.

After this evaluation, for the next analysis of stratification, an unpaired student t test was used for WBC that followed a normal distribution. For CV sample that not followed a normal distribution, I used a non parametric test called "*Mann-Whitney test*". The analyses done in this chapter follows the statistical analysis for LINE-1 methylation study of *Khan MFJ. et al.,2017* [71].

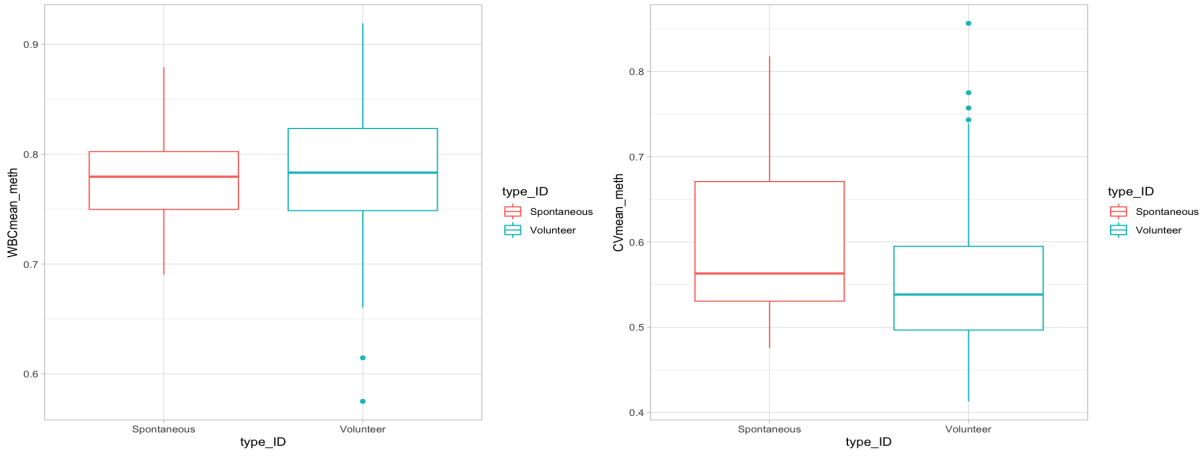

FIGURE 4.8: **Global methylation of LINE-1** in WBC on the left and CV on the right

Shapiro-Wilk test LINE-1 methylation							
RPL WBC		**RPL CV**		**VTP WBC**		**VTP CV**	
W	***p-value***	***W***	***p-value***	***W***	***p-value***	***W***	***p-value***
0.98	0.64	0.89	3.96*10-5	0.97	0.08	0.93	0.0006

TABLE 4.12: **Normality test for LINE-1 methylation** of RPL and VTP of WBC and CV

LINE-1 methylation	*T student test in WBC*		*Mann-Whitney test in CV*	
	Mean SD	**p-value**	**Mean SD**	**p-value**
RPL	77% ± 0.41%		60% ± 1.0%	
VTP	78% ± 0.62%		55% ± 0.83%	
RPL-VTP		*p= 0.70*		***p=0.0010***

TABLE 4.13: **Global LINE-1 methylation level:** Mean and standard deviation (SD) for WBC and CV in RPL and VTP.

4.4.1 Stratification of LINE-1 methylation analysis based on information from the Clinical Database

After the analysis of the global DNA methylation in LINE-1, the next step focus on the stratification of the DNA methylation with the different parameter present in clinical database, where there is a comparison between RPL and VTP samples of CV and WBC on the average of LINE-1 methylation level.

The parameters analyzed were the age of mothers, BMI of mothers, smoking exposure, folic acid intake and embryo age and sex (table 4.14). The criteria of selection of threshold are already describe in the previous chapter. For comparison of this parameters, the unpaired student's t-test for WBC and Mann-Whitney test for CV were performed.

LINE-1 methylation	*Analysis level of methylation in WBC*			*Analysis level of methylation in CV*		
	Mean ± SD	*Mean ± SD*	*p-value*	*Mean ± SD*	*Mean ± SD*	*p-value*
Mother's age 35 years	**>=35**	**<35**		**>=35**	**<35**	
RPL	77 % ±0.40%	78% ±0.40%	*p= 0.46*	61% ±0.94%	60% ±1.05%	*p=0.53*
VTP	76% ±0.53%	78% ±0.64%	*p=0.20*	54% ±0.78%	55% ±0.86%	*p=0.87*
Mother's BMI	***Normal weight***	***Overweight***		***Normal weight***	***Overweight***	
RPL	78% ±0.37%	76% ±0.48%	*p=0.080*	59% ±1.0%	62% ±1.0%	*p=0.40*
VTP	78% ±0.54%	76% ±0.70%	*p=0.20*	54% ±0.78%	55% ±0.76%	*p=0.51*
Smoking exposure	***"Yes"***	***"No"***		***"Yes"***	***"No"***	
RPL	78%±0.38%	77%±0.39%	*p=0.70*	61% ±1.0%	59% ±0.95%	*p=0.59*
VTP	79% ±0.46%	75 % ±0.69%	*p=0.060*	54% ±0.89%	54% ±0.71%	*p=0.79*
Embryo's age	**>70**	**<=70**		**>70**	**<=70**	
RPL	78% ±0.42%	77% ±0.41%	*p=0.36*	58% ±0.9%	62% ±1.0%	*p=0.06*
VTP	79% ±0.5%	76% ±0.6%	*p= 0.06*	53% ±0.9%	56% ±0.8%	*p=0.10*
Folic acid	***"Yes"***	***"No"***		**"Yes"**	**"No"**	
RPL	77% ±0.45%	77% ±0.38%	*p= 0.97*	57% ±0.81%	62% ±1.0%	***p= 0.030***
VTP	NA	NA	*NA*	NA	78% ±0.56%	*NA*
Embryo's sex	***"Female"***	***"Male"***		***"Female"***	***"Male"***	
RPL	78% ±0.41%	77%±0.45%	*p=0.76*	61% ±1.0%	60% ±0.80%	*p=0.86*
VTP	78% ±0.54%	77 % ±0.67%	*p=0.40*	56% ±1.0%	54% ±0.68%	*p=0.67*

TABLE 4.14: **Evaluation of LINE-1 methylation** in subsets based on the clinical data.

