FACULTY OF MEDICINE AND HEALTH SCIENCES

SCREENING FOR GENETIC ABNORMALITIES OF INFERTILITY

Femke Machtelinckx
Student number: 01603646

Supervisor(s): Prof. Björn Heindryckx, Ph.D. Annekatrien Boel, M.Sc. Arantxa Cardona Barberán

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Master of Medicine in Medicine

Academic year: 2019 – 2021

GHENT UNIVERSITY

FACULTY OF MEDICINE AND HEALTH SCIENCES

SCREENING FOR GENETIC ABNORMALITIES OF INFERTILITY

Femke Machtelinckx
Student number: 01603646

Supervisor(s): Prof. Björn Heindryckx, Ph.D. Annekatrien Boel, M.Sc. Arantxa Cardona Barberán

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Master of Medicine in Medicine

Academic year: 2019 – 2021

GHENT UNIVERSITY

Preface

I would like to thank my promotor prof. Björn Heindryckx and co-promotors Annekatrien Boel and Arantxa Cardona Barberán for their guidance and support during the last two years. I truly enjoyed our collaboration. Furthermore, I would like to thank the Faculty of Medicine and Health Sciences and the G-FaST laboratory for enabling this research. A last word of gratitude goes to my family and friends for supporting me during these years writing the thesis.

Before this thesis, I already had a strong interest in genetics and reproduction. I was definitely delighted when I found out I would be able to work with this subject. The last few months have only enriched that fascination. I can hereby proudly present this thesis, in which we screened patients with a history of infertility for genetic abnormalities.

Table of contents

List of abbreviations

A	Adenine
Ala	Alanine
AOA	Assisted oocyte activation
ART	Assisted reproductive technology
BiHM	Biparental hydatidiform moles
Bp	Base pairs
BR	Blastocyst rate
BWS	Beckwith-Wiedemann syndrome
C	Cytosine
Ca^{2+}	Calcium
Cas9	CRISPR-associated protein 9
cDNA	Complementary DNA
CPL	Cytoplasmic lattice
CRISPR	Clustered regularly interspaced short palindromic repeat
cRNA	Complementary ribonucleic acid
DAG	Diacylglycerol
ddNTPs	Dideoxynucleotides
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotides
dup	Duplication
EDA	Embryo developmental arrest
ER	Endoplasmic reticulum
F-actin	Filamentous actin
FAF1	Fas-associated protein factor
FF	Fertilization failure
FR	Fertilization rate
G	Guanine
G-FaST	Ghent-Fertility and Stem cell Team
Gln	Glutamine
His	Histidine
HOCA	Human oocyte calcium analysis
ICSI	Intracytoplasmic sperm injection
ins	Insertion
IP_3(R)	Inositol 1,4,5- trisphosphate (receptor)
IVF	In vitro fertilization
KHDC3L	KH domain containing 3 like
Leu	Leucine
MAF	Minor allele frequency
MATER	Maternal antigen that embryos require
MEGs	Maternal-effect genes
Met	Methionine
MII	Metaphase II
MLID	Multilocus imprinting disorders
MOAT	Mouse oocyte activation test
MOCA	Mouse oocyte calcium analysis

mtDNA	Mitochondrial DNA
NACHT	NAIP (neuronal apoptosis inhibitor protein)
NGS	Next-generation sequencing
NLRPs	NOD-like receptors with pyrin domain
NLRP2	NLR family pyrin domain containing 2
NLRP5	NLR family pyrin domain containing 5
NLRP7	NLR family pyrin domain containing 7
NOD	Nucleotide-binding oligomerization domain
NT	Nuclear transfer
OAD	Oocyte activation deficiency
PADI6	Peptidyl arginine deiminase type 6
PCR	Polymerase chain reaction
PIP_2	Phosphoinositide 4,5-biphosphate
PKA	Protein kinase
PLCζ	Phospholipase C zeta
RHM	Recurrent hydatidiform moles
RNA	Ribonucleic acid
SCMC	Subcortical maternal complex
Ser	Serine
SNP	Single-nucleotide polymorphism
SOAF	Spermatozoon-released oocyte-activation factor
T	Thymine
Thr	Threonine
TLE6	Transducin-like enhancer of split 6
TUBB8	Tubulin Beta 8 Class VIII
VUS	Variant of uncertain significance
WES	Whole exome sequencing
WGS	Whole genome sequencing
WHO	World Health Organization

Abstract

Background

Infertility affects millions of couples worldwide. Although ICSI (intracytoplasmic sperm injection) has become very efficient, fertilization failure (FF) still occurs in 1-5% of the cycles. This is largely due to an insufficient oocyte activation, regulated by the male sperm factor PLCζ (phospholipase C ζ), resulting in none or few embryos available for transfer. PLCζ induces Ca^{2+} oscillations in the oocyte, essential for fertilization. Defects can be overcome using assisted oocyte activation (AOA) which induces artificial calcium oscillations during ICSI. Alternatively, embryo developmental arrest (EDA) can occur before the blastocyst stage despite good fertilization rates. This is regulated by the subcortical maternal complex (SCMC), containing NLRP2, NLRP5, NLRP7, PADI6, TLE6 and KHDC3L proteins. Mutations in *PLCZ1* have been associated with FF after ICSI, and mutations in the *SCMC* genes with EDA.

Objective

The objectives of this study are to broaden the knowledge about possible mutations in these genes and their relation to these specific infertility phenotypes.

Methods

The DNA of 23 couples with EDA was extracted from saliva samples. The maternal genes *NLRP2, NLRP5, NLRP7, PADI6, TLE6* and *KHDC3L* were screened using Next Generation Sequencing, as well as the paternal gene *PLCZ1* to exclude a sperm-related deficiency. Interesting variants were classified using VarSome and confirmed with Sanger sequencing.

Results

In *PLCZ1*, the homozygous variant p.S500L was found, classified as a variant of uncertain significance (class 3) in one patient. This specific mutation was also reported in recent research, causing fertilization failure after routine ICSI. Surprisingly, this patient also showed embryonic development arrest, in contrast to the fertilization failure phenotype previously reported in patients with this mutation. The phenotype interpretation remains unclear, but might contribute to proof of causality of this mutation. Since a PLCζ defect is the underlying cause of infertility, the patient qualifies for AOA treatment. In the SCMC genes, five newly found mutations of uncertain significance were found (in *PADI6*, *NLRP2*, *NLRP5*, *NLRP7* and *TLE6*) in five female patients. None of these were previously reported, complicating phenotype interpretation. More functional analysis would be required to enable a better understanding.

Conclusions

This study did not prove pathogenicity of these mutations, but contributes to a wider knowledge of the genetic factors associated with infertility, more specific *PLCZ1* and the SCMC genes.

Samenvatting

Achtergrond

Infertiliteit raakt miljoenen koppels wereldwijd. Intracytoplasmatische sperma-injectie (ICSI) is reeds zeer efficiënt, maar bevruchting faalt in 1-5%, vaak door ontoereikende eicelactivatie. Eicelactivatie staat onder controle van spermafactor PLCζ die intracellulaire Ca^{2+} oscillaties induceert, essentieel voor bevruchting. Defecten in PLCζ worden omzeild met geassisteerde eicel activatie (AOA) therapie, door inductie van artificiële calcium oscillatie. Ook kan embryo-ontwikkeling stoppen in het blastocyststadium ondanks goede bevruchtingcijfers na ICSI. Het subcorticale maternale complex (SCMC) speelt hierin een rol en bestaat uit NLRP2, NLRP5, NLRP7, PADI6, TLE6 en KHDC3L. Mutaties in het gen *PLCZ1* zijn geassocieerd met falen van bevruchting en mutaties in de SCMC genen met falen van embryo-ontwikkeling (EDA).

Doelstelling

Het doel van deze studie is de bijdrage aan een bredere kennis over mogelijke mutaties in deze genen en hun relatie met deze specifieke fenotypes van infertiliteit.

Methode

Het DNA van 23 koppels met EDA werd geëxtraheerd uit speekselstalen. Maternele genen *NLRP2, NLRP5, NLRP7, PADI6, TLE6* en KHDC3L werden gescreend met Next Generation Sequencing, samen met *PLCZ1* om een spermafactor defect uit te sluiten. Interessante varianten werden geclassificeerd met VarSome en bevestigd met Sanger sequencing.

Resultaten

In *PLCZ1*, de homozygote variant p.S500L werd gevonden in één mannelijke patiënt en werd geclassificeerd als variant met onduidelijke significantie (klasse 3). Deze mutatie werd eerder gerapporteerd als pathogeen. Verrassend vertoonde deze patiënt staken van embryo-ontwikkeling, in tegenstelling tot eerder gerapporteerde gevallen met falen van bevruchting. De interpretatie blijft hierdoor onduidelijk, maar kan bijdragen tot bewijs van causaliteit van deze mutatie. Omdat een PLCζ defect hier oorzakelijk is, komt deze patiënt in aanmerking voor AOA therapie. In de SCMC genen werden vijf nieuw gevonden mutaties van onduidelijke significantie gevonden (in *PADI6*, *NLRP2*, *NLRP5*, *NLRP7* en *TLE6*) in vijf vrouwelijke patiënten. Functionele analyse is essentieel voor betere interpretatie, omdat geen van deze varianten eerder werd gepubliceerd.

Conclusie

Deze studie levert geen bewijs over de pathogeniteit van deze mutaties, maar kan toch zeker bijdragen aan de kennis over de associatie tussen infertiliteit en *PLCZ1* en de SCMC genen.

1. Introduction

1.1 Infertility: definition of disease

Fertility is the capacity to establish a clinical pregnancy [1] and requires normal oocyte maturation, fertilization and embryonic development [2]. Infertility is defined as a couple's inability to conceive after twelve months of frequent, unprotected sexual intercourse [3] due to an impairment of a person's capacity to reproduce, either as an individual or with his/her partner. According to the latest WHO definition, infertility is a disease which generates disability as an impairment of function [1]. It is a unique condition since it often affects a couple rather than an individual [3]. Subfertility could be used interchangeably with infertility. It is defined as any form or grade of reduced fertility in couples unsuccessfully trying to conceive [1].

The three major factors influencing the spontaneous probability of conception are the time period of negative clinical outcomes, the maternal age and the disease-related infertility [1]. Not only does the number of oocytes decrease throughout productive years, the remaining oocytes are poorer in quality, leading to chromosomal abnormalities and spontaneous abortions. This decrease in quality also occurs with the man's sperm and thus his ability to reproduce [3]. While the most significant negative predictive factor of fertility is increasing women's age at conception [1, 3], other factors including acute or chronic conditions, infectious diseases, genetic conditions [3], lifestyle and environmental factors are believed to play an increasing role. These factors will be presented as gender-specific or not [1].

Infertility is a potentially life-changing diagnosis for couples trying to conceive. A diagnosis of infertility and the associated management plan can lead to significant physical, financial and emotional hardships, such as psychological stress, anxiety and depression. Once a couple is determined to be infertile, prompt referral to a specialist is indicated. Treatment varies according to the cause. Infertility has become a growing public health concern [3].

1.2 Epidemiology of infertility

More than 48.5 million couples (or over 186 million people [2]) worldwide are infertile, which has a widespread global impact [4]. In developing countries, infertility has been attributed to a female factor in 35%-45%, to a male factor in 8%-30% or to combined factors in 35%-40%. A number of infertility causes remains unknown [3], with especially 50% of male infertility being unexplained. This is particularly worrying, as globally, male infertility may be attributed as the major causative factor underlying infertility [5].

Although the prevalence of infertility remained relatively stable in the past several decades, the demand for infertility services has increased substantially. This is partially due to delayed childbearing trends combined with advances in assisted reproductive technology [3]. Assisted reproductive technology (ART) have resulted in millions of children being born. However, pregnancy rates per embryo transfer remain fairly low, suggesting that other causative factors relating to abortive embryogenesis and recurrent implantation failure in the clinic are yet to be resolved [5].

1.3 IVF and ICSI

ART is now routine in the treatment of infertile individuals. With the development of ART, the number of IVF (in vitro fertilization) and ICSI (intracytoplasmic sperm injection) cycles has increased every year, and now more than five million babies have been delivered by ART [4].

IVF involves the retrieval of oocytes after ovarian stimulation, which are then *in vitro* exposed to the sperm to achieve fertilization. One or more embryos are then transferred into the uterine cavity for implantation. IVF is indicated for all causes of infertility after failing treatment with less invasive therapies [3].

ICSI consists of the injection of a single sperm into the oocyte [6]. It was first developed to help attain successful fertilization for couples with severe male factor deficiency, but now it is commonly used as the standard ART, even in up to 80% of all in vitro cycles in Belgium. During this procedure, all of the natural selection barriers to fertilization are bypassed [7]. In an IVF or ICSI cycle, the cumulus-oocyte complex is first isolated from the individual's follicular fluid, and if ICSI is pursued, the cumulus cells surrounding the oocytes are further removed. The oocytes are then cultured in vitro. Fertilization occurs within 14-16 hours after IVF or ICSI treatment. Normal fertilized oocytes continue to be cultured, and embryo quality is assessed three and five days after fertilization on the basis of cell number and morphology [4]. Even though ICSI technique leads to spectacular fertilization rates (± 70% to 80%), low or total failed fertilization still occurs in 1% to 5% of all ICSI cycles [5, 7].

The recurrent failure after ICSI can be categorized as either fertilization failure or early embryonic arrest following successful fertilization [2, 8].

1.4 Fertilization failure after ICSI cycles

The term ‘failed fertilization’ typically refers to failure of all the available mature metaphase II (MII) oocytes to fertilize [9]. This often occurs because of nucleus-cytoplasmic maturation asynchrony in the oocyte [10], but can also fail by chance when very few MII oocytes are available for ICSI, or can be due to the lack of motile spermatozoa on the day of oocyte retrieval. Still, the principle cause of failed fertilization has been attributed to an oocyte activation deficiency [9].

1.4.1 Oocyte activation by sperm-induced calcium oscillations

Currently, ICSI failure is believed to result mainly from a deficiency in a physiological mechanism fundamental to fertilization, termed oocyte activation [5]. The term ‘oocyte activation’ describes a series of biochemical and morphological events that mammalian eggs go through to prepare the egg for early embryo development after sperm-egg fusion [6]. Oocyte activation relies both on oocyte-related and sperm-related oocyte-activating factors that transduce the sperm-derived signal to the oocyte’s cell-cycle controlling systems [9].

At mammalian fertilization, the fundamental stimulus that triggers oocyte activation and initiation of early embryogenic development is an acute rise of the intracellular free calcium concentration inside the egg cytoplasm [6]. The first Ca^{2+} trigger is followed by a series of shorter high-frequency Ca^{2+} transients, the so-called Ca^{2+} oscillations that continue for several hours after fertilization or ICSI [9]. Ca^{2+} transients at oocyte activation are essential [5]. This striking phenomenon consequently orchestrates a series of further events of egg activation, such as cortical granule exocytosis, which blocks polyspermy, meiotic resumption and pronuclear development. Mammalian eggs are very sensitive to the precise pattern of Ca^{2+} oscillations. Considering that the cell cycle progression rate may be an indicator of normal embryogenesis, the profile of Ca^{2+} oscillations during mammalian fertilization may not only be necessary and sufficient for egg activation to occur but also equally important for subsequent embryogenic events [6].

1.4.2 Calcium oscillation induced by PLCζ

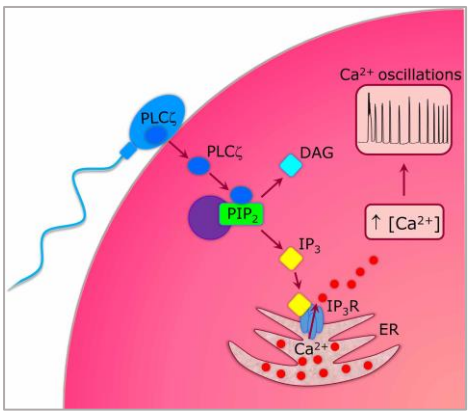


Figure 1. Schematic representation of egg activation triggered by sperm-specific PLCζ. Following sperm–egg membrane fusion, PLCζ is released from the sperm into the egg cytosol and targets a distinct intracellular vesicular membrane containing its membrane-bound substrate, phosphoinositide 4,5-bisphosphate (PIP_2). PLCζ-mediated PIP_2 hydrolysis produces two second messengers DAG (diacylglycerol) and IP_3 (inositol 1,4,5-trisphosphate), which lead to an interaction with its receptor (IP_3R) on the endoplasmic reticulum (ER) that triggers Ca^{2+} release from intracellular stores. This subsequently produces the characteristic pattern of Ca^{2+} oscillations that initiate egg activation [6].

So far, several sperm factors involved in oocyte activation have been identified [7]. However, multiple studies have found the specific PLC isozyme to be the most significant factor and it is believed to be primarily responsible for mammalian oocyte activation. It was first identified using mouse expressed sequence tag databases to describe a novel, and testis-specific phospholipase C, termed PLC zeta [5]. PLCζ is expressed during the transition of round spermatid to elongated spermatid and is localized in the acrosomal, equatorial and post-acrosomal regions of the human sperm head [7]. Sperm PLCζ catalyses the hydrolysis of phosphatidylinositol 4,5-biphosphate (PIP_2) triggering cytosolic Ca^{2+} oscillations *(Figure 1)* [6].

Sperm that consistently failed to fertilize oocytes following ICSI is unable to produce Ca^{2+} oscillations, or else they produce oscillations that are reduced in frequency and amplitude [5]. Previous studies showed that PLCζ expression in men with a history of failed fertilization is low [7], suggesting that defects in PLCζ may underlie cases of fertilization failure [5].

Mutating the catalytic domain of PLCζ completely abrogates Ca^{2+} release in mouse oocytes [5]. The first clinical correlates between PLCζ and male infertility were reported to result in reduced levels, total absence, or abnormal localization patterns of the PLCζ protein [11, 12]. Both heterozygous and homozygous mutations have been identified in the *PLCZ1* gene of patients with a sperm-related oocyte activation deficit [5]. SNPs (single-nucleotide polymorphisms) have been reported either within the coding sequence of *PLCZ1* or its associated bi-directional promoter in human patients [6]. Twenty-one pathogenic *PLCZ1* mutations have yet been identified and reported, which result in alteration of protein fold, truncated protein, frameshift truncation of protein and reduced substrate binding [12]. An abnormally reduced ability of sperm to induce calcium release could be linked to defective embryogenesis in cases of hydatidiform moles [13]. Levels and localization patterns of PLCζ are associated with sperm defects such as abnormal sperm parameters and morphology, DNA fragmentation, and cases of abnormal embryogenesis and absence of varicocele [12].