4.4.2 Age of mothers

After stratification on the age of mothers with a threshold of 35 years old, I don't found any significant difference in both WBC and CV, both in RPL and VTP. The average of methylation on WBC and CV is similar between the groups. In table 4.14, in the WBC, in the category with less of 35 years old, the mean of RPL is 78% (sd=0.40 %). The same category in VTP, the mean is 78% (sd=0.64%). In WBC over 35 years old, the mean of RPL is 77% (sd=0.40%), and a similar result for VTP with 76% of mean (sd=0.53%). The comparison inside the same group with unpaired student t test shows no significant result between RPL (t-student p-value=0.46). Also, for VTP the t-student test shows no significant result (p-value=0.20).

In CV, the mean of RPL over 35 years old is 61% (sd=0.94%) and for VTP is 54% (sd=0.78%).In the other category, the mean less of 35 years old for RPL is 60% (sd=1.05%) and for VTP is 55% (sd=0.86%). Inside the same group RPL of CV, there isn't significant result (Mann-Whitney test p-value= 0.53). Also, for CV of VTP for Mann-Whitney test shows no significant result (p-value=0.87).

4.4.3 BMI

Next parameter analyzed is on the BMI of the mothers, where no significant difference was found between WBC and CV groups. The normal BMI is in range of values of 18.5-24.9 and overweight over 24.9. In table 4.14 shows that in WBC, the mean of RPL in normal weight is 78% (sd=0.37%) and in overweight RPL the mean is of 76% (sd=0.48%). For the VTP, the mean of women with normal weight is 78% (sd=0.54%) and in overweight VTP the mean is of 76% (sd=0.70%). The difference inside the group RPL is not significant (t-student p-value= 0.080). Inside VTP is also is not significant (T student test p-value= 0.20).

In CV, the mean of RPL normal weight is 59% (sd=1.0%) and in RPL overweight, is 62 % (sd=1.0%). In VTP, the normal weight is 54% (sd=0.78%) and overweight is 55% (sd= 0.76%). In CV of RPL there isn't significant result (Mann-Whitney test p-value= 0.40). The same for VTP (Mann-Whitney test p-value= 0.51).

4.4.4 Smoking

Smoking exposure of mothers can influence methylation both in WBC and CV. I create two group: "No" with no exposure to smoke and "Yes" with exposure to smoke, also previous of pregnancy. As shows in table 4.14,in WBC, the mean of RPL in smoker is 78% (sd=0.38%) and in no-smoker RPL the mean is of 77% (sd=0.39%). Inside the group RPL the difference is not significant (t-student p-value= 0.70). In WBC of VTP, the mean in smoker is 79% (sd=0.46%) and in non-smoker is 75% (sd=0.69%). The difference inside VTP of WBC group is not significant (T student test p-value= 0.060).

In CV, the mean of RPL smoker is 61% (sd=1.0%) and the mean of RPL no smoker, is 59 % (sd=0.95%) so in CV there isn't a significant result (Mann-Whitney test p-value= 0.59). In VTP CV, the smoker is 54% (sd=0.89%) and in non smoker the mean is 54% (sd=0.71%) with no significant result (Mann-Whitney test p-value= 0.79).

4.4.5 Age of embryo

Afterwards, the analysis focus on the age of embryo, where the threshold used is 70 days old. Embryo's age was calculated as the difference between the date of the procedure and the last menstruation date. It needs to be considered that it is impossible to precisely know the time of the embryo death for RPL samples, nevertheless the difference we used is the best approximation. Embryos were subdivided in two groups using as threshold 70 days, because from literature it is known that the most miscarriages happens in first trimester at 10 weeks, and also embryo became fetus in that period [78]. The limit of emrbyo age is 90 days, as said before. As shows in table 4.14, in WBC more of 70 days, the mean of RPL is 78% (sd=0.42%) of CpG island is metilated against a mean of 77% (sd=0.41%) in WBC less 70 days. The difference from two group is not significant (t-student test p-value=0.36). The comparison inside the same group of VTP shows no significant result beetween VTP less (76%, sd=0.6%) and over 70 days(79%, sd=0.5%) (t-student test p-value=0.06).

In CV, the mean of RPL over 70 days is 58% (sd=0.9%) of CpG island is metilated and in RPL less of 70 days the mean of is 62% (sd=1.0%). The difference is again not significant (Mann-Whitney p-value= 0.06). Inside the same group VTP, in CV there isn't significant result between VTP over (53%, sd=0.9%) and less 70 days (56%, sd=0.8%) (Mann Whitney p-value= 0.10).

4.4.6 Folic acid intake

The folic acid (FA) intake is considered if women used folic acid during pregnancy or not. As shows in table 4.14, the mean of RPL in WBC with FA is 77% (sd=0.45%) against the RPL without intake of FA where mean is 77% (sd=0.38%), there is no difference (t-student p-value= 0.97).

In CV, the mean of RPL with FA is 57% (sd=0.81%) against the RPL without intake of FA where mean is 62% (sd=1.0%), there is a significant difference (Mann-Whitney p-value= 0.030), as shows in figure 4.9. In VTP, no women intakes folic acid, because the pregnancy probably was not planned, so I can't apply the Mann-Whitney test because I don't have any case of VTP with folic acid supplementation.

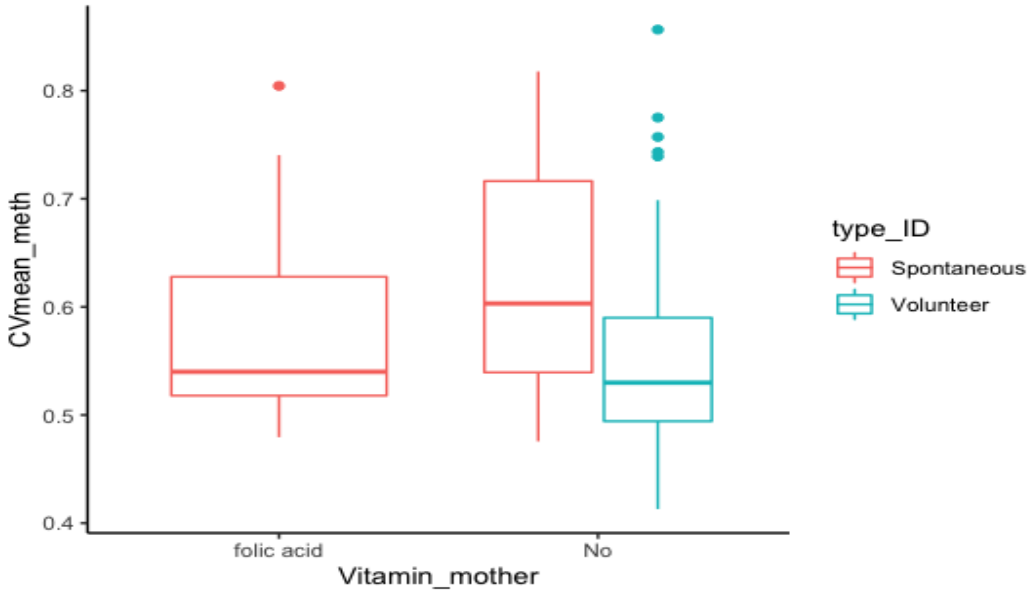

FIGURE 4.9: **Global methylation of LINE-1 in CV** based on stratification on folic acid intake in mothers during pregnancy

4.4.7 Sex of the embryo

The availability of gDNA from CV gave me the opportunity to determine the sex of the embryo. As shows in table 4.14, in WBC with female embryos, the mean of RPL is 78% (sd=0.41%) and in WBC with male embryos, the mean of RPL is 77% (sd=0.45%). The comparison shows no significant result between RPL (t-student p-value=0.76). In WBC of VTP, in the women with female embryos, the mean is 78% (sd=0.54%) and with male embryos the mean is 77% (sd=0.67%) with a significant p-value (T student test p-value=0.40).

In CV, the mean of RPL female is 61% (sd=1.0%) of CpG island is methylated and the mean of RPL male, is 60 % (sd=0.80%). Inside the same group RPL, in CV there isn't a significant result (Mann-Whitney p-value= 0.86). In the CV of VTP control, the mean of female embryos is 56% (sd=1.0%) and in male embryos, the mean is 54% (sd=0.68%) with non significant result (Mann-Whitney test p-value=0.67).

4.5 Analysis of LINE-1 methylation for MTHFR rs1801133 and rs1801131

The stratification present in this paragraph want to check the effect on the level of methylation in LINE-1 of the genotype of two different SNP of MTHFR on CV and WBC, both in RPL and VTP, to see if the genotype of that genes can influence the methylation (tab:4.15). From previous introduction, this genes are involved in metabolic folic acid cycle.
For WBC, there isn't a significant result for gene *MTHFR* rs1801133 in RPL WBC. I consider the dominant homozygous genotype as reference. The heterozygous C/T and recessive genotype T/T seems ipomethylated than homozygous genotype C/C (p=0.17, p=0.11). The level of methylation in heterozygous and recessive genotype is 79% (sd=0.75%, sd=0.41%) against 85% (sd=0.80%) in dominant genotype. In WBC of VTP, the comparison between wild type C/C with 84% (sd=0.72%) genotype and

heterozygous C/T 84% (sd=0.73%) do not shows any significant result (Test t p-value= 0.73). A significant result (Test t p-value=0.02) appears if I consider in the VTP control, the difference between wild type C/C 84% (sd=0.72%) and recessive genotype T/T 75% (sd=0.78%). In fact the recessive genotype T/T result ipomethylated then other genotype.

I don't observe any significant association when considering other SNP of *MTHFR* rs1801131 in all group.

	T student in WBC on LINE-1 methylation				***Mann-Whitney test in CV on LINE-1 methylation***			
SNP	***Mean ± SD RPL***	***p-value***	***Mean ± SD VTP***	***p-value***	***Mean ± SD RPL***	***p-value***	***Mean ± SD VTP***	***p-value***
rs1801133								
DD	85%* ± 0.80%		84%* ± 0.72%		61%* ± 1.0%		55%* ± 0.89%	
DR	79% ± 0.75%	p=0.17	84% ± 0.73%	p= 0.73	59% ± 0.90%	p=0.33	55% ± 0.74%	p= 0.86
RR	79% ± 0.41%	p=0.11	75% ± 0.78%	**p=0.02**	64% ± 0.80%	p=0.34	54% ± 1.0%	p=0.45
rs1801131								
DD	77%* ± 0.41%		78%* ± 0.58%		62%* ±0.98%		53%* ± 0.84%	
DR	79%* ± 0.42%	p=0.061	76 %± 0.62%	p=0.10	58% ±0.93%	p=0.062	57% ± 0.84%	p=0.06
RR	76%* ± 0.39 %	p=0.48	82% ± 0.54%	p=0.11	62% ± 1.0 %	p=0.98	55%± 0.64%	p=0.17

TABLE 4.15: **Evaluation of LINE-1 methylation** in CV and WBC based on MTHFR gene variants.