Abnormalities in sperm PLCζ levels may underlie not only infertility through fertilization failure, but also cases of male subfertility, whereby adequate PLCζ may be delivered to oocytes to cause activation only, but which may be insufficient for embryonic competence [6]. However, such correlations between *PLCZ1* mutations and levels of PLCζ within sperm are currently poorly understood, particularly the contributory effects that heterozygous mutation and SNPs may exert upon the apparent inherent variation between different fertile males in levels of sperm PLCζ [5].

1.5 Embryo arrest causing ICSI failure

Normal embryonic development is the key to establish a successful pregnancy. Of all human IVF-produced embryos, 40% to 70% is viable, whereas others arrest at different stages of development. Developmental arrested embryos are discarded whereas viable embryos are cryopreserved, cultivated to blastocysts, or implanted directly into the individual's uterus. IVF and ICSI cycles fail if all of an individual's embryos are arrested at early stages of development. Much of human embryonic developmental potential is determined before fertilization by the content of the oocyte and is encoded by so-called maternal-effect genes (MEGs) [4]. Although the zygote is formed by the fusion of paternal and maternal pronuclei, early development, consisting the first 3 days in human, is therefore essentially under maternal command from the MEGs in the oocyte ooplasm. MEGs code for unique transcripts for proteins that contribute to the SCMC [14], required for early cleavage events post fertilization [15].

1.5.1 SCMC as the major maternal factor in early embryo development

Numerous recent studies have identified maternal-effect factors that play essential roles in preimplantation embryonic development [16] through the degradation of maternal RNA and proteins and the activation of the embryonic genome. This group of maternal factors called the subcortical maternal complex (SCMC) functions as a spindle positioner through regulation of subcortical F-actin *(Figure 2)* and its presence is required for proper symmetric cell division [15]. The complex consists of at least following proteins: *transducing-like enhancer of split 6 (TLE6), KH domain containing 3 like (KHDC3L), protein-arginine deiminase type 6 (PADI6), oocyte-expressed protein homologue, NLR family pyrin domain containing 2 (NLRP2), NLR family pyrin domain containing 5 (NLRP5)* and *NLR family pyrin domain containing 7 (NLRP7)* [2, 8, 17]. All these genes were found responsible for primary female infertility and early development arrest during embryonic cleavage stages after ART [18].

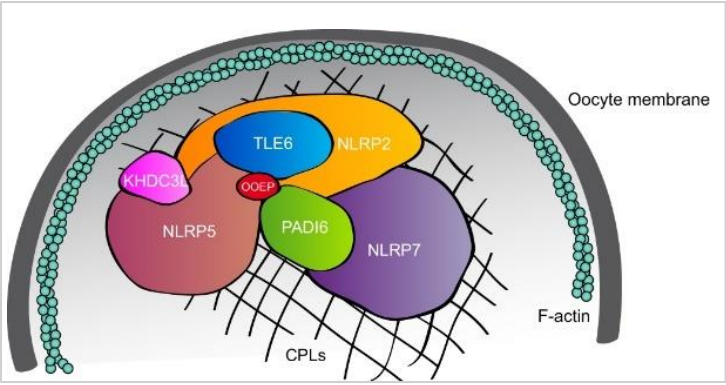


Figure 2. Structure of SCMC, interacting with cytoplasmic lattice (CPL) and F-actin in the oocyte [19].

1.5.1.1 PADI6

PADI6 is a maternal-effect gene and a member of the SCMC that has been shown to have recessive variants that affect gene function. *PADI6* codes for peptidylarginine deiminase, a member of a class of Ca^{2+}-dependent enzymes that are responsible for citrullination, which consists of post-translational modifications of arginine residues within peptides (peptidylarginine) to citrulline [18]. PADI6 plays a role in *de novo* protein synthesis prior to the maternal-to-embryonic transition and in microtubule-mediated organelle positioning and movement. It is critical to the formation of the oocyte-restricted fibrous cytoskeletal structure, namely the cytoplasmic lattice (CPL) [8, 16]. It plays important roles in the formation of rigid structures such as hair, skin and myelin sheath [18]. Compared to other PADI family members, PADI6 is uniquely localized in oocytes and early embryos [4].

PADI6 mutations are identified as the cause of female-specific infertility due to an embryonic arrest [20]. *PADI6*-null embryos do not develop past the two-cell stage [16]. The lack of CPL causes abnormal organelle positioning and redistribution, reduced amounts of ribosomal and messenger RNA, altered *de novo* protein synthesis and impaired embryonic genome activation [18]. This impaired zygotic genome activation in patients' embryos is indicated by the reduced levels of its faithful readout phosphorylated RNA polymerase II [2, 20]. Previous studies identified 14 different mutations in *PADI6*, including 10 compound-heterozygous mutations and 4 homozygous mutations, among which both nonsense and missense compound-heterozygous mutations and frameshift, nonsense and missense homozygous mutations [4, 8, 16]. Homozygous nonsense mutations in *PADI6* are predicted to cause a truncated protein. Alternatively, this mutation could also trigger nonsense-mediated decay. *PADI6* activity in mutated individuals is completely eliminated [4]. However, only a limited number of cases with *PADI6* mutations has been reported, which indicates the need for further research, since the genetic basis remains poorly understood [16].

1.5.1.2 TLE6

Transducin-like enhancer of split-6 (TLE6) also contributes to the SCMC. TLE6 is a putative PKA substrate. This cAMP-dependent protein kinase (PKA) is a critical physiological inhibitor of meiotic resumption in oocytes. PKA regulates the phosphorylation of TLE6, which is essential for meiotic maturation. The special-temporal localization of PKA and TLE6 within the oocyte is also consistent with interaction between these proteins. This phosphorylation could be related to the roles of *KHDC3L* and *NLRP5*, in spindle assembly and migration during oocyte maturation [21], since *TLE6* provides the formation of the cytoplasmic F-actin fine meshwork in the murine zygote subcortex. Absence leads to asymmetric cell division and cleavage-stage embryonic death [15] on Day 3 and failure to form blastocysts [22].

Mutations were identified as responsible for early embryonic arrest, leading to human preimplantation embryo lethality [2, 8, 22]. The mutation abrogates TLE6 phosphorylation and impairs its binding to components of the SCMC, which in turn affects the stability of the SCMC [23]. A homozygous frameshift mutation was detected in *TLE6* that led to the truncated protein, but also a homozygous missense (affecting phosphorylation) and compound-heterozygous mutations have been identified [8].

However, these mutations could only account for a small number of cases, and other new mutant genes or novel mutations in this and other SCMC genes responsible for the phenotypes still need to be identified [8].

1.5.1.3 NLRP7

Nucleotide-binding oligomerization domain (NOD)-like receptors with a pyrin domain (PYD), NLRPs, are pattern recognition receptors, well recognized for their important roles in innate immunity, inflammatory responses and apoptosis [17, 24]. Only recently were specific NLRPs identified as maternal-effect genes. They integrate multiple functions to modulate zygotic genome activation, cell division, and organelle rearrangement. Several reports suggest that NLRPs are associated with imprinting and other diseases, e.g. cancer [17].

The human *NLRP7* gene evolved from duplication of *NLRP2* during evolution, which explains their similar function [17, 24, 25]. Colocation with *PADI6* and *KHCD3L* confirms that *NLRP7* is also a SCMC gene [17, 18]. Within oocytes, NLRP7 localizes to the cytoskeleton and Golgi apparatus and is abundant in the cortical region [26, 27]. The functions of *NLRP7* consist of chromatin reprogramming, DNA methylation during early embryonic development, but also

embryonic cellular immune response, crucial for embryonic development [17]. The expression of NLRP7 is at its lowest level by Day 3 and increases again by Day 5. This pattern corresponds with the embryonic genome activation and blastocyst stages, respectively [17, 24].

NLRP7 biallelic sequence variants, genomic deletions and complex rearrangements such as splice site and frameshift variants have been observed in affected females [22, 28, 29]. Homozygous or compound heterozygous *NLRP7* missense and nonsense mutations have been reported in both genders, but male mutations do not affect reproductive outcomes, which indicates that *NLRP7* regulates female reproduction specifically [18, 24, 27].

Mutations in *NLRP7* play a major role in recurrent hydatidiform moles (RHM) and reproductive waste (48-70% of patients with RHM [26, 27]), and 59 autosomal recessive identified mutations affect reproduction by a two-hit mechanism [2, 8, 22, 24, 28-30]. A hydatidiform mole is an abnormal human pregnancy, characterized by non-existent, or abnormal, embryonic development [28]. Although research is limited for this gene, there has recently been a study showing a link between embryonic development in a patient with *NLRP7* mutations and embryo arrest, where a pathogenic variant was found [27]. NLRP7 is important before and after implantation, since mutated women with *NLRP7* mutations have normal ovulation and several pregnancies, but show impairment after 7 to 14 weeks of gestation [17]. However, multiple studies reported some controversial outcomes regarding the association of *NLRP7* with RHM, but this might be due to a too small study group [24, 25, 30-32]. The lack of methylation marks in moles, supports the hypothesis that *NLRP7* mutations have a direct negative impact on the acquisition of imprinting marks [32]. Rather it would be explained by a polygenetic and multifactorial effect [31].

1.5.1.4 KHDC3L

KH domain containing protein 3 like (KHDC3L, C6orf221 or FILIA), as a part of the SCMC, is expressed in human epiblasts and ensures their genome stability and viability. In response to DNA damage, KHDC3L also localizes to DNA damage sites and facilitates homologous recombination mediated DNA repair [33]. Since KHDC3L and NLRP7 might perform similar functions in embryonic development [8], KHDC3L might take part in methylation [32] related to the establishment or maintaining maternal epigenetic marks during oogenesis and development [15, 30]. The colocation of these two genes suggests that KHDC3L also plays a role in the intracellular trafficking and secretion of cytokines [26].

Mutations in *KHDC3L* and recurrent complete hydatidiform moles might be associated [8, 15], which, as mentioned above, results in maternally-inherited, recurrent pregnancy losses [14, 33]. In patients with RHMs but negative for *NLRP7* mutations, 14% was found to have a *KHDC3L* mutation [26]. In contrast to *NLRP7*, *KHDC3L* is a minor gene involved in biparental RHM (BiHM) and seven autosomal recessive mutations in *KHDC3L* (initiation codon mutations, splicing mutations and frameshift mutations) have been reported in nine families, both missense homozygous and compound heterozygous, which suggest that, similar to other genes, heterozygosity is not sufficient to cause BiHM [8, 22, 29]. Interestingly, a homozygous frameshift mutation (c.44delA) was identified and found to be related to embryo arrest [8]. Other than *NLRP7*, *KHDC3L* mutations can be inherited both maternally and paternally (or *de novo*) and the frequency and impact of this gene may be underestimated if the genetic screening is performed only in maternal genome [33].

However, mutations in these genes can only account for a small number of patients, and the genetic basis of early embryonic arrest remains largely unknown [2]. This rare occurrence might be the underlying issue in research with small sample size [33], where no mutations linked to RHMs were found. These articles claim that *KHDC3L* is not a (major) factor in recurrent miscarriages [25, 28, 30, 31], but these results may also confirm that *KHDC3L* is only a minor factor in developing BiHM [32].

1.5.1.5 NLRP2

NLRP2 controls age-associated maternal fertility, and both loss and overexpression of NLRP2 results in a significant reduction in viable embryos [2, 34]. Fas-associated protein factor (FAF1) is a specific binder partner of the NLRP2 protein. Both proteins are present throughout oogenesis and the development of preimplantation embryos, and are mainly localized to both the cytoplasm and nucleus, implying that they physically interact. In mice, this interaction between NLRP2 and FAF1 is probably essential for successful development of cleavage-stage embryos and failure of the NLRP2-FAF1 complex results in the inability to effectively degrade maternal proteins. This causes important development events to occur in an unregulated manner [35]. The expression pattern is similar to NLRP7, since it reaches the lowest point on Day 3 (embryonic genome activation) and increases to Day 5 (blastocysts stage) [17, 24].

Different truncating missense, frameshift and nonsense homozygous and compound heterozygous mutations in *NLRP2* were identified in five patients, and all of these mutations were responsible for early embryonic arrest and lower blastocyst rates. Mutations in *NLRP2* impair the protein stability [2]. All mutations showed indications for a Mendelian recessive

inheritance pattern [2, 17]. *NLRP2* has been associated with the imprinting disorder Beckwith-Wiedemann syndrome (BWS) [25, 32, 34, 36]. BWS has similarities to (R)HM: the hydropic placenta and the hyperfunction of paternally relative to maternally expressed genes in imprinted regions. Most often this is caused by abnormal imprinting, point mutations or duplications/deletions in *NLRP2* regions [28].

Similar to *NLRP7* and *KHDC3L*, several studies showed no link with a pathogenic phenotype, but this can also be explained by the combination of rare appearance of these mutations and the small study group [25]. Still, these mutations were only reported in a limited number of cases, thus more evidence is required to understand the whole genetic influence on infertility.

1.5.1.6 NLRP5

NLRP5 (or MATER) is a member of the NLRP family of genes [35] and was initially identified as an oocyte-specific auto-antigen associated with auto-immune premature ovarian failure [17], although NLRP5 altered expression in abnormal human embryos and developmentally arrested two-cell embryos strongly suggests a role in preimplantation embryo development [15]. It is located in oocyte mitochondria and nucleoli, close to the nuclear pores, but is predominantly cytoplasmic throughout preimplantation development. This suggests that NLRP5 might have both cytoplasmic and nuclear functions. It interacts with anti-apoptotic pathways, which suggests its function in human follicular maturation. NLRP5 may be involved in translocation/tethering of a subset of organelles to subcortical or perinuclear regions, which is supported by the location of NLRP5 to the CPL network with PADI6 [17].

Multiple homozygous and compound heterozygous mutations in *NLRP5* were identified, both nonsense and missense, and were responsible for early embryonic arrest. All mutations showed indications for a Mendelian recessive inheritance pattern [2]. Mutations in *NLRP5* were associated with reproductive wastage and multilocus imprinting disorders (MLID) in humans, including miscarriages and reported molar pregnancy. Patients exposed to maternal *NLRP5* variants showed regions of methylation disturbance at known imprinted loci, with variable distribution and severity [15]. *NLRP5* mutations are not commonly reported, implicating that further research on the SCMC genes is yet to be done.

1.5.2 Distinguishing oocyte-related and sperm-related deficiencies in ICSI failure

1.5.2.1 Evaluation of oocyte activation by MOAT

The mouse oocyte activation test (MOAT) can be carried out by injecting human spermatozoa into mouse oocytes to evaluate their activating capacity. Such a diagnostic test prior to assisted oocyte activation (AOA) is of a major importance to assess the underlying cause of ICSI failure [9], since it can differ sperm- and oocyte-related factors [37]. Distinguishing the activation deficiency due to a depressed activity of a sperm factor from the impaired ability of the oocyte to respond to this factor in order to be activated is crucial [9]. After micromanipulations, oocytes were cultured and checked for two-cell formation (activation rate) after 24 hours. Depending on the activation rate, patients are classified into three different groups. MOAT group 1 includes patients with a sperm-related OAD (oocyte activation deficiency) causing low activation (activation rate 0-20%); MOAT group 2 includes all patients with a diminished sperm-related oocyte-activating capacity (activation rate 21-84%); MOAT group 3 covers patients with a suspected oocyte-related deficiency and thus a high activation rate (activation rate ≥85%) [38, 39]. AOA has been proven very beneficial for treating fertilization failure caused by sperm factors, but most patients with a suspected oocyte factor do not benefit from AOA. Fertilization rates obtained after AOA are higher in MOAT group 1, followed by MOAT group 2 and finally MOAT group 3 [38]. It is still not clear why some MOAT 3 patients benefit from AOA while others don't. The timing of the oocyte factor affected in the activation pathway might play a role. If there is a mutation in an oocyte factor that acts after the Ca^{2+} oscillations are induced, AOA treatment will not overcome the oocyte activation failure.

Importantly, MOAT serves as a valuable tool for appropriate patient counselling, particularly regarding gamete donation. Moreover, the transmission risk of a PLCζ-associated OAD to children born after AOA can be estimated with more accuracy [38].

1.5.2.1 Calcium pattern analysis by MOCA

Despite the added value of MOAT, more sensitive diagnostic tests, such as Ca^{2+} pattern analysis of the patient's spermatozoa in mice (mouse oocyte calcium analysis; MOCA) or human oocytes (human oocyte calcium analysis; HOCA) would aid in predicting the success of AOA and guiding further clinical management of fertilization failure after ICSI [38]. The calcium pattern analysis uses a Ca^{2+}-sensitive fluorescent dye and is scored per oocyte and per patient individually based on the presence of calcium spikes and their frequency and amplitude. MOCA should definitely be considered for patients with a MOAT group 2 result, since it can confirm or refute a sperm-related activation deficiency [40].

1.5.3 Current treatment options for ICSI failure

1.5.3.1 AOA treatment to overcome fertilization failure

Currently, cases of ICSI failure are routinely clinically resolved through assisted oocyte activation (AOA), involving artificial induction of Ca^{2+} release following ICSI [5]. Artificial activating agents such as ethanol, Ca^{2+} ionophores and electrical pulses cause a single and prolonged rise in Ca^{2+} in animal oocytes [9]. The AOA protocol differs amongst different study centers. Mechanical AOA, calcium ionophores, a combination of both and strontium chloride are all possible, although calcium ionophores are the most commonly used artificial agents. The chosen protocol diverges in ionophore concentration used, the duration of exposure, the moment of ionophore exposure following ICSI and the number of ionophore exposures. Specifically, in the Department of Reproductive Medicine from UZ Gent, the following protocol is used: the combination of mechanical AOA and calcium ionophores is performed at the moment of ICSI with an injection of 0.1 mol/l $CaCl_2$ together with spermatozoa followed by a 2-fold ionomycin (10 µmol/l) exposure afterwards, 10 minutes each and 30 minutes apart [39].