Chapter 5

Discussion

In this project I want to investigate genetics and epigenetics of recurrent pregnancy loss through the analysis of genetic markers from genes which have been demonstrated to play an important role in this pathology. This study is quite unique, as it is one of few which had investigated gDNA methylation in tissue from CV of recurrent spontaneous abortion cases using samples from VTP as reference controls.

5.1 Realization of a biobank of DNA extracted from chorionc villi

In the first part of the project, I realized a biobank through the processing of samples of chorionic villi. I researched the best way to process CV gDNA and found no significant difference between two methods for tissue homogenization and DNA extraction: from dry tissues or from tissues in a solution with lysis buffer. I also found that there are no significant difference on DNA quality when using three kits for DNA extraction available in commerce (Instagene Matrix from Bio-Rad, QIAamp DNA Blood Mini Kit from Qiagen, and Nucleon BACC1 kit from GE Healthcare UK). All this methods can be used for an extraction to obtained good quality of DNA from CV.

I analyzed the clinical data collected together with the biological samples. For many of them, are available anthropometric and medical information. There is no significant difference with a reference set of mothers that did voluntary termination of pregnancy (VTP), with the expected exception for the significantly higher age of mothers presenting recurrent pregnancy loss. This aspect has been fully demonstrated in literature to be correlated to increase the risk of RPL [9]. Notably I see a trend of higher education level correlated with advanced age, most likely because the age of pregnancy tend to be higher than other women for social reasons.

Finally, part of the collected samples were also used for another project, a quite unique study as it is one of few which had extracted and investigated whole-genome sequence of gDNA in tissue from CV. Prior to sequencing samples were screened for chromosomal aneuploidies, and it was found that the 70% of the samples has aneuploidies [79].

5.2 Genetic of miscarriage

This preliminary study suggests that risk of RPL is influenced by maternal's *MTHFR* c.677C>T genotypes. In particular, c.677T allele seems to have an effect on the risk of RPL, with triplicate risk in T/T homozygous women and half risk in heterozygous. I hypothesize that this evidence could be related to a reduced synthesis of 5- methyl-Tetrahydrofolate in T/T homozygotes [80], which could eventually dysregulates the methylation of genomic DNA, ultimately leading to epigenetic modifications, which could affect the survival of embryos. In fact, in my epigenetic study, related to comparison between *MTHFR*c.677C>T and global level of LINE-1, I see a trend of significance difference in methylation between wild type genotype and heterozygous and recessive homozygous genotype.

This finding is robust to stratification for several parameters. I observe that the T/T genotype is associated with increased risk of recurrent spontaneous abortion in mothers without intake of folic acid during pregnancy (p= 0.050) and without any exposure to smoke (p=0.026). This information is very interesting because probably the lack of the use of supplement of folic acid, increase the risk of RPL in women with T/T genotype due to less activity of MTHFR enzyme, as said before.

Taking into account the sex of embryos, the preliminary results for *MTHFR*c.677C>T suggest that the genotype-associated risk of RPL has different trends in the two sexes, both in WBC and CV. In particular, I observe increased risk in mothers with male embryos (p= 0.012) in comparison with mothers with female embryos (p= 0.19). Instead, if I consider the embryos, the female CV with C/T genotype, shows a significant increase of risk of RPL. For T/T female, no significant result,probably due to the small size of compared groups. This result can be related to the observation that the aborted fetuses from RPL with unexplained causes, were preferentially female [81]. Also the study of Del Fabro et al., 2011 [82], confirm that the female embryo might be more susceptible to RPL than male ones during embryogenesis, implantation, and initial fetal development. This evidence needs further investigations in a larger case-control cohort, especially because the allele T might be causes increase of risk RPL in population, especially in Italian population, if women are pregnant of female embryos. In fact, this could be related to increased prevalence of T/T genotype in Italian population, as shown by study of Wilcken et al., 2003 [83], where he saw in 12-15% of cases of T/T genotype in southern Europe (Spain and northern Italy), peaking in southern Italy (20-26% in Campania and Sicily).

The study of *McCoy et al., 2015* [45] discovered that *PLK4* rs2305957 gene play a role in the centriole duplication and can alter the result of mitosis through a deregulation of centriole areas and contribute to mitotic-origin aneuploidy risk during human early embryo development. Later, the study of *Zhang et al., 2017* [44], shows that the recessive genotype A/A was correlated with higher risk of RPL and it probably causes this deregulation during mitosis of centriole area. In my study for the *PLK4* rs2305957 gene, I observe a significant increased of the risk of RPL when I consider women with G/A genotype, with smoke exposure (p=0.010) and without folic acid intake during pregnancy (p=0.024). But in my study, only the heterozygous genotype G/A is significant correlated with higher increase of risk of RPL. The effect of A/A genotype, however, is not statistically significant probably due to the small size of compared groups. But it is clear that the allele A of this *PLK4* rs2305957 gene can

cause problem during early embryo development.

In my study, an important role in reduce the risk of RPL in the embryo, seems to play the *NKG2D* gene (rs2617170). In the analysis of ORs based on alleles, the recessive allele A gives protection (p=0.025). I can see the same effect on recessive genotype A/A where the effect of reduction of RPL be increase (p=0.0067). When the analysis is stratified on clinical parameters, I find the same protective effect on CV with mothers without any exposure to smoke (p=0.0029) and normal weight (p=0.015). In a study from Hizem et al.,2014 [39], the recessive genotype A/A in mothers is correlated with higher cytotoxic activity. Here for the first time I observe the effects on embryos finding is opposite, i.e. protection from RPL.

In a normal pregnancy, the role of *NKG2D* is to stimulate decidual NK cells to produce a type of cytokines that induce angiogenesis, such as vascular endothelial that contributes to the uterine vascular re-modeling. Indeed, the fetus itself is never in direct contact with uterine tissues and maternal blood. The role to invade the decidua and form the placenta is of fetal trophoblast and the cytokines, secreted by maternal and fetal cells at the site of implantation, contribute to maintain the communication between trophoblast and decidual cells. This help to maintain in balance the fetal interface and the maternal immune system in normal pregnancy.