Numerous studies have proven that AOA following ICSI could lead to improved fertilization and pregnancy rates [7, 38, 41, 42]. However, a major concern is the non-physiological Ca^{2+} transients induced within the oocyte, since it does not mimic precisely the physiological fertilization process [43]. This may affect the activity of specific proteins and enzymes downstream and also the cellular homeostasis, which in turn may exert long-term genetic/epigenetic, biochemical and physiological effects [5].

However, the follow-up results of children born following ICSI with AOA are reassuring although conflicting data still exists. That is why couples undergoing AOA should be informed [43]. Neonatal, developmental as well as behavioral outcome were not altered in children born after ICSI-AOA. Nevertheless, the sample sizes of the studies were low and definite conclusions cannot be drawn [43, 44].

It is imperative that a more endogenous clinical treatment is pursued to overcome such controversial aspects of AOA, and to this degree, PLCζ may represent a viable method to clinically treat cases for which a *PLCZ1* defect represents the underlying cause of fertilization failure [5]. PLCζ cRNA (complementary RNA) injections are a possibility [45], but have a risk of incorporation into the genome, which limits their use [6, 10, 39]. Therefore, the production of purified, enzymatically-active, recombinant human PLCζ protein has been a goal for many laboratories around the world, to varying degrees of success, to enable dose-controlled delivery of functional viable PLCζ protein [6]. Cases with defective PLCζ may be resolved by the utilization of this purified protein as a more physiological replacement for current AOA protocols, which is fast coming closer to realization [5].

1.5.3.2 Correction of genetic mutations using gene editing

Genome editing tools, such as the clustered regularly interspaced short palindromic repeat (CRISPR)/Cas9, have enabled the correction of a genetic mutation in the human germline [46]. To date, the main purpose of germline genome editing would be to correct heritable causes of monogenic diseases. In the light of infertility, genome editing could not only be a part of the infertility treatment, but would also prevent transmission of infertility to offspring. *PLCZ1* represents a valuable candidate and it is conceivable that more genes involved in embryogenesis or recurrent implantation failure could be corrected with CRISPR/Cas9 [47].

The actual application of CRISPR/Cas9 in a clinical setting is currently limited by both technical flaws and ethical considerations (i.a. “designer baby” concerns). Technically, both mosaicism and off target editing has been reported, but the actual use of gene editing in premature oocytes is currently unjustifiable, given that germline genome editing can potentially affect progeny with the substantial risk of off-target effects [46-48]. The absence of a reasonable alternative is a common argument to stand behind gene editing [47].

1.6 Overcoming embryo arrest with AOA and assessment of its molecular determinants

It has been demonstrated that calcium oscillations not only provide for short-term, but also for long-term developmental events [37]. Oocyte activation is a regulated process triggered by the sperm entry and the oocyte must be released from the metaphase II-arrest to complete initiation and progression of the cell cycle and the subsequent cell divisions [39, 49]. The master key to initiate all cytological changes in fertilized oocytes, and thus its developmental competence [49], is a series of intracellular calcium rises that start short after sperm-oocyte fusion [39]. It seems like a reasonable hypothesis that AOA, while creating artificial calcium oscillations, might positively affect the pregnancy rates of cases of failed fertilization caused by an embryo arrest due to inadequate oocyte activation.

More research has yet to investigate the results of this procedure, but two studies have been done to date in this regard: the Ebner study [50], published in 2015, and the Bonte study, which took place in our own study center.

1.6.1 The Ebner Study and its research on AOA treatment in early developmental arrest

Ebner et al. (2015) stated that a calcium ionophore treatment (A23187, calcimycin) improves embryo development and outcome in patients with a history of developmental problems/arrest [50]. Application of A23187 led to increased rates of cleavage to 2-cell stage, blastocyst formation and clinical pregnancy/live birth.

This prospective, multicenter (five Austrian centers), uncontrolled intervention study (duration 1 year) included 57 patient. Inclusion criteria were complete embryo arrest in a previous cycle, complete developmental delay, or reduced blastocyst formation on Day 5 (≤15%). Severe male factor patients and patients with <30% fertilization rate after ICSI were excluded because they would be routine indications for ionophore usage. The total of the 57 immediately preceding cycles in the same patients constituted the control cycles/control group. In the treatment cycles, all metaphase II-oocytes were exposed to a commercially available ready-to-use ionophore for 15 minutes immediately after ICSI. After a three-step washing procedure, *in vitro* culture was performed as in the control cycles, up to blastocyst stage when achievable.

Fertilization rate did not differ significantly (75.4% vs 73.2%), but further cleavage to 2-cell stage was significantly higher in the ionophore group (98.5% vs 91.9%), as was the number of formed blastocyst on Day 5 (47.6% vs 5.5%). This was associated with a significant increase in implantation rate (44.4% vs 12.5%), clinical pregnancy (45.1% vs 12.8%) and live birth (45.1% vs 12.8%). All babies were healthy. This is the first evidence that developmental incompetence of embryos is an additional indication for ionophore treatment [50].

While *Ebner et al. (2015)* claims to have significant results, there are some unmentioned limitations concerning the inclusion of patients [50]. The ionophore treatment had significant results with patients with a history of developmental problems/arrest, but it was never excluded that the couples did not suffer from Ca^{2+} oscillation failure. Patients were included in the study after one failed ICSI cycle, without MOAT/MOCA testing to distinguish possible sperm and oocyte factors causing infertility. Since there can be a substantial cycle to cycle variability, the patient inclusion criteria were not strict enough. Also, in this study there were changes in the ovarian stimulation protocol of the patients included, which could have had an effect on the results. This raises doubts whether the study results are valid.

1.6.2 Reassessing the Ebner results in Bonte 1.0

The Bonte Study repeated the Ebner study while avoiding the patient recruitment bias and established three main objectives. First, the study investigated whether AOA could improve embryo development in patients with good fertilization (≥60%) with a minimum of 10 zygotes but low blastocyst rates on Day 5 (≤15% of at least early blastocyst quality). This result may affect the clinical use of the current AOA treatment. Secondly, it was studied with MOCA whether embryo delay/arrest is linked with aberrant calcium patterns induced by the spermatozoa. Finally, the Bonte Study examined if the embryo delay/arrest condition is caused by mutations in genes important for successful embryo development, which have only been identified in a very limited number of patients [38], more specifically the SCMC genes and *PLCZ1*.

In this retrospective cohort study, AOA cycles and previous ICSI cycles in couples experiencing low or total failed fertilization after ICSI were compared. Different from the Ebner Study [50], all 18 patients were examined before AOA using MOCA to exclude a sperm-related oocyte-activating deficiency [38]. The overall fertilization rates after AOA were even lower than after conventional ICSI (69.7% vs 76.7%), however blastocyst rates increased slightly after AOA compared to previous attempts (17.4% vs 10.1%) as did the pregnancy rates (45.5% vs 38.7%). Many AOA cycles ended up without blastocysts and were not included to calculate pregnancy rate. In these patients, AOA could not restore embryo development. Further details can be found in *Table 1*. All sperm samples displayed calcium oscillatory patterns on MOCA comparable to the control sample, thus a sperm-factor defect can be ruled out.

For the male patient, only PLCζ is tested, since this gene is the major genetic factor in male infertility according to research [5-7]. However, no mutations were expected since failed fertilization did not occur. To investigate a female factor, multiple genes, previously associated with female infertility, were screened: *PADI6, TLE6, NLRP7, KHDC3L, NLRP2* and *NLRP5*. As mentioned before, all these genes contribute to the formation of the SCMC [2].

Outcomes of ICSI with AOA were compared to a standard ICSI treatment in couples with a history of embryo development arrest. The results showed only a slight difference in blastocyst and pregnancy rates using AOA, and definitely less significant outcomes than the Ebner Study. Among the 18 tested female patients, seven heterozygous variants of uncertain significance were reported in six patients. More details about the interesting variants reported in Bonte 1.0 can be found in *Additional file 1*. Further research for AOA as a treatment for early development arrest is thus not likely to be repeated in our center, but proceeding the genetic research could be beneficial.

	ICSI cycles, n	Fertilization rate (2PN/MII)	Blastocyst rate (blastocyst/2PN)	Pregnancy rate (+hCG/n ET)	*AOA cycles, n*	Fertilization rate (2PN/MII)	Blastocyst rate (blastocyst/2PN)	Pregnancy rate (+hCG/n ET)	MOCA AxF score
P1	*3*	90.32% (28/31)	0.00% (0/28)	na	*1*	94.44% (17/18)	0.00% (0/17)	na	93.3
P2	*4*	73.68% (28/38)	14.29% (4/28)	0.00% (0/2)	*1*	33.33% (5/15)	40.00% (2/5)	100.00% (1/1)	101.0
P3	*3*	66.04% (35/53)	8.57% (3/35)	100.00% (1/1)	*1*	81.48% (22/27)	9.09% (2/22)	0.00% (0/2)	na
P4	*2*	83.33% (10/12)	10.00% (1/10)	0.00% (0/1)	*1*	57.14% (4/7)	0.00% (0/4)	na	76.7
P5	*2*	76.92% (10/13)	10.00% (1/10)	50.00% (1/2)	*1*	88.89% (8/9)	0.00% (0/8)	na	52.0
P6	*2*	73.68% (14/19)	7.14% (1/14)	0.00% (0/1)	*1*	71.43% (5/7)	0.00% (0/5)	na	90.8
P7	*2*	69.23% (9/13)	11.11% (1/9)	100.00% (1/1)	*1*	83.33% (10/12)	10.00% (1/10)	0.00% (0/1)	50.4
P8	*2*	66.67% (12/18)	8.33% (1/12)	0.00% (0/1)	*1*	54.55% (6/11)	0.00% (0/6)	na	56.2
P9	*2*	75.00% (9/12)	11.11% (1/9)	0.00% (0/1)	*1*	58.33% (7/12)	14.29% (1/7)	100.00% (1/1)	123.2
P10	*5*	74.65% (53/71)	13.21% (7/53)	60.00% (3/5)	*1*	61.54% (8/13)	25.00% (2/8)	0.00% (0/1)	107.6
P11	*4*	78.38% (29/37)	6.90% (2/29)	25.00% (1/4)	*1*	66.67% (4/6)	25.00% (1/4)	0.00% (0/1)	77.0
P12	*2*	75.00% (15/20)	13.33% (2/15)	0.00% (0/2)	*1*	62.50% (5/8)	20.00% (1/5)	100.00% (1/1)	na
P13	*2*	93.75% (15/16)	6.67% (1/15)	100.00% (1/1)	*1*	71.43% (10/14)	40.00% (4/10)	50.00% (1/2)	na
P14	*2*	66.67% (10/15)	10.00% (1/10)	100.00% (1/1)	*1*	75.00% (6/8)	16.67% (1/6)	100.00% (1/1)	na
P15	*4*	71.88% (23/32)	13.04% (3/23)	33.33% (1/3)	*1*	25.00% (1/4)	0.00% (0/1)	na	na
P16	*2*	77.78% (14/18)	7.14% (1/14)	0.00% (0/1)	*1*	60.00% (6/10)	0.00% (0/6)	na	na
P17	*3*	90.00% (18/20)	11.11% (2/18)	100.00% (2/2)	*1*	77.78% (7/9)	42.86% (3/7)	na	na
P18	*2*	92.00% (23/25)	17.39% (4/23)	0.00% (0/2)	*1*	87.50% (7/8)	85.71% (6/7)	na	na
Total	***48***	**76.67% (355/463)**	**10.14% (36/355)**	**38.71% (12/31)**	***18***	**69.70% (138/198)**	**17.39% (24/138)**	**45.45% (5/11)**	

Table 1. Comparison of preimplantation characteristics and embryo transfer outcomes after previous ICSI and subsequent AOA treatment and MOCA outcome in patients with embryo developmental problems in the Bonte 1.0 study.

2. Objectives

This master thesis will discuss the Bonte 2.0 study. Patients that suffered from embryo developmental arrest will retrospectively be included for further genetic screening of variants in earlier mentioned genes. The aim of this study is to broad the knowledge of mutations in the maternal effect genes (*NLRP2, NRLP5, NLRP7, PADI6, TLE6,* KHDC3L) and in *PLCZ1*, but also their clinical relevance, in particular their association with and impact on infertility.

3. Methods

3.1 Patient recruitment

The Bonte Study 2.0 retrospectively examined patients with a history of embryo developmental problems/arrest from January 2018 to September 2019 at Ghent University Hospital. Patients were only included when they underwent at least two fresh ICSI cycles at the fertility center at Ghent University, following earlier mentioned Bonte inclusion criteria: patients with good fertilization (≥60%) with at least 10 zygotes, but low blastocyst rates (≤15%).

As shown in *Additional file 2*, 45 couples met the inclusion criteria, among which 6 need to be called again, 6 did not consent and 5 did not deliver a sample after consent. The saliva kit still needs to be sent to 2 couples. In the end, 26 couples returned the saliva kit, of which 24 were screened. Due to problems with the informed consent, one couple is not reported in this Master Thesis. The total study group thus consists of 23 females and 22 males, since there was one lesbian couple included in the study, and 2 couples still need to be screened. The saliva samples were sent to the G-FaST (Ghent-Fertility and Stem cell Team) laboratory at Ghent University Hospital. Screening started in 2018 but some of the patients had treatments in UZ years before 2018.

Ethical approval was obtained. The Ghent University Hospital EC approval number for Bonte 2.0 is EC2017-0819.

3.2 DNA extraction and quantification

Saliva samples were collected in Oragene OG-500 saliva kits. DNA purification of the samples was executed according to the DNA Genotek PrepIT•L2P laboratory protocol. After incubation in 50°C in a water incubator for 1 hour, 500 µl of the sample was mixed with 20 µl of PT-L2P. This mixture was cooled down and then centrifuged, facilitating the transfer of the supernatant as a manner of sample cleaning. Then 600 µl of 95-100% ethanol was added, leading to clotting of the DNA fibers. Afterwards, centrifugation of the mixture led to easy removal of the supernatant. An ethanol wash was later performed with 70% ethanol and was removed after 1 minute, which resulted in only the clotted DNA being left in the tube. After complete air drying, 50 µl of TE solution was added and mixed with the clotted DNA. At last, the sample was incubated at 50°C for 1 hour with occasional vortexing and then incubated at room temperature overnight.

The extracted DNA was quantified using the Qubit™ 1X dsDNA HS Assay Kit (ThermoFisher). The coding exons and flanking intronic sequences of the following genes were examined: *PLCZ1* (NM_033123.4), *NLRP2* (NM_001174081.1), *NLRP5* (NM_153447.4), *NLRP7* (NM_001127255.1), *TLE6* (NM_001143986.1), *PADI6* (NM_207421.4) and *KHDC3L* (NM_001017361.2). These genes were amplified by PCR and subjected to targeted next-generation sequencing (MiSeq).

3.3 PCR protocol

Concerning primer design, PrimerXL was recently developed and generates high quality primers with the possibility to include 5' and 3' intronic regions in the target sequenced, forming the exonic region, the extra-included 3' and 5' intronic regions and intronic near-target as different possible targets [51]. In this study, only the exon and the exon-intron boundaries were amplified. The input sequence consists of the unique coding sequences of all corresponding transcripts. Features such as SNPs and secondary structures are masked in the DNA sequence since it is known to have a negative effect on amplification [51].

Each PCR reaction consisted of 5 µl of the Kapa2G Robust Master Mix, 2.5 µl gDNA of the patient's sample and 2.5 µl of the desired primer. This process was repeated for every primer, each in a different well of the PCR plate. The plate was sealed and placed in the PCR machine from Applied Biosystems, Model Gene-Amp PCR System 2700. The Ford-program was followed (Kapa Robust 60). More details can be found in *Table 2*.

	Time	Temperature
1. Activation	3 min	95°C
2. Denaturation	15 s	95°C
3. Annealing	10 s	60°C
4. Elongation	15 s	72°C
Repeat step 1-4 for 35 cycles		
5. Final	1 min	72°C

Table 2. Ford PCR-program: 35 cycles of activation, denaturation, annealing and elongation.

The quality of the PCR products are controlled by Fragment Analyzer, using microchip-electrophoresis. When one clear spike of the expected length is visible on this electrophoresis, the results are good. A blank sample will not show any visible spikes (aside from those resulting from the primer dimers).

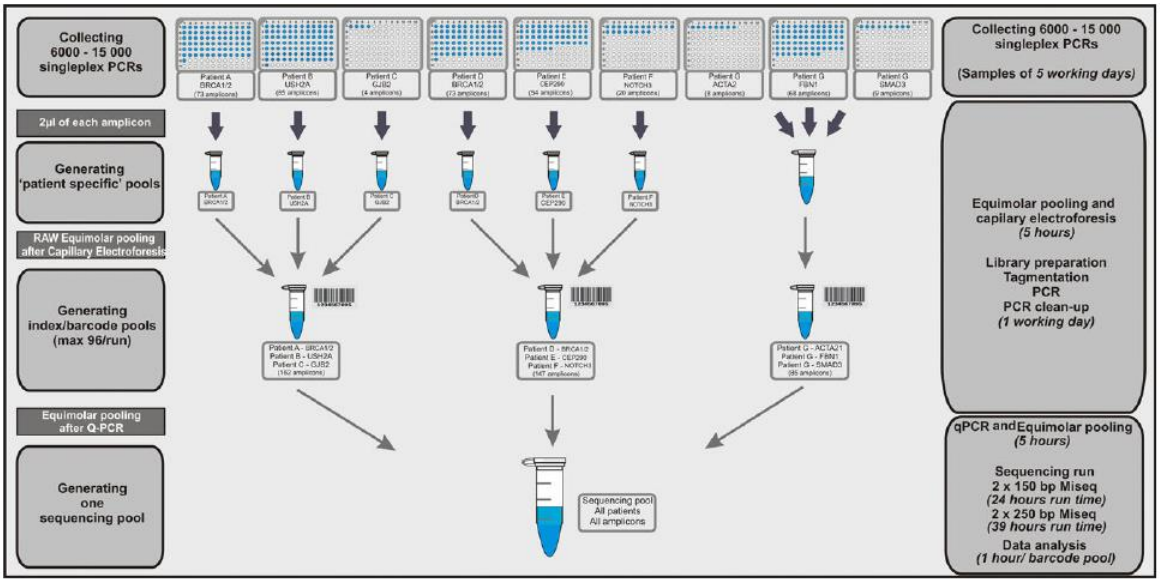


Figure 3: Post-PCR protocol scheme [52].