Syncytiotrophoblast cells produce and secrete NKG2D ligands via exosomes. The NKG2D ligand-loaded exosomes in normal pregnancy serum is able to interact with NKG2D and downregulate the receptor expression on peripheral blood mononuclear cells (PBMC), with the following inhibition of the NKG2D-dependent cytotoxic response. The presence of this genotype A/A in embryos, probably can contribute to maintain this secretion by Syncytiotrophoblast cells [84].

5.3 Epigenetic in miscarriage

Epigenetic refers to the study of alterations in gene function that are mitotically and/or meiotically heritable and that do not entail a change in the DNA sequence. Epigenetic mechanisms are essential for development and differentiation, but can be disrupted by exogenous factors [58]. Scientific evidence increasingly suggests that exposures during the intrauterine period can increase the danger of developing disease in later life [60].

In this study, I examined the DNA methylation in my samples of gDNA from both in WBC and CV. I want to study DNA methylation because I want to understand if the life style of the mothers (e.g. smoke, alcool, folic acid intake) can influence the environment of pregnancy. In particular, I want to understand if a bad life style can cause RPL. The importance of epigenetic is that, became possible to act and change bad life style of mothers, while it is impossible to do the same on genetic because are stable elements.

I evaluated the role of epigenetic in RPL through the analysis of LINE-1 methylation. Because of their high genome dissemination, LINE-1 methylation is a surrogate marker of global DNA methylation level. This study is quite unique, as it is one of

few which had investigated gDNA methylation in tissue from CV of recurrent pregnancy loss cases.

The preliminary results show no significant LINE-1 methylation difference between gDNA from RPL and VTP in mother's white blood cells (WBC). However, a significant increase of 5% higher methylation level (p-value 0.0010) can be seen in RPL cases compared to VTP cases when considering gDNA from chorionic villi. My results, although preliminary, provide for the first time evidence of an increased level of LINE-1 methylation in RPL cases, suggesting that altered epigenetics of the embryo could play a role in determining the recurrent pregnancy loss.

This difference was not influenced by mother's age, smoking exposure and BMI or embryo's sex and age. Among VTP cases the stratification for FA supplementation exposure was not possible, due to the lack of positive cases (only one VTP case was supplemented with FA).

Folate is an important precursor of one-carbon units required for DNA methylation. Thus, folate metabolism has been suggested to influence epigenetic alterations in cancer. High folic acid might contribute to the maintenance of global methylation through an equal supply of one-carbon units for the methylation machinery, thereby stabilizing the genome [61]. In my study an effect was observed when considering the exposure to periconceptional supplementation with folic acid (FA) (400 g/day). As shown in Table 4.14, the LINE-1 DNA methylation among RPL cases in CV is significantly higher in cases not supplemented with FA (p-value = 0.030). Among VTP cases the stratification for FA supplementation exposure was not possible, due to the lack of positive cases. In this study, the result of epigenetic analysis is indicating a possible role of folate bioavailability in causing RPL, and suggesting that intervention with folate supplementation could be beneficial in preventing this pathology. In addition, to support this thesis, from stratification on different SNP, in WBC, the *MTHFR* rs1801133 shows a trend of difference in mean of methylation among RPL cases, where the comparison between wild type C/C and the heterozygous and recessive genotype T/T shows a decrease of methylation in recessive genotype of 6% (p=0.17, p=0.11). The same patterns of demethylation is seen in VTP and it became significant (p=0.01), with a decrease of methylation of 9% in recessive genotype in VTP. This can be related to minor efficiency of enzyme activity in T/T genotype, so the lower level of methylation increase activity of the gene.

Chapter 6

Conclusion

This preliminary study indicate that risk of RPL is influenced by maternal *MTHFR* c.677C>T genotypes, suggesting a possible interaction between embryo's sex and *MTHFR* genotypes. In conclusion this study suggests a possible causal role due to unbalanced folate metabolism, and lays the groundwork for innovative measures aimed at preventing recurrence of spontaneous abortion.
Notably, the mothers in which I find significant association are in healthy, have normal BMI, and no exposition to smoke. This is important because I want to focus on this type of women, without an apparent cause of miscarriage.
In the embryos, available for the first time, I see a protective effect at the *NKG2D* gene. This study give the important information that it is important to study, not only the maternal genome, but also the fetal genome, an aspect that has been neglected until now.

My results, although preliminary, provide for the first time evidence of an increased level of LINE-1 methylation in RPL cases, suggesting that altered epigenetics of the embryo could play a role in determining the recurrent pregnancy loss. Moreover, I provide first evidence of an influence of mother's periconceptional exposure to folic acid supplementation on the global genomic methylation level of embryo's chorionic villi, indicating a possible role of folate bioavailability in causing RPL, and suggesting that intervention with folate supplementation could be beneficial in preventing this pathology.

As future direction, this study shows the importance to focus on the trios (mother, father and fetus) during pregnancy in order to prevent miscarriage. Acting on mother life style where this are not suitable for good result of pregnancy. This results lay the foundations for identifying an effective protocol for the treatment of women with recurrent spontaneous abortion without apparent cause, that currently is absent.

Bibliography

[1] Hebe Campaña et al. "Association between a maternal history of miscarriages and birth defects". In: *Birth defects research* 109.4 (2017), pp. 254–261.

[2] Boran Du et al. "Polymorphisms of methalenetetrahydrofolate reductase in recurrent pregnancy loss: an overview of systematic reviews and meta-analyses". In: *Journal of assisted reproduction and genetics* (2019), pp. 1–14.