After amplification, we perform singleplex PCRs for all required amplicons for every patient. In a second step, 2µl of each PCR product is pooled per patient per amplicon. Next, we gather all the PCR-mixes of one patient, for every patient. In a final step, all index pools undergo library preparation and are pooled in a single tube prior to sequencing. These PCR mixtures are sequenced on a MiSeq instrument and variants are listed electronically [52].

3.4 DNA sequencing using MiSeq

The DNA sequencing was first performed using MiSeq, a Next Generation Sequencing (NGS) platform. MiSeq provides end-to-end sequencing solutions and includes onboard cluster generation and data analysis [53], using sequencing by synthesis. Before sequencing, the library splices into single strands with the help of linearization enzyme and is grafted to the flowcell, followed by bridge amplification to form clusters which contain clonal DNA fragments. Four kinds of nucleotides (ddATP, ddGTP, ddCTP, ddTTP) which contain different cleavable fluorescent dye and a removable blocking group would complement the template one base at a time, meanwhile the signal is captured [54]. MiSeq is one of the smallest benchtop sequencers that can perform onboard cluster generation, amplification, genomic DNA sequencing, and data analysis in a single run. It performs both single- and paired-end runs with adjustable read lengths. MiSeq provides an ideal platform for rapid turnaround time and is a cost-effective tool for various analysis focused on a targeted gene sequencing, metagenomics and gene expression studies [53].

The developments in novel sequencing technologies and its overall cost are beneficial for larger amounts of gene sequencing. The power and utility of NGS is based on its massively parallel interrogation of nucleic acids. The ability to simultaneously evaluate millions of base pairs allows clinicians and researchers to ask and answer novel and important questions. However, requiring relatively low-throughput dideoxy sequencing limits its utility [55].

3.5 Variant classification using VarSome criteria

The received data of mutations in all patients was listed in excel files. Each variant found was examined via multiple online databases: VarSome, Polyphen, SIFT, Grantham score, GnomAD, ClinVar and Human Splice Finder. Those results were gathered in another excel file and the conclusion was based on all available information in the used databases. Gene variants were classified according to the nomenclature recommended by the Human Genome Variation Society (HGVS). Five groups were identified: ‘pathogenic’, ‘likely pathogenic’, ‘uncertain significance’, ‘likely benign’ and ‘benign’ [56]. Variants classified as ‘pathogenic’, ‘likely pathogenic’ and ‘uncertain’ were confirmed by Sanger sequencing.

VarSome is a search engine, aggregator and impact analysis tool for human genetic variation and a community-driven project aiming at sharing global expertise on human variants. It gathers information from more than 30 external databases and its own database consists of more than 33 billion data points describing excess of 500 million variants [57]. By collecting over-all knowledge, it helps oversee the identified and reported mutations and their classification. VarSome is in constant change and new information is added every day. It is possible that one variant is classified as class 3 one day and class 2 another day.

Variant pathogenicity is reported using an automatic variant classifier that evaluates the submitted variant according to the ACMG guidelines, classifying it as ‘pathogenic’ (class 5), ‘likely pathogenic’ (class 4), ‘uncertain significance’ (class 3), ‘likely benign’ (class 2) and ‘benign’ (class 1) [56]. Population frequency, pathogenicity predictions, clinically relevant information and associated phenotypes are factors that contribute to the classification. All data and information is retrieved from other databases and is gathered, allowing a combined expertise of the community to be organized and shared [57]. More details on the criteria for each class and how classification happens, is found in *Additional file 3* and *Additional file 4*.

3.6 Mutation confirmation with Sanger sequencing

Sanger sequencing is a manual single-strand DNA sequencing method. Its most important advantages are its high sensitivity, easy workflow and cost-effective benefits. Further, the sensitivity of NGS may drop compared with Sanger sequencing as not all exons may be sufficiently covered [52], therefore Sanger is still the golden standard. Although NGS is becoming even more beneficial, some variants detected can have serious medical implications for the tested proband and their family. Performing a second orthogonal validation with Sanger is therefore appropriate [55].

In sequencing, a primer attaches to a denatured PCR fragment in a reaction mixture with the four dNTPs (dATP, dCTP, dGTP, dTTP). Nucleotide by nucleotide, DNA is synthesized until a fluorochrome-marked ddNTP attaches. The result is a mixture of fragments of different lengths, all ending in a fluorochrome-marked ddNTP. The DNA sequence can be read from comparing which ddNTP corresponds with which location on electrophoresis.

In our own research lab at UZ Ghent, the BigDye ® Terminator Cycle Sequencing kit is used. After PCR-product cleaning, a master mix is prepared, containing 0.5µl ready reaction, 2µl ABI-buffer, 2µl primer and 4.5µl water per sequence reaction. Every tube contains 9µl master mix and 2 µl purified PCR product. The cycle sequencing reaction follows exact settings: 5 minutes at 95°C, then 25 consecutive cycles (each cycle exists of 10 seconds at 95°C, 5 seconds at 55°C and 2 minutes at 60°C), followed by 10 minutes at 15°C.

Magnetic DTR beads and 85% ethanol are added to the sequence reaction for purification. The DNA fragments attach to the beads, which move to the bottom of the tube after being placed on a magnet, enabling easy removal of the supernatant. This is repeated twice. Before analysis, 40µl water is added to the tubes, causing the elution of the DNA from the beads. 30µl of the mixture is used in the microtiter ABI-plate for sequencing.

4. Results

4.1 Clinical characteristics of semen and previous ICSI cycles

In total, 23 couples were included in the Bonte 2.0 study and were screened for *PADI6*, *NLRP2*, *NLRP5*, *NLRP7*, *KHDC3L* and *TLE6* (female patients) or *PLCZ1* (male patients). Clinical details of these patients are found in *Table 3* and *Table 4*, female and male patient characteristics respectively. For confidential reasons, the patients are coded from 101 to 124, both partners of one couple will be named after this number. Due to an absence of informed consent, P115 is excluded from this thesis.

Patient	Age (at last ICSI cycle)	ICSI cycles, n	Fertilization rate (2PN/MII)	Blastocyst rate (blastocyst/2PN)	Pregnancy rate (+hCG/n ET)	Live birth (yes/no)
P101	35	3	73.08% (19/26)	5.26% (1/19)	100.00% (1/1)	No
P102	28	2	100.00% (10/10)	0.00% (0/10)	0.00% (0/0)	No
P103	32	3	90.00% (27/30)	14.81% (4/27)	50.00%(2/4)	Yes
P104	43	15	70.42% (50/71)	14.00% (7/50)	40.00% (2/5)	Yes
P105	34	5	71.43% (20/28)	15.00% (3/20)	0.00% (0/3)	No
P106	36	5	75.00% (30/40)	13.33% (4/30)	25.00% (1/4)	Yes
P107	27	2	57.89% (11/19)	0.00% (0/19)	0.00% (0/0)	No
P108	34	2	89.47% (17/19)	11.76% (2/17)	0.00% (0/2)	No
P109	33	3	75.00% (36/48)	8.33% (3/36)	50.00% (1/2)	Yes
P110	41	4	62.50% (20/32)	5.00% (1/20)	0.00% (0/1)	No
P111	31	3	58.82% (30/51)	6.67% (2/30)	50.00% (1/2)	Yes
P112	30	3	83.33% (25/30)	4.00% (1/25)	0.00% (0/0)	No
P113	36	2	84.62% (11/13)	0.00% (0/11)	0.00% (0/0)	No
P114	36	2	64.86% (24/37)	4.17% (1/24)	100.00% (1/1)	Yes
P115	No informed consent, this patient will not be reported in this thesis.					
P116	34	3	82.76% (24/29)	0.00% (0/24)	0.00% (0/0)	No
P117	26	2	86.67% (13/15)	0.00%(0/13)	0.00% (0/0)	No
P118	25	2	74.07% (20/27)	5.00% (1/20)	0.00% (0/1)	No
P119	36	4	70.37% (19/27)	10.53%(2/19)	50.00% (1/2)	No
P120	34	2	81.82% (18/22)	11.11% (2/18)	0.00% (0/3)	No
P121	30	2	80.56% (29/36)	13.79%(4/29)	100.00% (3/3)	Yes
P122	36	4	71.60% (58/81)	10.34% (6/58)	50.00% (2/4)	Yes
P123	33	2	71.43% (15/21)	6.67% (1/15)	0.00% (0/0)	No
P124	34	2	76.92% (20/26)	10.00% (2/20)	50.00% (1/2)	Yes
Total (mean)	33.08	3.29	74.21% (564/760)	8.33 % (47/564)	38.10% (16/42)	-

Table 3. Preimplantation development and embryo transfer outcomes from the ICSI cycles of patients with embryo developmental problems. These patients were only included in the genetic screening.

Most patients showed normal fertilization (mean FR of 74.2%), but with only a limited number of blastocysts formed (mean blastocyst rate is 8.3%). Early embryo arrest is thus common in these patients. Live birth was observed in 9 out of 23 patients (39.1%) after ICSI treatment. More details can be found in *Table 3*.

Patient	Age (at last ICSI cycle)	Ejaculate volume (ml)	Sperm concentration (M/ml)	Sperm motility (a+b) %	Morphology (Normal shape, %)	Origin (partner/donor)
P101	40	2.50	0.50	5	2	Partner
P102	33	4.00	64.50	42	NA	Partner
P103	37	1.80	50.46	28.6	2	Partner
P104	44	2.20	32.84	26.4	5	Partner
P105	37	2.90	91.18	81.8	NA	Partner
P106	38	1.10	77.81	24.3	4	Partner
P107	29	6.2	32.33	50	5	Partner (3 cycles with TESE sample)
P108	37	2.00	77.55	27.3	5	Partner
P109	38	6.00	6.15	10.1	3	Partner
P110	48	4.80	22.79	12.1	1	Partner
P111	31	0.40	0.00	0	0	Partner (3 cycles with TESE sample)
P112	30	1.20	20.15	11.1	6	Partner
P113	36	4.00	67.92	29.3	4	Partner
P114	NA	0.5	75.43	26	NA	Donor
P115	No informed consent, this patient will not be reported in this thesis.					
P116	43	2.10	67.80	20.3	4	Partner
P117	29	1.40	51.50	12	0	Partner
P118	28	3.20	7.00	30	2	Partner
P119	37	2.70	21.65	32.7	8	Partner
P120	31	3.30	0.10	37	0	Partner
P121	34	3.00	4.66	14.3	1	Partner
P122	36	-	-	-	-	Partner (TESE)
P123	44	-	-	-		Partner (TESE)
P124	33	3.20	6.53	53	2	Partner
Total (mean)	35.61	2.76	25.17	28.82	3	-

Table 4. Sperm sample characteristics used in the last ICSI cycle of the patients with embryo developmental problems included in the genetic screening.

WHO reported a list of reference values for human semen and suggests 1.2ml as a lower limit for normal semen volume, for normal concentration 9M/ml, for normal motility 34% and for normal morphology 3% [58]. The WHO semen characteristics are the most commonly used in IVF clinics.

4.2 Data analysis and classification of interesting variants

All found variants using MiSeq were analyzed and categorized, based on population frequency and VarSome criteria as described before. If a variant was found to be class 3 or higher, Sanger was performed to confirm the mutation. More details about interesting variants listed in *Table 5* can be found in *Additional file 5*. Variants categorized lower than class 3 are assumed to be non-pathogenic and will therefore not be further discussed in this thesis.

Code	Gene	c.Nomenclature	Classification criteria	Classification	Sanger confirmed?
101A	NLRP7	c.930G>T	PM1, PM2, BP4	Uncertain significance	Yes
102A	NLRP2	c.2401G>A	BS1, BP1, BP4	Likely benign	
107A	NLRP5	c.623-12_623-11insTTC	BP4	Uncertain significance	Yes
108A	NLRP2	c.1075T>C	BP1, BP4	Likely benign	
111A	NLRP7	c.467G>A	BM2, BP1, BP4	Benign	
112A	NLRP7	c.1104T>C	BA1, BP4, BP6, BP7	Benign	
113A	NLRP2	c.1060A>G	BP4, BP6	Likely benign	
118B	PLCZ1	c.1499C>T	BA1, BP4, PM1, PP2	Benign	Yes
119A	TLE6	c.893C>T	PM2, BP4	Uncertain significance	Yes
119A	NLRP7	c.1257G>C	PM2, BP4, BP7	Likely benign	
121A	NLRP7	c.2094G>A	BS2, BP4, BP7	Likely benign	
121A	PADI6	c.286G>A	PM2, BP4	Uncertain significance	Yes
122A	NLRP2	c.1060A>G	BS1, BP4, BP6	Likely benign	
123A	NLRP2	c.1346C>T	BP4	Uncertain significance	Yes

Table 5. Variants of interest found in screening for variants in the SCMC genes and PLCZ1 in 23 couples and their classification based on VarSome criteria.

Out of all detected variants, five variants of uncertain significance (class 3) were found in five female patients, and all were confirmed on Sanger sequencing. An interesting variant in *PLCZ1* was found in one male patient and, although it was classified as class 1, will be further discussed in section 4.5.

4.3 Variant in *NLRP7*: c.930G>T

4.3.1 Patient and variant information

In couple 101, in the female gene *NLRP7*, we found a heterozygous variant with an uncertain significance and is thereby classified as “Class 3 (uncertain significance)” based on two ‘pathogenic moderate’ and one ‘benign supporting’ criteria *(Table 6)*.

Patient	P101A			
Gene	*NLRP7*			
Name	c.930G>T	p.Gln310His	rs145973556	NM_001127255
Hetero-/homozygous?	Heterozygous			
Consequence	Missense			
Classification	Uncertain significance (class 3)			
Criteria	PM1: mutational hot spot or well-studied functional domain without benign variation PM2: absent in population databases BP4: multiple lines of computational evidence suggest no impact on gene/gene product			
Frequency (MAF)	0.0017	**Frequency non-Finish Europe (MAFnfe)**		0.003

Table 6: Details of the heterozygous missense variant of NLRP7, found in patient 101A

This variant was first found with MiSeq sequencing with good coverage (376) with a variant allele frequency of 43.88298%, suggesting a heterozygosity in this patient. The presence of the variant was later confirmed using Sanger sequencing *(Figure 4)*.

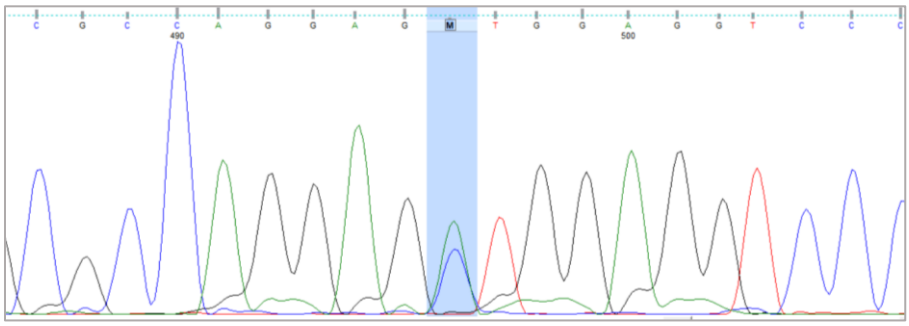


Figure 4. Results Sanger sequencing of NLRP7 in patient 101A: Variant c.930G>T, resulting in a heterozygous missense (substitution of G by T resulting in a Gln substitution by His) could be confirmed.

Clinical information of P101A showed a fertilization rate of 73.08% (19/26) with a blastocyst rate of 5.26% (1/19) and a pregnancy rate of 100.00% (1/1) but no live birth in either of the three ICSI cycles.

4.3.2 Impact of mutation on gene function and its correlation to the phenotype

NLRP7 consists of three different domains (*Figure 5*) and is 1038bp long, with mutation p.Gln310His (p.Q310H) located on bp310, in exon 4 on position 578 of 1579 (coding), on chromosome 19 (19q13.42). This mutation affects the NACHT domain, termed after four proteins containing an NTPase domain with significant similarities. The NACHT domain controls programed cell-death during development by regulating cytochrome c efflux from the mitochondria, which stimulates apoptosis [59].

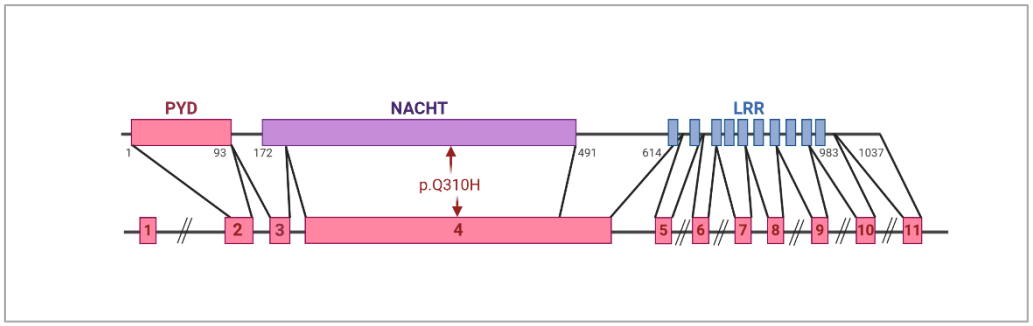


Figure 5. Structure of NLRP7, p.Q310H is located in exon 4 responsible for the NACHT domain. Figure created by BioRender.

As a part of the SCMC, NLRP7 is known to play a role in chromatin reprogramming, DNA methylation during early embryonic development and embryonic cellular immune response. A frameshift mutation (p.Gln310HisfsX38) has previously been described on this exact location, as well as another mutation with substitution of Gln by Arg (p.Q310R), both mutations were associated with infertility [60]. Mutations in *NLRP7* have commonly been reported as responsible for recurrent hydatidiform moles, and affected women show pregnancy issues after 7 to 14 weeks of gestation.

This phenotype agrees with the clinical outcomes in P101A (pregnancy occurs, but is interrupted).

4.4 Variant in *NLRP5*: c. 623-12_623-11insTTC

4.4.1 Patient and variant information

In couple 107, in the female gene *NLRP5*, we found a heterozygous variant with an uncertain significance and is thereby classified as “Class 3 (uncertain significance)” based on one ‘benign supporting’ criteria *(Table 7)*.

Patient	P107A			
Gene	*NLRP5*			
Name	c.623-12_623-11insTTC	c.623-9_623-7dup	rs538671525	NM_153447
Hetero-/homozygous?	Heterozygous			
Consequence	Non coding (intronic)			
Classification	Uncertain significance (class 3)			
Criteria	BP4: multiple lines of computational evidence suggest no impact on gene/gene product			
Frequency (MAF)	0.003	**Frequency non-Finish Europe (MAFnfe)**		0.002

Table 7. Details of the heterozygous intronic variant of NLRP5, found in patient 107A

This variant was first found with MiSeq sequencing with good coverage (198) with a variant allele frequency of 45.9596%, suggesting a heterozygosity in this patient. The presence of the variant was later confirmed using Sanger sequencing *(Figure 6)*.

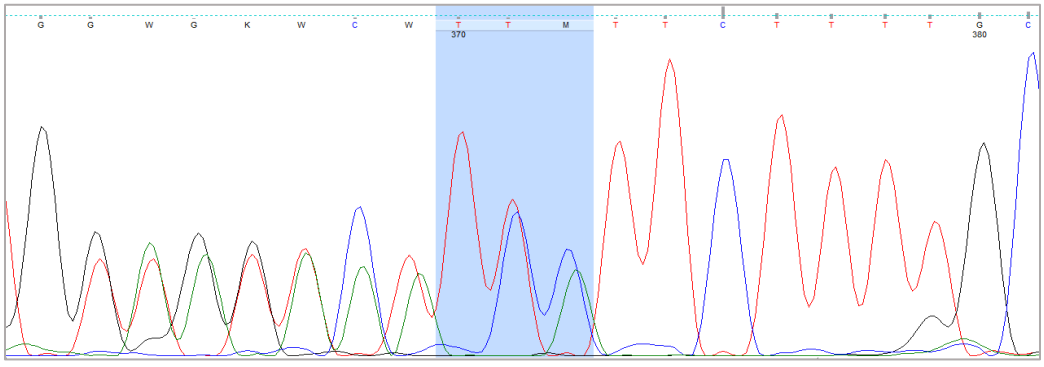


Figure 6. Results Sanger sequencing of NLRP5 in patient 107A: Variant c.623-12_623-11insTTC, resulting in a heterozygous intronic insertion (TTC insertion) could be confirmed. ‘M’ codes for ‘either A or C’.

Clinical information of P107A showed a fertilization rate of 57.89% (11/19) with a blastocyst rate of 0.00% (0/11), a pregnancy rate of 0.00% (0/0) and no live birth in either of the two ICSI cycles.

4.4.2 Impact of mutation on gene function and its correlation to the phenotype

NLRP5 consists of three different domains (*Figure 7*) and is 1201bp long, with mutation c.623-9_623-7dup located on intron 5 before position 971 of 976 (non-coding, splicing), 7bp upstream from a splice site, on chromosome 19 (19q13.43). Its intronic location is not known to affect a specific domain of the *NLRP5* gene.

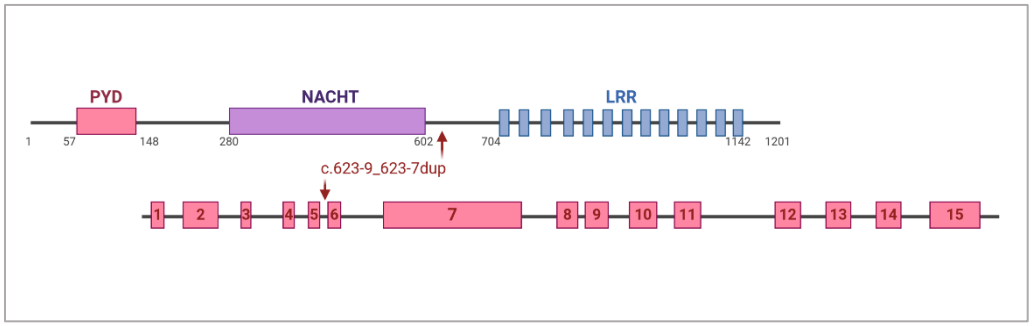


Figure 7. Structure of NLRP5, c.623-9_623-7dup located in intron 5. Figure created by BioRender.

As a member of the SCMC, the NLRP5 protein has a role in preimplantation embryo development and has both cytoplasmic and nuclear functions, interacting with anti-apoptotic pathways and organelle translocation. Mutations in *NLRP5* were earlier associated with oocyte maturation defects [22]. Comparing this phenotype with our patient's clinical outcomes, the low fertilization rate and the extremely low blastocyst rate could find an explanation within this pathway, since defects in mouse *Nlrp5* and human *NLRP5* earlier described developmental arrest at the two-cell stage [15, 17]. However, the mutation was found in an intronic region and is thus most likely not pathogenic.

4.5 Variant in *PLCZ1*: c.1499C>T

4.5.1 Patient and variant information

In couple 118, in the male gene *PLCZ1*, we found a homozygous variant classified as “Class 1 (benign)” based on two pathogenic and two benign criteria *(Table 8)*. However, MAF is only too high in Ashkenazi Jewish population group, which does not include our patient. BA1 criteria is unjustified, which results in the classification of this variant as “Class 3 (uncertain significance)”. Further, this specific variant has been found in high frequency in patients with fertilization failure, in 9 out of 37 (24.3%) patients analyzed by the same study [61]. This also indicates an association with the phenotype.

Patient	P118B			
Gene	*PLCZ1*			
Name	c.1499C>T	p.Ser500Leu	rs10505830	NM_033123
Hetero-/homozygous?	Homozygous			
Consequence	Missense			
Classification	Benign (class 1) / Uncertain significance (class 3)			
Criteria	PM1: mutational hot spot or well-studied functional domain without benign variation PP2: missense in gene with low rate of benign missense variants and pathogenic missenses common BP4: multiple lines of computational evidence suggest no impact on gene/gene product (BA1: MAF is too high for disorder)			
Frequency (MAF)	0.0312	**Frequency non-Finish Europe (MAFnfe)**		0.0462

Table 8. Details of the homozygous missense variant of PLCZ1, found in patient 118B

In the first MiSeq screening, coverage of the *PLCZ1* gene was too low and the variant was yet to be found. Sanger sequencing was performed for this gene and the variant was identified and confirmed *(Figure 8)*. To confirm the result, new DNA was extracted from the saliva sample of this patient and Sanger was repeated. The mutation was found again.

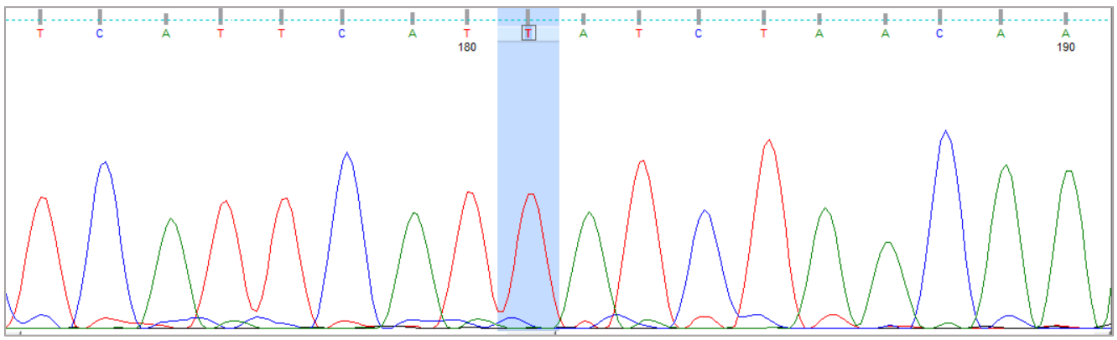


Figure 8. Results Sanger sequencing of PLCZ1 in patient 118B: Variant c.1499C>T, resulting in a homozygous missense (substitution of C by T resulting in a Ser substitution by Leu) could be confirmed.

Clinical information of P118B showed a sperm concentration of 7.00M/ml, a sperm motility of 30% with 2% showing a normal morphology. All of these parameters indicate a higher risk of fertilization failure. However, surprisingly, the fertilization rate in this patient was normal (74%), indicating a normal activating capacity of the sperm. This stands in contrast to most patients with this mutation, of which low or total fertilization failure was reported. A possible explanation lies in the low embryo development (5%), which might be caused by insufficient calcium release, resulting in insufficient oscillations.

4.5.2 Impact of mutation on gene function and its correlation to the phenotype

The mutation p.S500L was found at the C2 N-terminal domain (*Figure 9*) and is located on exon 13 (position 38 of 130) of the *PLCZ1* gene (protein length 609bp) on chromosome 12 (12p12.3). The C2 domain *(Figure 10)* is required for PLCζ binding to PI_3 and PI_5P-containing liposomes [61]. Correct functioning of the PLCζ protein enables adequate Ca^{2+} oscillations and oocyte activation, as earlier explained in sections 1.4.1 and 1.4.2.

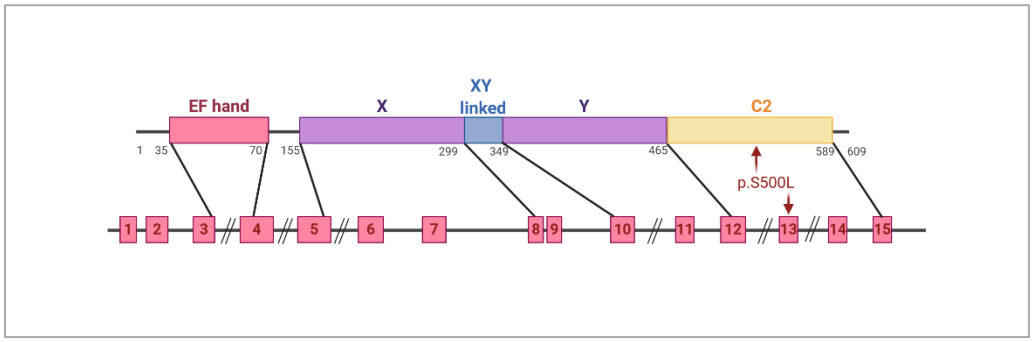


Figure 9. Structure of PLCZ1. p.S500L is located on the C2 domain. Figure created by BioRender.

The p.S500L mutation shows an in vitro phenotype of an alteration of local protein fold, resulting in oocyte activation deficiency with reduced, absent or wrongly located PLCζ [12]. Compared to clinical information and the patient's fertilization failure, this mutation could certainly be causal.

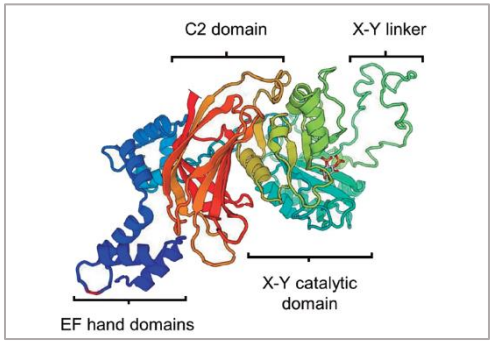


Figure 10. Wild-type PLCζ [61]

4.6 Variant in *TLE6*: c.893C>T

4.6.1 Patient and variant information

In couple 119, in the female gene *TLE6*, we found a heterozygous variant with an uncertain significance and is thereby classified as “Class 3 (uncertain significance)” based on one ‘pathogenic moderate’ and one ‘benign supporting’ criteria *(Table 9)*.

Patient	P119A			
Gene	*TLE6*			
Name	c.893C>T	p.Thr298Met	rs141406529	NM_001143986
Hetero-/homozygous?	Heterozygous			
Consequence	Missense			
Classification	Uncertain significance (class 3)			
Criteria	PM2: absent in population databases BP4: multiple lines of computational evidence suggest no impact on gene/gene product			
Frequency (MAF)	0.00004	**Frequency non-Finish Europe (MAFnfe)**		0.000027

Table 9. Details of the heterozygous missense variant of TLE6, found in patient 119A

This variant was first found with MiSeq sequencing with good coverage (1081) with a variant allele frequency of 49.58372%, suggesting a heterozygosity in this patient. The presence of the variant was later confirmed using Sanger sequencing *(Figure 11)*.

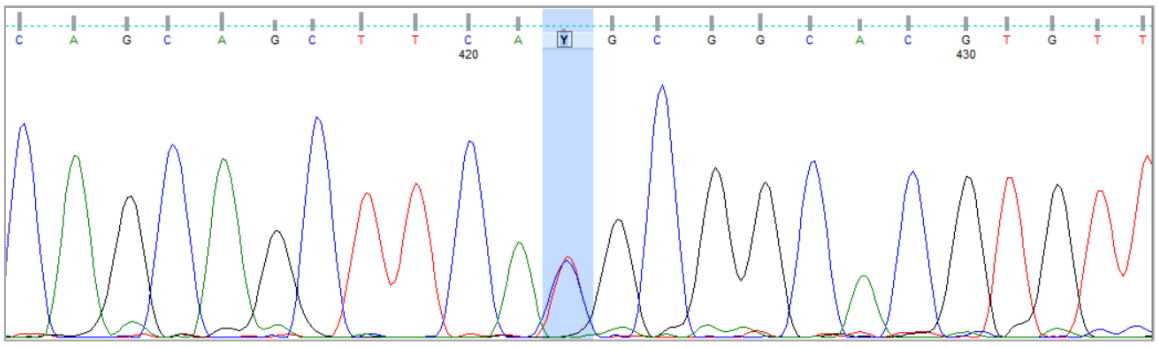


Figure 11. Results Sanger sequencing of TLE6 in patient 119A: Variant c.893C>T, resulting in a heterozygous missense (substitution of C by T resulting in a Thr substitution by Met) could be confirmed.

Clinical information of P119A showed a fertilization rate of 70.37% (19/27) with a blastocyst rate of 10.53% (2/19) and a pregnancy rate of 50.00% (1/2) but no live birth in either of the four ICSI cycles.

4.6.2 Impact of mutation on gene function and its correlation to the phenotype

TLE6 consists of seven WD40 repeats (protein 573bp long) [23], with mutation p.Thr298Met (p.T298M) located on bp298, in exon 12 on position 153 of 253 (coding) on chromosome 19 (19p13.3).

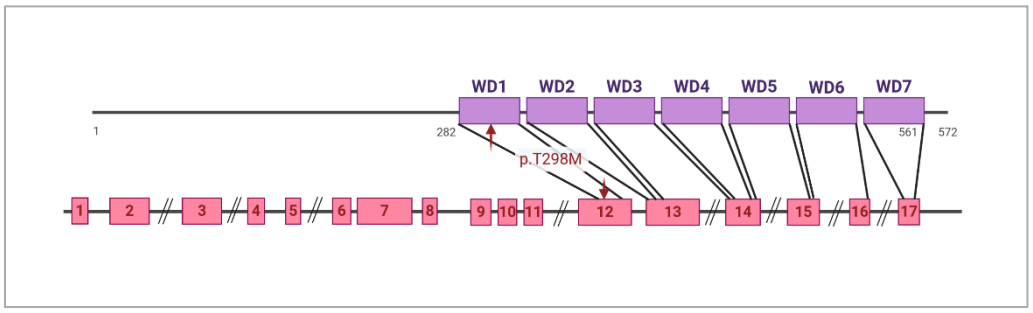


Figure 12. Structure of TLE6, consisting of seven repeats of WD40 between bp282 and bp561. p.T298M is located on the WD1 domain. Figure created by BioRender.

As explained in 1.5.1.2, TLE6 is known to be a member of the SCMC (*Figure 13*) and a PKA substrate, which is in turn responsible for inhibition of meiotic oocyte resumption and for phosphorylating TLE6, essential for meiotic maturation (spindle assembly and migration). Absence impairs its binding to the SCMC and leads to preimplantation embryonic lethality.

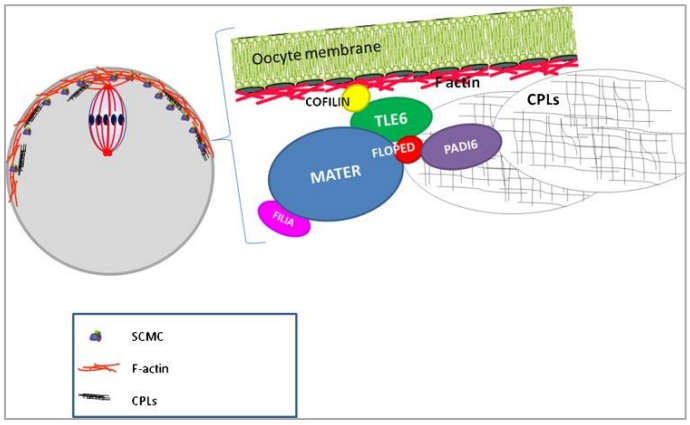


Figure 13. Schematic representation of the hypothesized structure of the SCMC and localization within the MII oocyte, showing TLE6 as a crucial part in its function [15]

In this patient, early embryo arrest could be a possible explanation for the clinical outcomes, since blastocyst rate was low. However, two blastocysts were formed and one pregnancy occurred, implying another factor may be in the picture. Another potential hypothesis states that the mutation only partially affects TLE6 function, which might explain why some blastocyst, although in low range, were formed.

4.7 Variant in *PADI6*: c.286G>A

4.7.1 Patient and variant information

In couple 121, in the female gene *PADI6*, we found a heterozygous variant with an uncertain significance and is thereby classified as “Class 3 (uncertain significance)” based on one ‘pathogenic moderate’ and one ‘benign supporting’ criteria *(Table 10)*.

Patient	P121A			
Gene	*PADI6*			
Name	c.286G>A	p.Ala96Thr	rs181841908	NM_207421
Hetero-/homozygous?	Heterozygous			
Consequence	Missense			
Classification	Uncertain significance (class 3)			
Criteria	PM2: absent in population databases BP4: multiple lines of computational evidence suggest no impact on gene/gene product			
Frequency (MAF)	0.001099	**Frequency non-Finish Europe (MAFnfe)**		0.001857

Table 10. Details of the heterozygous missense variant of PADI6, found in patient 121A

This variant was first found with MiSeq sequencing with good coverage (495) with a variant allele frequency of 46.86869%, suggesting a heterozygosity in this patient. The presence of the variant was later confirmed using Sanger sequencing *(Figure 14)*.