[3] Neha Sudhir et al. "Cytogenetic analysis in couples with recurrent miscarriages: a retrospective study from Punjab, north India". In: *Journal of genetics* 95.4 (2016), pp. 887–894.

[4] A Oliver and C Overton. "Diagnosis and management of miscarriage." In: *The Practitioner* 258.1771 (2014), pp. 25–8.

[5] Sezin YAkuT et al. "Chromosome abnormalities identified in 457 spontaneous abortions and their histopathological findings/457 spontan abortus materyalinde tespit edilen kromozomal anomaliler ve histopatolojik bulguları". In: *Turkish Journal of Pathology* 31.2 (2015), pp. 111–118.

[6] ISTAT. *La salute riproduttiva della donna*. 2017.

[7] M.A. Hil. *Placenta development*. https://embryology.med.unsw.edu.au/embryology/index.php/Lecture_-_Placenta_Development.2020./. 2020.

[8] Livio Zanoio, Eliana Barcellona, and Gabrio Zacché. *Ginecologia e ostetricia*. Elsevier Health Sciences, 2013.

[9] Maria C Magnus et al. "Role of maternal age and pregnancy history in risk of miscarriage: prospective register based study". In: *bmj* 364 (2019), p. l869.

[10] Ana Vičić et al. "Trisomy 1 in an early pregnancy failure". In: *American Journal of Medical Genetics Part A* 146.18 (2008), pp. 2439–2441.

[11] Anna Kenyon and Susan Bewley. "The effect of age on obstetric (maternal and fetal) outcomes". In: *Reproductive ageing. RCOG Press, London* (2009), pp. 125–40.

[12] Narges Kamkar, Fariba Ramezanali, and Marjan Sabbaghian. "The relationship between sperm DNA fragmentation, free radicals and antioxidant capacity with idiopathic repeated pregnancy loss". In: *Reproductive biology* 18.4 (2018), pp. 330–335.

[13] Channa N Jayasena et al. "Reduced testicular steroidogenesis and increased semen oxidative stress in male partners as novel markers of recurrent miscarriage". In: *Clinical chemistry* 65.1 (2019), pp. 161–169.

[14] Merel MJ van den Berg et al. "Genetics of early miscarriage". In: *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* 1822.12 (2012), pp. 1951–1959.

[15] Francisco Arredondo and Luis S Noble. "Endocrinology of recurrent pregnancy loss". In: *Seminars in reproductive medicine*. Vol. 24. 01. Copyright© 2006 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New ... 2006, pp. 033–039.

[16] Howard JA Carp, Carlo Selmi, and Yehuda Shoenfeld. "The autoimmune bases of infertility and pregnancy loss". In: *Journal of autoimmunity* 38.2-3 (2012), J266–J274.

[17] Nagwa EA Gaboon. "Recurrent spontaneous abortion: an overview of genetic backgrounds and impact of male factors: a review". In: *International Journal of Human Genetics* 13.2 (2013), pp. 79–83.

[18] Ineke Krabbendam et al. "Thrombophilias and recurrent pregnancy loss: a critical appraisal of the literature". In: *European Journal of Obstetrics & Gynecology and Reproductive Biology* 118.2 (2005), pp. 143–153.

[19] Deborah B Nelson et al. "First trimester levels of BV-associated bacteria and risk of miscarriage among women early in pregnancy". In: *Maternal and child health journal* 19.12 (2015), pp. 2682–2687.

[20] Corey M Benedum et al. "Risk of spina bifida and maternal cigarette, alcohol, and coffee use during the first month of pregnancy". In: *International journal of environmental research and public health* 10.8 (2013), pp. 3263–3281.

[21] Fiorella Miceli et al. "Effects of nicotine on human luteal cells in vitro: a possible role on reproductive outcome for smoking women". In: *Biology of reproduction* 72.3 (2005), pp. 628–632.

[22] Roberta B Ness et al. "Cocaine and tobacco use and the risk of spontaneous abortion". In: *New England Journal of Medicine* 340.5 (1999), pp. 333–339.

[23] Hamid Reza Nakhai-Pour et al. "Use of nonaspirin nonsteroidal anti-inflammatory drugs during pregnancy and the risk of spontaneous abortion". In: *Cmaj* 183.15 (2011), pp. 1713–1720.

[24] Anne-Marie Nybo Andersen et al. "Maternal age and fetal loss: population based register linkage study". In: *Bmj* 320.7251 (2000), pp. 1708–1712.

[25] Tae Yeong Choi et al. "Spontaneous abortion and recurrent miscarriage: A comparison of cytogenetic diagnosis in 250 cases". In: *Obstetrics & gynecology science* 57.6 (2014), pp. 518–525.

[26] María del Rosario Rodríguez-Guillén et al. "Maternal MTHFR polymorphisms and risk of spontaneous abortion". In: *Salud publica de Mexico* 51 (2009), pp. 19–25.

[27] Fabio Coppedè. "The complex relationship between folate/homocysteine metabolism and risk of Down syndrome". In: *Mutation Research/Reviews in Mutation Research* 682.1 (2009), pp. 54–70.

[28] Rémy Slama et al. "Influence of paternal age on the risk of spontaneous abortion". In: *American Journal of Epidemiology* 161.9 (2005), pp. 816–823.

[29] Jana L Allison and Danny J Schust. "Recurrent first trimester pregnancy loss: revised definitions and novel causes". In: *Current Opinion in Endocrinology, Diabetes and Obesity* 16.6 (2009), pp. 446–450.

[30] Elizabeth E Puscheck and Rajasingam S Jeyendran. "The impact of male factor on recurrent pregnancy loss". In: *Current Opinion in Obstetrics and Gynecology* 19.3 (2007), pp. 222–228.

[31] Isabelle Bray, David Gunnell, and George Davey Smith. "Advanced paternal age: How old is too old?" In: *Journal of Epidemiology & Community Health* 60.10 (2006), pp. 851–853.

[32] Siaw-Cheok Liew and Esha Das Gupta. "Methylenetetrahydrofolate reductase (MTHFR) C677T polymorphism: epidemiology, metabolism and the associated diseases". In: *European journal of medical genetics* 58.1 (2015), pp. 1–10.

[33] Nam Keun Kim et al. "Influence of combined methylenetetrahydrofolate reductase (MTHFR) and thymidylate synthase enhancer region (TSER) polymorphisms to plasma homocysteine levels in Korean patients with recurrent spontaneous abortion". In: *Thrombosis research* 117.6 (2006), pp. 653–658.

[34] Izabela Nowak et al. "Possible role of HLA-G, LILRB1 and KIR2DL4 gene polymorphisms in spontaneous miscarriage". In: *Archivum immunologiae et therapiae experimentalis* 64.6 (2016), pp. 505–514.

[35] Fateme Arjmand et al. "The balance of the immune system between HLA-G and NK cells in unexplained recurrent spontaneous abortion and polymorphisms analysis". In: *Immunologic research* 64.3 (2016), pp. 785–790.

[36] SY Al Omar and L Mansour. "Association of HLA-G 14-base pair insertion/deletion polymorphism with breast cancer in Saudi Arabia". In: ().

[37] Yuko Hayashi et al. "Genotyping analyses for polymorphisms of ANXA5 gene in patients with recurrent pregnancy loss". In: *Fertility and sterility* 100.4 (2013), pp. 1018–1024.

[38] Bei Zhang et al. "The association between annexin A5 (ANXA5) gene polymorphism and left ventricular hypertrophy (LVH) in Chinese endogenous hypertension patients". In: *Medicine* 96.44 (2017).

[39] Sondes Hizem et al. "Polymorphisms in Genes Coding for the NK-Cell Receptor NKG 2D and its Ligand MICA in Recurrent Miscarriage". In: *American Journal of Reproductive Immunology* 72.6 (2014), pp. 577–585.

[40] Eva Cochery-Nouvellon et al. "Interleukin 10 gene promoter polymorphisms in women with pregnancy loss: preferential association with embryonic wastage". In: *Biology of reproduction* 80.6 (2009), pp. 1115–1120.

[41] RH Qaddourah et al. "IL-10 gene promoter and intron polymorphisms and changes in IL-10 secretion in women with idiopathic recurrent miscarriage". In: *Human reproduction* 29.5 (2014), pp. 1025–1034.

[42] Maneesh Kumar Misra et al. "Association of functional genetic variants of CTLA4 with reduced serum CTLA4 protein levels and increased risk of idiopathic recurrent miscarriages". In: *Fertility and sterility* 106.5 (2016), pp. 1115–1123.

[43] Stephanie M Engel et al. "Polymorphisms in folate metabolizing genes and risk for spontaneous preterm and small-for-gestational age birth". In: *American journal of obstetrics and gynecology* 195.5 (2006), 1231–e1.

[44] Qian Zhang et al. "Maternal common variant rs2305957 spanning PLK4 is associated with blastocyst formation and early recurrent miscarriage". In: *Fertility and sterility* 107.4 (2017), pp. 1034–1040.

[45] Rajiv C McCoy et al. "Common variants spanning PLK4 are associated with mitotic-origin aneuploidy in human embryos". In: *Science* 348.6231 (2015), pp. 235–238.

[46] Per Magne Ueland et al. "Biological and clinical implications of the MTHFR C677T polymorphism". In: *Trends in pharmacological sciences* 22.4 (2001), pp. 195–201.

[47] Philippe Goyette et al. "Human methylenetetrahydrofolate reductase: isolation of cDNA, mapping and mutation identification". In: *Nature genetics* 7.2 (1994), pp. 195–200.

[48] Saeedeh Salimi et al. "The early-onset preeclampsia is associated with MTHFR and FVL polymorphisms". In: *Archives of gynecology and obstetrics* 291.6 (2015), pp. 1303–1312.

[49] Cristina Camprubí et al. "Lack of association of MTHFR rs1801133 polymorphism and CTCFL mutations with sperm methylation errors in infertile patients". In: *Journal of assisted reproduction and genetics* 30.9 (2013), pp. 1125–1131.

[50] Arijit Biswas et al. "Recurrent abortions in Asian Indians: no role of factor V Leiden Hong Kong/Cambridge mutation and MTHFR polymorphism". In: *Clinical and Applied Thrombosis/Hemostasis* 14.1 (2008), pp. 102–104.

[51] ZJ Foka et al. "Factor V Leiden and prothrombin G20210A mutations, but not methylenetetrahydrofolate reductase C677T, are associated with recurrent miscarriages". In: *Human Reproduction* 15.2 (2000), pp. 458–462.

[52] Duo Su et al. "Association of interleukin-10 gene promoter polymorphisms with recurrent miscarriage: a meta-analysis". In: *American Journal of Reproductive Immunology* 76.2 (2016), pp. 172–180.

[53] Chongjuan Gu et al. "Association of interleukin-10 gene promoter polymorphisms with recurrent pregnancy loss: a meta-analysis". In: *Journal of assisted reproduction and genetics* 33.7 (2016), pp. 907–917.

[54] Juanjuan Zhang et al. "The synergic effects of CTLA-4/Foxp3-related genotypes and chromosomal aberrations on the risk of recurrent spontaneous abortion among a Chinese Han population". In: *Journal of human genetics* 63.5 (2018), pp. 579–587.

[55] Fabio Coppedè. "The genetics of folate metabolism and maternal risk of birth of a child with Down syndrome and associated congenital heart defects". In: *Frontiers in genetics* 6 (2015), p. 223.

[56] Abee L Boyles et al. "Neural tube defects and folate pathway genes: family-based association tests of gene–gene and gene–environment interactions". In: *Environmental health perspectives* 114.10 (2006), pp. 1547–1552.