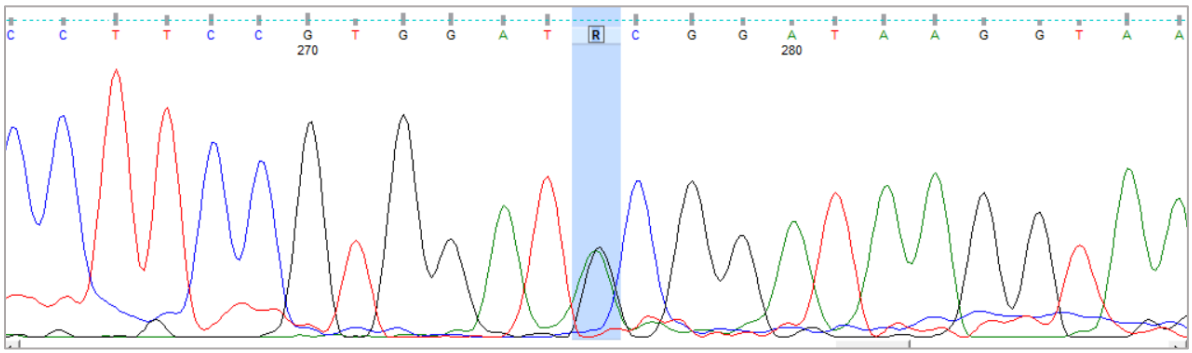


Figure 14. Results Sanger sequencing of PADI6 in patient 121A: Variant c.286G>A, resulting in a heterozygous missense (substitution of G by A resulting in a Ala substitution by Thr) could be confirmed.

Clinical information of P121A showed a fertilization rate of 80.56% (29/36) with a blastocyst rate of 13.79% (4/29) and a pregnancy rate of 100.00% (3/3) but live birth after two ICSI cycles.

4.7.2 Impact of mutation on gene function and its correlation to the phenotype

PADI6 consists of three domains (*Figure 15*) and is 695bp long, with mutation p.Ala96Thr (p.A96T) located on bp96, in exon 2 on position 170 of 178 (splicing, coding) on chromosome 1 (1p36.13). This mutation affects the PAD N-terminal domain, resulting in a loss-of-function of the PADI6 protein, involved in cytoskeletal reorganization in egg and early embryo.

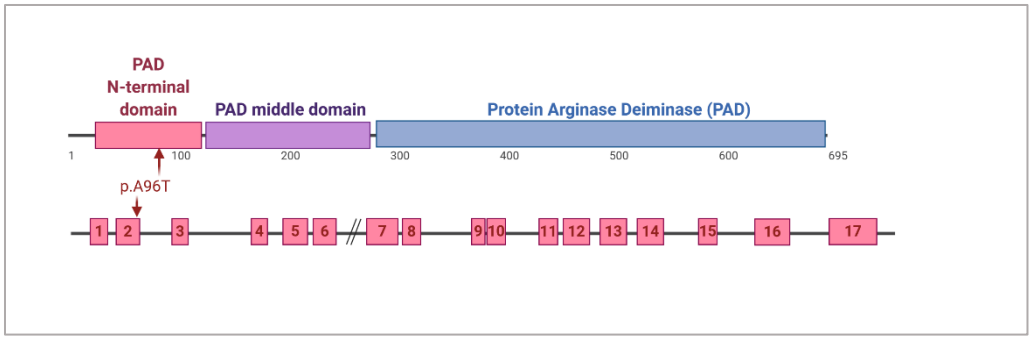


Figure 15. Structure of PADI6 and its domains. p.A96T is located on the PAD N-terminal domain. Figure created by BioRender.

PADI6 also contributes to the SCMC and plays a role in *de novo* protein synthesis and organelle positioning and movement. Pathogenic mutations in *PADI6* have therefore been associated with a disturbed morphology during early embryo development (*Figure 16*).

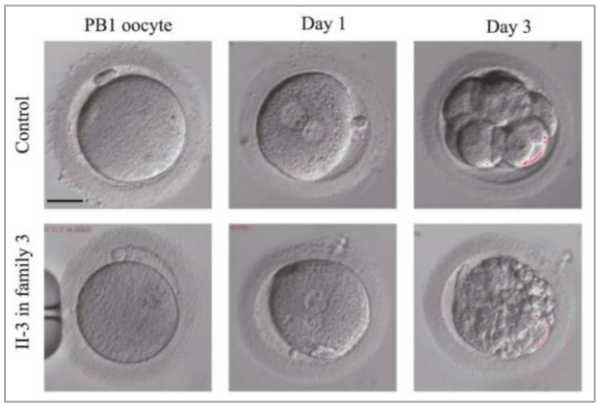


Figure 16. Morphology of a control oocyte and an oocyte with a pathogenic mutation in PADI6 during early development [8]

Mutations in *PADI6* have previously been associated with preimplantation embryonic lethality (*Figure 16*). The patient however, successfully completed pregnancy, resulting in live birth after two ICSI cycles, indicating a subfertility rather than total protein activity loss.

4.8 Variant in *NLRP2*: c.1346C>T

4.8.1 Patient and variant information

In couple 123, in the female gene *NLRP2*, we found a heterozygous variant with an uncertain significance and is thereby classified as “Class 3 (uncertain significance)” based on one ‘benign supporting’ criteria *(Table 11)*. Since VarSome is in constant change, this variant was later classified as class 2, but is still considered as an interesting variant and is therefore discussed.

Patient	P123A			
Gene	*NLRP2*			
Name	c.1346C>T	p.Thr449Met	rs144271525	NM_001174081
Hetero-/homozygous?	Heterozygous			
Consequence	Missense			
Classification	Uncertain significance (class 3)			
Criteria	BP4: multiple lines of computational evidence suggest no impact on gene/gene product			
Frequency (MAF)	0.000182	**Frequency non-Finish Europe (MAFnfe)**		0.000356

Table 11. Details of the heterozygous missense variant of NLRP2, found in patient 123A

This variant was first found with MiSeq sequencing with good coverage (929) with a variant allele frequency of 49.83854%, suggesting a heterozygosity in this patient. The presence of the variant was later confirmed using Sanger sequencing *(Figure 17)*.

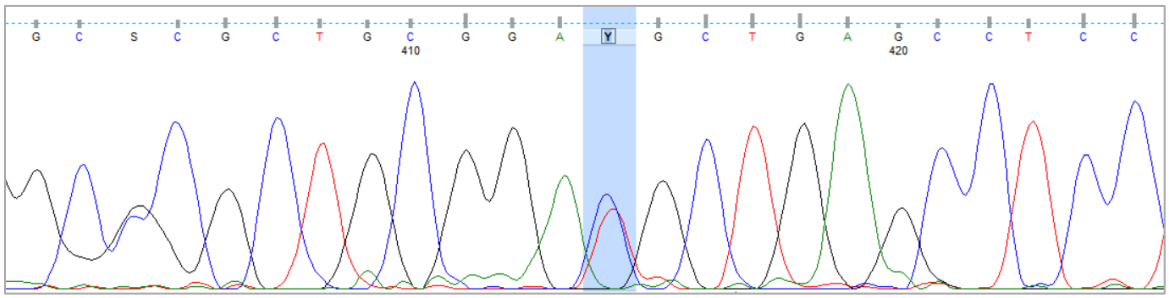


Figure 17. Results Sanger sequencing of NLRP2 in patient 123A: Variant c.1346C>T, resulting in a heterozygous missense (substitution of C by T resulting in a Thr substitution by Met) could be confirmed.

Clinical information of P123A showed a fertilization rate of 71.43% (15/21) with a blastocyst rate of 6.67% (1/15) and a pregnancy rate of 0.00% (0/0) with no live birth in either of the two ICSI cycles.

4.8.2 Impact of mutation on gene function and its correlation to the phenotype

NLRP2 consists of three domains (*Figure 18*) and is 1063bp long, with mutation p.Thr449Met (p.T449M) located on bp449, in exon 6 on position 883 of 1567 (coding) on chromosome 19 (19q13.42). The mutation is located in the NACHT domain, that can be compared to the NACHT domain in *NLRP7*, controlling programed cell death during development.

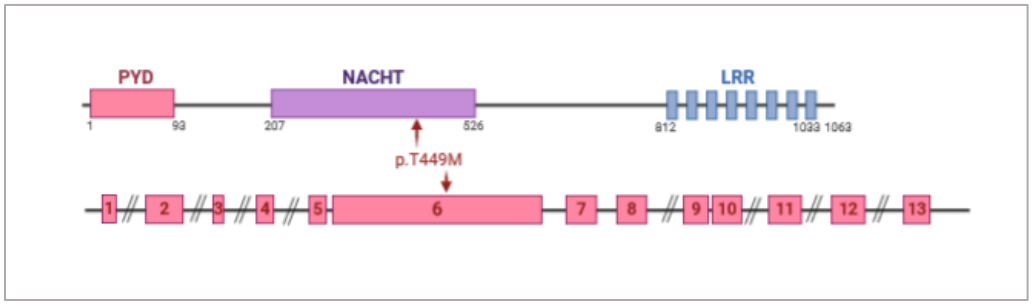


Figure 18: Structure of NLRP2, containing three domains. The found variant p.T449M is located in the NACHT domain. Figure created by BioRender.

NLRP2, as a part of the SCMC, is essential for the effective degradation of maternal proteins, resulting in a regulated pathway of early embryonic development. As in previous research where mutations in *NLRP2* resulted in low blastocyst rates caused by impaired protein stability, patient P123A also showed a low blastocyst rate. This variant could therefore possibly be involved in the pathogenicity, resulting in preimplantation developmental arrest.

5. Discussion

5.1 Study limitations

The interpretation of the found variants is not easy, since no definite conclusion can be drawn for now. Only the variants classified as class 3 were discussed in Results, but VarSome criteria change day by day and the investigated genes are still poorly understood. Therefore, it is possible that a causal mutation found in this research was not discussed or even listed as an 'interesting variant'.

Secondly, only exonic regions were screened. It is noteworthy that intronic fragments such as promotor regions also have an influence on the according gene. However, we did not screen for this, which does not allow a full understanding of gene function and activity. Whole exome sequencing and whole genome sequencing could be an interesting future perspective.

Thirdly, only a limited number of patients was included in our study. A larger research group could provide more insight in the association between the gene function and its (so far suspected) phenotype.

At last, the first broad targeted screening was performed using Next Generation Sequencing. Although the quality of this technology improves every day, sensitivity is lower than when using Sanger sequencing. This may result in some variants being missed in the process of sequencing. Further, the MiSeq reported low coverage of one or more genes in certain patients. These target regions were later repeated, but if low coverage occurred again, there was no third repetition of the sequencing. Hereby, the study might have missed some interesting genes that might be (partly) causal for infertility.

5.2 Previous reports of *PLCZ1* mutation p.S500L

Interestingly, the *PLCZ1* variant p.Ser500Leu (p.S500L) has previously been reported in multiple patients [61-63]. *Ferrer-Vaquer et al. (2016)* reported one heterozygous patient with fertilization failure after ICSI [62]. Also, *Yoon et al. (2008)* reported one homozygous patient with fertilization failure, although there is low, but remaining Ca^{2+} oscillatory activity [63]. Contrarily, the homozygous p.S500L mutation has been reported in 5/3347 healthy Dutch fathers [64]. Although the mutation occurs in 3.1% of natural reference population and even in 4.6% of the non-Finish European population, we cannot exclude that the affected men are not subfertile or infertile.

Torra-Massana et al. (2019) reported nine patients with this mutation out of 37 screened. All nine patients were known with fertilization failure, eight of which presented oocyte activation failure. Functional analysis was performed by cRNA injection of wild-type and mutated forms of *PLCZ1* into human in vitro matured metaphase II oocytes and fertilization outcomes were scored later. The p.S500L mutation was predicted to be detrimental for PLCζ/*PLCZ1* function, although surprisingly oocyte activation did not significantly decrease (60%) compared to wild-type *PLCZ1*. Of the nine affected patients, one patient was homozygous for the mutation, which correlated with a severe fertilization failure in three cycles [61], consistent with a previous report that sperm homozygous for this mutation triggers low calcium oscillation activity when injected into mouse oocytes [63]. Since a *PLCZ1* defect is known to result in oocyte activation deficiency, patients with this mutation qualify for AOA treatment by artificially mimicking PLCζ function. Eight out of the nine affected patients underwent clinical intervention with AOA, resulting in five live births out of seven for heterozygous patients and no pregnancy for the homozygous patient [61].

Excitingly for our research lab, P118B with this mutation hereby also qualifies for AOA treatment, although the phenotype of our patient didn't show fertilization failure, but embryo developmental arrest instead. The sperm showed good oocyte activating capacity on MOCA, which is surprising since the patient has a homozygous mutation in *PLCZ1.* The calcium release being sufficient for oocyte activation, but not for further development, might be an acceptable hypothesis. The patient should consider clinical intervention by AOA, hopefully resulting in a live birth.

5.3 Further investigations for variants of uncertain significance

Since all variants we found are class 3 or lower, there is no significant causal female gene found in our study. It is difficult to create a clear conclusion about these variants, since not much research is available. Functional analysis is necessary to obtain evidence for pathogenic classification of the variants. There are some other approaches to investigate the class 3 variants. However, labor intensity and cost limit further research for these patients.

Better understanding and identification of genetic variants may be useful to patients on several levels. Knowledge about the causal mutation could not only lead to personalized and targeted treatments, but could also enable screening of prospective IVF embryos to ensure pathogenic infertility variants are not transmitted to future children [65].

5.3.1 Trio analysis to identify non-pathogenic inherited variants

Another approach to the further investigation of the found class 3 variants is trio analysis, where the parents of the infertile couple are also screened for the same variants. This might give more insight in whether the mutations are *de novo* or inherited. Presence of the variant in one or both of the patient's parents suggests a benign character and the variant can therefore be reclassified as probably not pathogenic. Causality of these variants can be proven or ruled out, the latter indicating that one or more other variants could also play a role in the pathogenicity.

5.3.2 Detection of hidden mutations using Whole Exome Sequencing

None of the female patients in our research population have a clear clinical diagnosis, all have negative or uncertain test results for the screened genes known to be associated with infertility. Normally, heterozygosity should be a protective factor against a mutant phenotype, but since all patient included in our study have fertility issues and since there is no certainty about the causality of the found variants, a multifactorial hypothesis is suspected. We must raise the question whether the patients are truly heterozygous, as only targeted gene sequencing of exons and a limited part of the introns were screened. Further intronic or genome wide screening for variants in e.g. promotor sites could be beneficial. Consequently, the power of Whole Exome Sequencing and Whole Genome Sequencing for the discovery of new disease genes or loci is unseen and has found applications in many unsolved cases [52]. It has not only been used in infertility research [65], but recently, WES has made its entrance in prenatal genetics [66, 67]. However, ethics about variants of unknown significance, privacy and incidental findings limit its use [66].

For WGS, the genome is sequenced without prior selection, whereas for WES, currently the more commonly used method for diagnostic applications, DNA regions containing the protein-coding exons are first selectively captured [67]. WES is more cost-effective than WGS and is preferred because the ability to interpret intronic regions of the genome is currently extremely limited [66]. There are some drawbacks however, mostly its high cost for sequencing, storage, analysis and interpreting data. Furthermore, WES and WGS increase the chance of incidental findings, which limits their current use [52].

Unfortunately, they are both restricted in use since knowledge about these sites is limited and intronic variants are even more complicated to interpret. Therefore, certainty about whether a variant is truly heterozygous or rather compound heterozygous is currently not possible.

5.3.3 Immunostaining as an assessment of protein location and quantity

Another approach for further genetic research is immunostaining, using an antibody against a specific targeted protein and resulting in a visualization of both its location and quantity.

Immunostaining has already been suggested as a potential therapeutic aid [68-70]. More specifically for infertility, PLCζ might be a possible prognostic biomarker for oocyte activation. Immunofluorescence analysis indicates that PLCζ levels are significantly higher in fertile sperm compared with OAD sperm [68]. This may help in determining the criteria and requirements of the fertility treatment, significantly decreasing the number of cycles needed for successful pregnancy [70]. However, a large degree of variability was reported regarding total PLCζ levels in human sperm among control subjects and possibly within samples from the same man, potentially leading to a misdiagnosis in a clinical setting [68]. Clinical use is therefore still limited, although it had been widely regarded as a worldwide cost-effective approach [70].

Further, immunostaining of female factors has also been performed, partly in the identification of the SCMC genes and their interaction within the oocyte, but also in function analysis of variants of uncertain significance. The quantity of a SCMC protein in the oocyte and the morphology of the oocyte itself indicate to what extent the mutation eliminates gene function and activity [4]. Immunostaining has been performed to assess activity of multiple genes associated with infertility, such as *PADI6* and *TUBB8* [4, 71, 72], for which oocyte and spindle morphology were evaluated. Associations between mutated genes and abnormal morphology were reported [72].

This procedure may contribute to a better knowledge about the variants found in our patients, since it might give more insight in the level of remained activity, and therefore might play a role in future therapeutic options. Therapeutic use of immunostaining for our patients is currently limited due to scarcity of IVF oocytes for research purposes.

5.3.4 Gene knockout models in mouse oocytes

Experimental research, within the concept of ‘trial and error’, has proven or rejected hypotheses for years using animal testing. For logistic and economic reasons, mice and rats are the most widely used laboratory animals. Their genomic, proteomic and metabolomic profiles as well as their organic functions and behavior are better known than for other species, although some results differ in animals and humans due to a greater human complexity [73].

The found variants in our research may be further investigated using knockout models in animals. Knockout models of the SCMC genes (i.a. *PADI6*, *NLRP2*, *NLRP5*, *KHDC3L*, *TLE6*) in mice have already been used in previous research [34, 36, 74-78], likewise for *PLCZ1 [63, 79, 80]*, and has been proven to cause embryo developmental issues. Not only a clearer knowledge about in what stage embryo development is arrested and hereby identifying future therapeutic possibilities is achievable, but this can be done at a lower cost and avoids the intensive process of obtaining human oocytes.