[57] Jill A McKay et al. "Genetic and non-genetic influences during pregnancy on infant global and site specific DNA methylation: role for folate gene variants and vitamin B12". In: *PloS one* 7.3 (2012), e33290.

[58] Lode Godderis et al. "Global methylation and hydroxymethylation in DNA from blood and saliva in healthy volunteers". In: *BioMed research international* 2015 (2015).

[59] D Perrin et al. "Specific hypermethylation of LINE-1 elements during abnormal overgrowth and differentiation of human placenta". In: *Oncogene* 26.17 (2007), pp. 2518–2524.

[60] Carrie V Breton et al. "Prenatal tobacco smoke exposure affects global and gene-specific DNA methylation". In: *American journal of respiratory and critical care medicine* 180.5 (2009), pp. 462–467.

[61] MingJi Jin et al. "Different histological types of non-small cell lung cancer have distinct folate and DNA methylation levels". In: *Cancer science* 100.12 (2009), pp. 2325–2330.

[62] Antonella Agodi et al. "Low fruit consumption and folate deficiency are associated with LINE-1 hypomethylation in women of a cancer-free population". In: *Genes & nutrition* 10.5 (2015), p. 30.

[63] Daniel Ardeljan et al. "The human long interspersed element-1 retrotransposon: an emerging biomarker of neoplasia". In: *Clinical chemistry* 63.4 (2017), pp. 816–822.

[64] Zhi-ming He et al. "Transition of LINE-1 DNA methylation status and altered expression in first and third trimester placentas". In: *PloS one* 9.5 (2014).

[65] Svetlana Fa et al. "Assessment of global DNA methylation in the first trimester fetal tissues exposed to maternal cigarette smoking". In: *Clinical epigenetics* 8.1 (2016), p. 128.

[66] Alexander J Nash et al. "Interaction between plasma homocysteine and the MTHFR c. 677C> T polymorphism is associated with site-specific changes in DNA methylation in humans". In: *The FASEB Journal* 33.1 (2019), pp. 833–843.

[67] Jyoti Mishra et al. "Differential global and MTHFR gene specific methylation patterns in preeclampsia and recurrent miscarriages: A case-control study from North India". In: *Gene* 704 (2019), pp. 68–73.

[68] Lee P Shulman et al. "Direct analysis of uncultured cytotrophoblastic cells from second-and third-trimester placentas: An accurate and rapid method for detection of fetal chromosome abnormalities". In: *American journal of obstetrics and gynecology* 163.5 (1990), pp. 1606–1609.

[69] Steven E Bates. "Classical cytogenetics: karyotyping techniques". In: *Human Pluripotent Stem Cells*. Springer, 2011, pp. 177–190.

[70] Erick C Castelli et al. "Insights into HLA-G genetics provided by worldwide haplotype diversity". In: *Frontiers in immunology* 5 (2014), p. 476.

[71] Mohammad Faisal J Khan et al. "Evaluating LINE-1 methylation in cleft lip tissues and its association with early pregnancy exposures". In: *Epigenomics* 10.1 (2018), pp. 105–113.

[72] Colin Delaney, Sanjay K Garg, and Raymond Yung. "Analysis of DNA methylation by pyrosequencing". In: *Immunosenescence*. Springer, 2015, pp. 249–264.

[73] I Stanghellini et al. "Quantitation of fetal DNA in maternal serum during the first trimester of pregnancy by the use of a DAZ repetitive probe". In: *Molecular human reproduction* 12.9 (2006), pp. 587–591.

[74] Geraldine M Clarke et al. "Basic statistical analysis in genetic case-control studies". In: *Nature protocols* 6.2 (2011), pp. 121–133.

[75] Jan Graffelman. "Exploring diallelic genetic markers: the hardy weinberg package". In: *Journal of Statistical Software* 64.3 (2015), pp. 1–23.

[76] Lyndsay Ammon Avalos, Claudia Galindo, and De-Kun Li. "A systematic review to calculate background miscarriage rates using life table analysis". In: *Birth Defects Research Part A: Clinical and Molecular Teratology* 94.6 (2012), pp. 417–423.

[77] Carmen Garrido-Gimenez and Jaume Alijotas-Reig. "Recurrent miscarriage: causes, evaluation and management". In: *Postgraduate medical journal* 91.1073 (2015), pp. 151–162.

[78] F Cunningham et al. *Williams obstetrics, 24e*. Mcgraw-hill, 2014.

[79] M Goddijn and NJ Leschot. "Genetic aspects of miscarriage". In: *BEST PRACTICE AND RESEARCH IN CLINICAL OBSTETRICS AND GYNAECOLOGY* 14.5 (2000), pp. 855–865.

[80] Paul F Jacques et al. "Relation between folate status, a common mutation in methylenetetrahydrofolate reductase, and plasma homocysteine concentrations". In: *Circulation* 93.1 (1996), pp. 7–9.

[81] Takashi Kano, Takahide Mori, and Akinori Kimura. "Gender ratio distortion in abortuses and live births from patients with recurrent spontaneous abortion". In: *American Journal of Reproductive Immunology* 62.3 (2009), pp. 125–127.

[82] Anna Del Fabro et al. "Fetal gender ratio in recurrent miscarriages". In: *International journal of women's health* 3 (2011), p. 213.

[83] B Wilcken et al. "Geographical and ethnic variation of the 677C> T allele of 5, 10 methylenetetrahydrofolate reductase (MTHFR): findings from over 7000 newborns from 16 areas world wide". In: *Journal of medical genetics* 40.8 (2003), pp. 619–625.

[84] Bing Mei et al. "Defects in NKG2D ligand expression result in failed tolerance induction at the maternal–fetal interface: A possible cause for recurrent miscarriage". In: *Medical hypotheses* 79.4 (2012), pp. 465–467.