5.3.5 Functional analysis by injection of mutant cDNA into HeLa cells

Five variants of uncertain significance were found in five female patients, but the pathogenicity of these mutations should be further analyzed in order to create a better understanding of their impact on the specific infertility phenotype. Functional analysis can be performed by incorporation of these mutations into full-length cDNA (complementary DNA), with a FLAG tag attached to the C-terminal end. Both mutant and wild-type cDNA are then transfected into different HeLa cells. The variants' impact on gene function can be assessed by comparing expression levels of the affected protein to the expression of the wild-type protein. Western blotting separates these proteins on gel electrophoresis and visualizes them using anti-FLAG antibodies, enabling the comparison [2]. This manner of variant analysis has recently been used in research of mutated SCMC genes such as *NLRP2, NLRP5, TLE6* and *PADI6*, in which protein defects or decreased expression levels caused by mutations in these genes could be demonstrated [2, 4, 23]. However, this procedure has earlier been performed to identify the multiple components of the SCMC and their interaction within oocytes [24, 36]. The latter has also been investigated by immunostaining, as mentioned in section 5.3.3.

5.4 Risks of consanguineous families and their role in genetic research

Although many of our patients have an infertile phenotype, only a limited number of possible causal variants was found. This information, accompanied by previous research, implies that genetic causes of infertility are multifactorial, and that infertility in our and many other patients is probably caused by multiple mutations affecting one or more genes. This hypothesis is confirmed by reports of the association between infertility and consanguinity. Consanguinity, defined as the intermarriage of two individuals who have at least one ancestor in common, is frequently seen in Mediterranean, and Muslim Mediterranean populations in particular.

Consanguinity may increase the rate of homozygous genotype expression, making offspring of consanguineous marriages at increased risk for recessively inherited disorders [81]. Recent studies suggest that it is highly correlated with rare genetic sperm-defect syndromes [81, 82], but also appears to have an adverse impact on ovarian reserve in the female descendants of consanguine couples [82]. Consequently, ART, particularly ICSI, should be practiced with caution and be accompanied by accurate sperm investigation and genetic counseling, since these syndromes are incurable and can be genetically transmitted to male offspring [81].

Research in infertility has been challenging, since high genetic heterogeneity and the rarity of variants makes their interpretation as causal challenging. Consanguinity studies might overcome both of these problems, providing genetic evidence for involvement in disease [65].

5.5 Future perspectives

5.5.1 Nuclear transfer in overcoming embryo development arrest

For couples who struggle with early embryo arrest, oocyte donation is currently the only possibility for pregnancy, renouncing to genetic parenthood. In this context, nuclear transfer might offer a solution. The nucleus from an oocyte collected from an infertile woman is transferred to a donor oocyte that has its own nucleus removed [83]. Germline nuclear transfer (NT) techniques are currently being applied in humans to prevent the transmission of mutated mitochondrial DNA (mtDNA) diseases [83, 84]. Yet, there is also growing interest in the translational use of NT for treating infertility and improving IVF outcomes [84], since embryogenesis and implantation have been proven to be influenced by cytoplasmic factors [85]. Damaged maternal mitochondria, often related with age-related female infertility, may play a vital role in some unresolvable cases of infertility which could be overcome by the application of NT technologies [83].

It is shown that, in couples that experienced repeated implantation failure as a result of poor embryo development, ooplasm transfer from donor oocytes at metaphase II (MII) stage into patient MII oocytes can be compatible with fertilization and pregnancy [85]. Nevertheless, direct scientific evidence to support such applications is currently lacking and it is unclear which infertility indications may benefit [83, 84]. Further, the benefit and safety of the procedures remains unknown [84].

Although these techniques remain experimental, these novel reproductive technologies may provide new potential avenues for genetic parenthood for patients facing age-related infertility and early embryo developmental arrest [84].

5.5.2 Injection of wild-type cRNA/recombinant protein to restore normal embryo development

Recent research on genes associated with infertility suggests the therapeutic use of wild-type cRNA or recombinant protein injection. *Sang et al. (2018)* determined whether the phenotype of fertilization failure, caused by a *WEE2* mutation, could be rescued. Direct injections of wild-type *WEE2* cRNA into mutated oocytes showed significant successes in fertilization rates after ICSI compared to the control oocytes (100.0% vs 0.0%) [86]. This sets an example for further research on the possibilities of this procedure in overcoming the impact of infertility-associated mutations. As mentioned in section 1.5.3.1, this technique has also been proposed for overcoming PLCζ deficiencies. The recombinant protein is even more preferable, since cRNA has a risk of incorporation in the genome [6]. It is not unrealistic to suggest the same approach for overcoming the effects of pathogenic mutations in the maternal SCMC genes, more specifically overcoming early developmental arrest.

6. Conclusion

The objectives of this study were to broaden the knowledge of mutations in the SCMC genes and in *PLCZ1* and its relation to infertility. The SCMC genes are associated with early embryo development and *PLCZ1* is associated with oocyte activation failure. Screening of *PADI6, TLE6, NLRP2, NLRP5, NLRP7* and *KHDC3L* in 23 female patients with fertility issues showed multiple variants, of which five were classified as class 3. For *PLCZ1*, one variant was classified as class 3 and was reported in previous research as a causal and pathogenic mutation.

Our research did not prove pathogenicity for newly found mutations, but might help expanding the knowledge about genetic causes of infertility since still little research has been done for these genes. Especially for *PLCZ1*, this study contributes to the proof that the p.S500L mutation is indeed causal. This may lead to a better understanding of the mutation and its consequences in affected patients, a more personalized and targeted approach in their treatment and lastly, a manner to prevent inheritance of the mutation to offspring. However, we found this mutation in a patient with embryo developmental arrest but good fertilization rates, which is contradictory to previous research. Further examination of this phenomenon should be initiated to completely understand the underlying pathway.

More research needs to be to done to fully understand the impact of genetics on infertility in general, but more specifically on embryo developmental arrest caused by mutations in the SCMC genes, or oocyte activation failure caused by mutations in *PLCZ1*. Moreover, genetic research might contribute to a better knowledge and treatment of unsolved cases of infertility.

7. References

1. Vander Borght, M. and C. Wyns, *Fertility and infertility: Definition and epidemiology.* Clin Biochem, 2018. **62**: p. 2-10.
2. Mu, J., et al., *Mutations in NLRP2 and NLRP5 cause female infertility characterised by early embryonic arrest.* J Med Genet, 2019. **56**(7): p. 471-480.
3. Cunningham, J., *Infertility: A primer for primary care providers.* Jaapa, 2017. **30**(9): p. 19-25.
4. Xu, Y., et al., *Mutations in PADI6 Cause Female Infertility Characterized by Early Embryonic Arrest.* Am J Hum Genet, 2016. **99**(3): p. 744-752.
5. Kashir, J., M. Nomikos, and F.A. Lai, *Phospholipase C zeta and calcium oscillations at fertilisation: The evidence, applications, and further questions.* Adv Biol Regul, 2018. **67**: p. 148-162.
6. Nomikos, M., J. Kashir, and F.A. Lai, *The role and mechanism of action of sperm PLC-zeta in mammalian fertilisation.* Biochem J, 2017. **474**(21): p. 3659-3673.
7. Khakpour, S., et al., *Zeta method: A noninvasive method based on membrane charge for selecting spermatozoa expressing high level of phospholipaseCzeta.* Andrologia, 2019. **51**(5): p. e13249.
8. Wang, X., et al., *Novel mutations in genes encoding subcortical maternal complex proteins may cause human embryonic developmental arrest.* Reprod Biomed Online, 2018. **36**(6): p. 698-704.
9. Heindryckx, B., et al., *Efficiency of assisted oocyte activation as a solution for failed intracytoplasmic sperm injection.* Reprod Biomed Online, 2008. **17**(5): p. 662-8.
10. Neri, Q.V., et al., *Understanding fertilization through intracytoplasmic sperm injection (ICSI).* Cell Calcium, 2014. **55**(1): p. 24-37.
11. Amdani, S.N., C. Jones, and K. Coward, *Phospholipase C zeta (PLCzeta): oocyte activation and clinical links to male factor infertility.* Adv Biol Regul, 2013. **53**(3): p. 292-308.
12. Kashir, J., *Increasing associations between defects in phospholipase C zeta and conditions of male infertility: not just ICSI failure?* J Assist Reprod Genet, 2020.
13. Nikiforaki, D., et al., *Sperm involved in recurrent partial hydatidiform moles cannot induce the normal pattern of calcium oscillations.* Fertility and sterility, 2014. **102**(2): p. 581-588.e1.
14. Bebbere, D., et al., *Expression of maternally derived KHDC3, NLRP5, OOEP and TLE6 is associated with oocyte developmental competence in the ovine species.* BMC Dev Biol, 2014. **14**: p. 40.
15. Bebbere, D., et al., *The subcortical maternal complex: multiple functions for one biological structure?* J Assist Reprod Genet, 2016. **33**(11): p. 1431-1438.
16. Zheng, W., et al., *New biallelic mutations in PADI6 cause recurrent preimplantation embryonic arrest characterized by direct cleavage.* J Assist Reprod Genet, 2020. **37**(1): p. 205-212.
17. Amoushahi, M., L. Sunde, and K. Lykke-Hartmann, *The pivotal roles of the NOD-like receptors with a PYD domain, NLRPs, in oocytes and early embryo developmentdagger.* Biol Reprod, 2019. **101**(2): p. 284-296.
18. Qian, J., et al., *Biallelic PADI6 variants linking infertility, miscarriages, and hydatidiform moles.* Eur J Hum Genet, 2018. **26**(7): p. 1007-1013.
19. Monk, D., M. Sanchez-Delgado, and R. Fisher, *NLRPs, the subcortical maternal complex and genomic imprinting.* Reproduction (Cambridge, England), 2017. **154**(6): p. R161-r170.
20. Maddirevula, S., et al., *The human knockout phenotype of PADI6 is female sterility caused by cleavage failure of their fertilized eggs.* Clin Genet, 2017. **91**(2): p. 344-345.
21. Duncan, F.E., et al., *Transducin-like enhancer of split-6 (TLE6) is a substrate of protein kinase A activity during mouse oocyte maturation.* Biol Reprod, 2014. **90**(3): p. 63.
22. Yatsenko, S.A. and A. Rajkovic, *Genetics of human female infertilitydagger.* Biol Reprod, 2019. **101**(3): p. 549-566.
23. Alazami, A.M., et al., *TLE6 mutation causes the earliest known human embryonic lethality.* Genome Biol, 2015. **16**: p. 240.
24. Li, G., et al., *NLRP7 is expressed in the ovine ovary and associated with in vitro pre-implantation embryo development.* Reproduction, 2019. **158**(5): p. 415-427.
25. Aghajanova, L., et al., *No evidence for mutations in NLRP7, NLRP2 or KHDC3L in women with unexplained recurrent pregnancy loss or infertility.* Hum Reprod, 2015. **30**(1): p. 232-8.
26. Reddy, R., et al., *Report of four new patients with protein-truncating mutations in C6orf221/KHDC3L and colocalization with NLRP7.* Eur J Hum Genet, 2013. **21**(9): p. 957-64.

27. Sills, E.S., et al., *Pathogenic variant in NLRP7 (19q13.42) associated with recurrent gestational trophoblastic disease: Data from early embryo development observed during in vitro fertilization.* Clin Exp Reprod Med, 2017. **44**(1): p. 40-46.
28. Andreasen, L., et al., *Mosaic moles and non-familial biparental moles are not caused by mutations in NLRP7, NLRP2 or C6orf221.* Mol Hum Reprod, 2012. **18**(12): p. 593-8.
29. Ulker, V., et al., *Novel NLRP7 mutations in familial recurrent hydatidiform mole: are NLRP7 mutations a risk for recurrent reproductive wastage?* Eur J Obstet Gynecol Reprod Biol, 2013. **170**(1): p. 188-92.
30. Manokhina, I., et al., *Maternal NLRP7 and C6orf221 variants are not a common risk factor for androgenetic moles, triploidy and recurrent miscarriage.* Mol Hum Reprod, 2013. **19**(8): p. 539-44.
31. Andreasen, L., et al., *NLRP7 or KHDC3L genes and the etiology of molar pregnancies and recurrent miscarriage.* Mol Hum Reprod, 2013. **19**(11): p. 773-81.
32. Mahadevan, S., et al., *No evidence for mutations in NLRP7 and KHDC3L in women with androgenetic hydatidiform moles.* Prenat Diagn, 2013. **33**(13): p. 1242-7.
33. Zhang, W., et al., *KHDC3L mutation causes recurrent pregnancy loss by inducing genomic instability of human early embryonic cells.* PLoS Biol, 2019. **17**(10): p. e3000468.
34. Peng, H., et al., *Nlrp2, a maternal effect gene required for early embryonic development in the mouse.* PLoS One, 2012. **7**(1): p. e30344.
35. Peng, H., et al., *NLRP2 and FAF1 deficiency blocks early embryogenesis in the mouse.* Reproduction, 2017. **154**(3): p. 245-251.
36. Mahadevan, S., et al., *Maternally expressed NLRP2 links the subcortical maternal complex (SCMC) to fertility, embryogenesis and epigenetic reprogramming.* Sci Rep, 2017. **7**: p. 44667.
37. Heindryckx, B., et al., *Treatment option for sperm- or oocyte-related fertilization failure: assisted oocyte activation following diagnostic heterologous ICSI.* Hum Reprod, 2005. **20**(8): p. 2237-41.
38. Bonte, D., et al., *Assisted oocyte activation significantly increases fertilization and pregnancy outcome in patients with low and total failed fertilization after intracytoplasmic sperm injection: a 17-year retrospective study.* Fertil Steril, 2019. **112**(2): p. 266-274.
39. Vanden Meerschaut, F., et al., *Assisted oocyte activation following ICSI fertilization failure.* Reprod Biomed Online, 2014. **28**(5): p. 560-71.
40. Vanden Meerschaut, F., et al., *Diagnostic and prognostic value of calcium oscillatory pattern analysis for patients with ICSI fertilization failure.* Human reproduction (Oxford, England), 2013. **28**(1): p. 87-98.
41. Ebner, T., et al., *Live birth after artificial oocyte activation using a ready-to-use ionophore: a prospective multicentre study.* Reproductive biomedicine online, 2015. **30**(4): p. 359-65.
42. Montag, M., et al., *The benefit of artificial oocyte activation is dependent on the fertilization rate in a previous treatment cycle.* Reproductive biomedicine online, 2012. **24**(5): p. 521-6.
43. Vanden Meerschaut, F., et al., *Neonatal and neurodevelopmental outcome of children aged 3-10 years born following assisted oocyte activation.* Reprod Biomed Online, 2014. **28**(1): p. 54-63.
44. Mateizel, I., et al., *Obstetric and neonatal outcome following ICSI with assisted oocyte activation by calcium ionophore treatment.* Journal of assisted reproduction and genetics, 2018. **35**(6): p. 1005-1010.
45. Nomikos, M., et al., *Functional disparity between human PAWP and PLCzeta in the generation of Ca2+ oscillations for oocyte activation.* Mol Hum Reprod, 2015. **21**(9): p. 702-10.
46. Ishii, T., *Reproductive medicine involving genome editing: clinical uncertainties and embryological needs.* Reprod Biomed Online, 2017. **34**(1): p. 27-31.
47. Kaser, D.J. and J.M. Franasiak, *Crossing the germline: CRISPR-Cas9 and our responsibility as reproductive endocrinology and infertility physicians.* J Assist Reprod Genet, 2018. **35**(3): p. 399-402.
48. Hershlag, A. and S.L. Bristow, *Editing the human genome: where ART and science intersect.* J Assist Reprod Genet, 2018. **35**(8): p. 1367-1370.
49. Anifandis, G., et al., *Artificial oocyte activation: physiological, pathophysiological and ethical aspects.* Syst Biol Reprod Med, 2019. **65**(1): p. 3-11.
50. Ebner, T., et al., *Treatment with Ca2+ ionophore improves embryo development and outcome in cases with previous developmental problems: a prospective multicenter study.* Hum Reprod, 2015. **30**(1): p. 97-102.
51. Lefever, S., et al., *High-throughput PCR assay design for targeted resequencing using primerXL.* BMC Bioinformatics, 2017. **18**(1): p. 400.

52. De Leeneer, K., et al., *Flexible, scalable, and efficient targeted resequencing on a benchtop sequencer for variant detection in clinical practice.* Hum Mutat, 2015. **36**(3): p. 379-87.
53. Ravi, R.K., K. Walton, and M. Khosroheidari, *MiSeq: A Next Generation Sequencing Platform for Genomic Analysis.* Methods in molecular biology (Clifton, N.J.), 2018. **1706**: p. 223-232.
54. Liu, L., et al., *Comparison of next-generation sequencing systems.* Journal of biomedicine & biotechnology, 2012. **2012**: p. 251364.
55. Beck, T.F., J.C. Mullikin, and L.G. Biesecker, *Systematic Evaluation of Sanger Validation of Next-Generation Sequencing Variants.* Clinical chemistry, 2016. **62**(4): p. 647-54.
56. Richards, S., et al., *Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology.* Genet Med, 2015. **17**(5): p. 405-24.
57. Kopanos, C., et al., *VarSome: the human genomic variant search engine.* Bioinformatics (Oxford, England), 2019. **35**(11): p. 1978-1980.
58. Cooper, T.G., et al., *World Health Organization reference values for human semen characteristics.* Human reproduction update, 2010. **16**(3): p. 231-45.
59. Slim, R. and E.P. Wallace, *NLRP7 and the Genetics of Hydatidiform Moles: Recent Advances and New Challenges.* Frontiers in immunology, 2013. **4**: p. 242.
60. Messaed, C., et al., *NLRP7 in the spectrum of reproductive wastage: rare non-synonymous variants confer genetic susceptibility to recurrent reproductive wastage.* Journal of medical genetics, 2011. **48**(8): p. 540-8.
61. Torra-Massana, M., et al., *Novel phospholipase C zeta 1 mutations associated with fertilization failures after ICSI.* Human reproduction (Oxford, England), 2019. **34**(8): p. 1494-1504.
62. Ferrer-Vaquer, A., et al., *PLCζ sequence, protein levels, and distribution in human sperm do not correlate with semen characteristics and fertilization rates after ICSI.* Journal of assisted reproduction and genetics, 2016. **33**(6): p. 747-56.
63. Yoon, S.Y., et al., *Human sperm devoid of PLC, zeta 1 fail to induce Ca(2+) release and are unable to initiate the first step of embryo development.* The Journal of clinical investigation, 2008. **118**(11): p. 3671-81.
64. Oud, M.S., et al., *A systematic review and standardized clinical validity assessment of male infertility genes.* Human reproduction (Oxford, England), 2019. **34**(5): p. 932-941.
65. Fakhro, K.A., et al., *Point-of-care whole-exome sequencing of idiopathic male infertility.* Genetics in medicine : official journal of the American College of Medical Genetics, 2018. **20**(11): p. 1365-1373.
66. Jelin, A.C. and N. Vora, *Whole Exome Sequencing: Applications in Prenatal Genetics.* Obstetrics and gynecology clinics of North America, 2018. **45**(1): p. 69-81.
67. Best, S., et al., *Promises, pitfalls and practicalities of prenatal whole exome sequencing.* Prenatal diagnosis, 2018. **38**(1): p. 10-19.
68. Kashir, J., et al., *Variance in total levels of phospholipase C zeta (PLC-ζ) in human sperm may limit the applicability of quantitative immunofluorescent analysis as a diagnostic indicator of oocyte activation capability.* Fertility and sterility, 2013. **99**(1): p. 107-117.e3.
69. Kashir, J., et al., *Antigen unmasking enhances visualization efficacy of the oocyte activation factor, phospholipase C zeta, in mammalian sperm.* Molecular human reproduction, 2017. **23**(1): p. 54-67.
70. Saleh, A., et al., *Essential Role of Sperm-Specific PLC-Zeta in Egg Activation and Male Factor Infertility: An Update.* Frontiers in cell and developmental biology, 2020. **8**: p. 28.
71. Chen, B., et al., *Novel mutations and structural deletions in TUBB8: expanding mutational and phenotypic spectrum of patients with arrest in oocyte maturation, fertilization or early embryonic development.* Human reproduction (Oxford, England), 2017. **32**(2): p. 457-464.
72. Feng, R., et al., *Mutations in TUBB8 and Human Oocyte Meiotic Arrest.* The New England journal of medicine, 2016. **374**(3): p. 223-32.
73. Garattini, S. and G. Grignaschi, *Animal testing is still the best way to find new treatments for patients.* European journal of internal medicine, 2017. **39**: p. 32-35.
74. Yurttas, P., et al., *Role for PADI6 and the cytoplasmic lattices in ribosomal storage in oocytes and translational control in the early mouse embryo.* Development (Cambridge, England), 2008. **135**(15): p. 2627-36.
75. Esposito, G., et al., *Peptidylarginine deiminase (PAD) 6 is essential for oocyte cytoskeletal sheet formation and female fertility.* Molecular and cellular endocrinology, 2007. **273**(1-2): p. 25-31.
76. Li, L., B. Baibakov, and J. Dean, *A subcortical maternal complex essential for preimplantation mouse embryogenesis.* Developmental cell, 2008. **15**(3): p. 416-25.

77. Tong, Z.B., et al., *Mater, a maternal effect gene required for early embryonic development in mice.* Nature genetics, 2000. **26**(3): p. 267-8.
78. Zheng, P. and J. Dean, *Role of Filia, a maternal effect gene, in maintaining euploidy during cleavage-stage mouse embryogenesis.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(18): p. 7473-8.
79. Nozawa, K., et al., *Sperm-borne phospholipase C zeta-1 ensures monospermic fertilization in mice.* Scientific reports, 2018. **8**(1): p. 1315.
80. Hachem, A., et al., *PLCζ is the physiological trigger of the Ca(2+) oscillations that induce embryogenesis in mammals but conception can occur in its absence.* Development (Cambridge, England), 2017. **144**(16): p. 2914-2924.
81. Inhorn, M.C., et al., *Consanguinity and family clustering of male factor infertility in Lebanon.* Fertility and sterility, 2009. **91**(4): p. 1104-9.
82. Lawrenz, B., et al., *Ethnical and sociocultural differences causing infertility are poorly understood-insights from the Arabian perspective.* Journal of assisted reproduction and genetics, 2019. **36**(4): p. 661-665.
83. Craven, L., et al., *Novel reproductive technologies to prevent mitochondrial disease.* Human reproduction update, 2017. **23**(5): p. 501-519.
84. Tang, M., et al., *Germline nuclear transfer in mice may rescue poor embryo development associated with advanced maternal age and early embryo arrest.* Human reproduction (Oxford, England), 2020. **35**(7): p. 1562-1577.
85. Zhang, J., et al., *Pregnancy derived from human zygote pronuclear transfer in a patient who had arrested embryos after IVF.* Reproductive biomedicine online, 2016. **33**(4): p. 529-533.
86. Sang, Q., et al., *Homozygous Mutations in WEE2 Cause Fertilization Failure and Female Infertility.* American journal of human genetics, 2018. **102**(4): p. 649-657.

8. Addendum

Additional file 1: Interesting variants reported in Bonte 1.0 Study

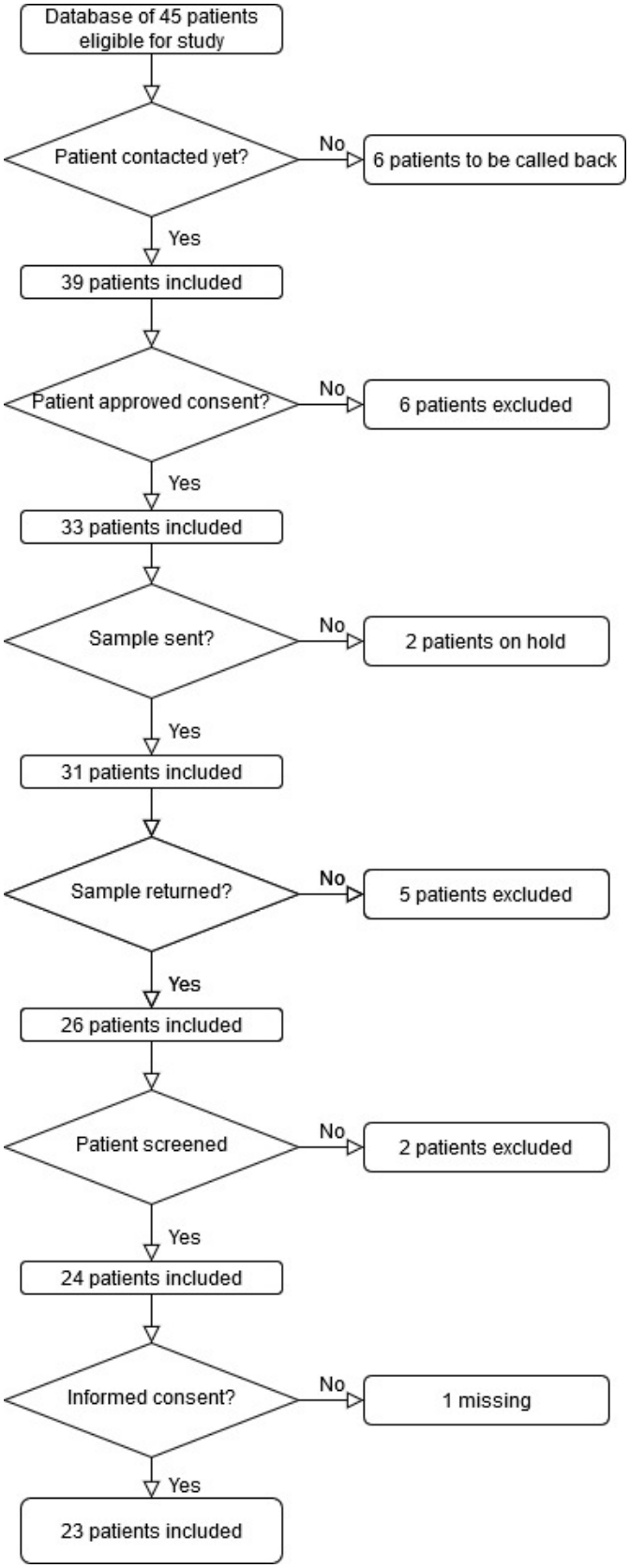
Code	Gene	g.Nomenclature	c.Nomenclature	p.Nomenclature	HET/cHET/HOM	RefSNP	Classification criteria	Classifiaction	MAF (GnomAD - all)	MAF (GnomAD - nfe)	Polyphen	SIFT	Grantham	Sanger confirmed
1A	*NLRP5*	Chr19(GRCh37):g.56531732_56531734dup	c.623-12_623-11insTTC	p.?	HET	rs538671525	BP4	Uncertain signficiance	0,006874	0,005891	NA	NA	NA	Yes
3A	*NLRP2*	Chr19(GRCh37):g.55494638T>C	c.1572T>C	p.Asp524=	HET	rs61735084	BS1, BP7, PS4	Likely benign **	0,006144	0,007444	NA	NA	0	Yes
4A	*TLE6*	Chr19(GRCh37):g.2989639C>T	c.1100C>T	p.Pro367Leu	HET	rs140669738	PP3	Uncertain signficiance	0,000113	0,000109	Probably damaging	Damaging	98	Yes
6A	*TLE6*	Chr19(GRCh37):g.2989123A>G	c.805A>G	p.Lys269Glu	HET	rs147398793	PM2, BP4	Uncertain signficiance	0,000071	0,000147	Benign	Tolerated	56	Yes
12A	*NLRP2*	Chr19(GRCh37):g.55494638T>C	c.1572T>C	p.Asp524=	HET	rs61735084	BS1, BP7, PS4	Likely benign **	0,006144	0,007444	NA	NA	0	Yes
13A	*NLRP2*	Chr19(GRCh37):g.55508725C>G	c.2920C>G	p.Leu974Val	HET	rs61735074	(BA1)*	Benign/Uncertain signficiance *	0,004792	0,000116	Probably damaging	Damaging	32	Yes
16A	*NLRP2*	Chr19(GRCh37):g.55494283C>G	c.1217C>G	p.Thr406Arg	HET	rs139903547	BS1, PP3	Uncertain signficiance	0,006266	0,009185	Probably damaging	Damaging	71	Yes
17A	*NLRP2*	Chr19(GRCh37):g.55494638T>C	c.1572T>C	p.Asp524=	HET	rs61735084	BS1, BP7, PS4	Likely benign **	0,006144	0,007444	NA	NA	0	Yes
18A	*NLRP2*	Chr19(GRCh37):g.55494668C>T	c.1602C>T	p.Asp534=	HET	rs144234287	BP7	Uncertain signficiance	0,000244	0,000472	NA	NA	0	Yes
18A	PADI6	Chr1(GRCh37):g.17723562T>C	c.1619-5T>C	p.?	HET	novel	BP4, PM2	Uncertain signficiance	not found	not found	NA	NA	NA	Yes

* The allele frequency in the African population is 0.05019, but is very low in all other populations. Therefore either benign, or VUS.

** The frequency of this variant in our patient population is higher than in the control population, which might point to it being VUS/(likely) pathogenic instead of likely benign

HET: heterozygous
cHET: compound heterozygous
HOM: homozygous
VUS: variant of uncertain significance
NA: not available
nfe: non-Finish European

Additional file 2: Patient recruitment



Additional file 3: VarSome criteria – The chart organizes all criteria by type of evidence as well as strength of the criteria for both benign as pathogenic assertion [56]

	Benign		Pathogenic			
	Strong	**Supporting**	**Supporting**	**Moderate**	**Strong**	**Very Strong**
Population Data	MAF is too high for disorder *BA1/BS1* **OR** observation in controls inconsistent with disease penetrance *BS2*			Absent in population databases *PM2*	Prevalence in affecteds statistically increased over controls *PS4*	
Computational And Predictive Data		Multiple lines of computational evidence suggest no impact on gene /gene product *BP4* Missense in gene where only truncating cause disease *BP1* Silent variant with non predicted splice impact *BP7*	Multiple lines of computational evidence support a deleterious effect on the gene /gene product *PP3*	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before *PM5* Protein length changing variant *PM4*	Same amino acid change as an established pathogenic variant *PS1*	Predicted null variant in a gene where LOF is a known mechanism of disease *PVS1*
Functional Data	Well-established functional studies show no deleterious effect *BS3*		Missense in gene with low rate of benign missense variants and path. missenses common *PP2*	Mutational hot spot or well-studied functional domain without benign variation *PM1*	Well-established functional studies show a deleterious effect *PS3*	
Segregation Data	Non-segregation with disease *BS4*		Co-segregation with disease in multiple affected family members *PP1*	Increased segregation data →		
De novo Data				*De novo* (without paternity & maternity confirmed) *PM6*	*De novo* (paternity & maternity confirmed) *PS2*	
Allelic Data		Observed in *trans* with a dominant variant *BP2* Observed in *cis* with a pathogenic variant *BP2*		For recessive disorders, detected in *trans* with a pathogenic variant *PM3*		
Other Database		Reputable source w/out shared data = benign *BP6*	Reputable source = pathogenic *PP5*			
Other Data		Found in case with an alternate cause *BP5*	Patient's phenotype or FH highly specific for gene *PP4*			

BS: benign strong
BP: benign supporting
FH: family history
LOF: loss-of-function
MAF: minor allele frequency
path.: pathogenic
PM: pathogenic moderate
PP: pathogenic supporting
PS: pathogenic strong
PVS: pathogenic very strong

Additional file 4: VarSome criteria - Rules for combining criteria to classify sequence variants [56]

Pathogenic

1. 1 Very Strong (PVS1) *AND*
   - **a.** ≥1 Strong (PS1–PS4) *OR*
   - **b.** ≥2 Moderate (PM1–PM6) *OR*
   - **c.** 1 Moderate (PM1–PM6) and 1 Supporting (PP1–PP5) *OR*
   - **d.** ≥2 Supporting (PP1–PP5)
2. ≥2 Strong (PS1–PS4) *OR*
3. 1 Strong (PS1–PS4) *AND*
   - **a.** ≥3 Moderate (PM1–PM6) *OR*
   - **b.** 2 Moderate (PM1–PM6) *AND* ≥2 Supporting (PP1–PP5) *OR*
   - **c.** 1 Moderate (PM1–PM6) *AND* ≥4 Supporting (PP1–PP5)

Likely Pathogenic

1. 1 Very Strong (PVS1) *AND* 1 Moderate (PM1–PM6) *OR*
2. 1 Strong (PS1–PS4) *AND* 1–2 Moderate (PM1–PM6) *OR*
3. 1 Strong (PS1–PS4) *AND* ≥2 Supporting (PP1–PP5) *OR*
4. ≥3 Moderate (PM1–PM6) *OR*
5. 2 Moderate (PM1–PM6) *AND* ≥2 Supporting (PP1–PP5) *OR*
6. 1 Moderate (PM1–PM6) *AND* ≥4 Supporting (PP1–PP5)

Benign

1. 1 Stand-Alone (BA1) *OR*
2. ≥2 Strong (BS1–BS4)

Likely Benign

1. 1 Strong (BS1–BS4) and 1 Supporting (BP1–BP7) *OR*
2. ≥2 Supporting (BP1–BP7)

---

* Variants should be classified as Uncertain Significance if other criteria are unmet or the criteria for benign and pathogenic are contradictory.

Additional file 5: Variants of interest, detected in all patients for variants in PADI6, NLRP2, NLRP5, NLRP7, TLE6, KHDC3L and PLCZ1.

Code	Gene	g.Nomenclature	c.Nomenclature	p.Nomenclature	HET/cHET/HOM	RefSNP	Classification criteria	Classifiaction	MAF (GnomAD - all)	MAF (GnomAD - nfe)	Polyphen	SIFT	Grantham	Sanger confirmed
101A	NLRP7	Chr19(GRCh37):g.55451257C>A	c.930G>T	p.Gln310His	HET	rs145973556	PM1, PM2, BP4	Uncertain significance	0,0017	0,003	NA	Deleterious	24	
102A	NLRP2	Chr19(GRCh37):g.55501424G>A	c.2401G>A	p.Ala801Thr	HET	rs117066658	BS1, BP1, BP4	Likely benign	0,0096	0,014	Benign	Tolerated	58	
107A	NLRP5	Chr19(GRCh37):g.56531732_56531734dup	c.623-12_623-11insTTC	p.?	HET		BP4	Uncertain significance		0,006	NA	NA	0	No, 201132, 201139, 201140 (failed)
108A	NLRP2	Chr19(GRCh37):g.55494141T>C	c.1075T>C	p.Phe359Leu	HET	rs62124644	BP1, BP4	Likely benign	0,007	0,011	Benign	Tolerated	22	
111A	NLRP7	Chr19(GRCh37):g.55451720C>T	c.467G>A	p.Arg156Gln	HET	rs61746625	BM2, BP1, BP4	Benign	0,0073	0,012	Benign	Tolerated	43	
112A	NLRP7	Chr19(GRCh37):g.55451083A>G	c.1104T>C	p.Ile368=	HET	rs16546 36	BA1, BP4, BP6, BP7	Benign	0,01	0,007	NA	NA	0	
113A	NLRP2	Chr19(GRCh37):g.55494126A>G	c.1060A>G	p.Ile354Val	HET		BP4, BP6	Likely benign	0,005026	0,007894	Benign	Tolerated	29	Yes (706), also found in Bonte 1.0
118B	PLCZ1	Chr12(GRCh38):g.18841115C>T	c.1499C>T	p.Ser500Leu	HOM	rs10505830	BA1, BP4, PM1, PP2	Benign	0,0312	0,0462	Possibly damaging	Deleterious, tolerated	145	
119A	TLE6	Chr19(GRCh37):g.2989211C>T	c.893C>T	p.Thr298Met	HET	rs141406529	PM2, BP4	Uncertain significance	0,00004	0,000027	Possibly damaging	Tolerated	81	Yes, 201132, 201139, 201140
119A	NLRP7	Chr19(GRCh37):g.55450930G>C	c.1257G>C	p.Ala419=	HET	rs151120858	PM2, BP4, BP7	Likely benign	0,000958	0,000771	NA	NA	0	
121A	NLRP7	Chr19(GRCh38):g.54938079G>A	c.2094G>A	p.His698=	HET	rs104895524	BS2, BP4, BP7	Likely benign	0,002397	0,002678	NA	NA	0	
121A	PADI6	Chr1(GRCh38):g.17373225G>A	c.286G>A	p.Ala96Thr	HET	rs181841908	PM2, BP4	Uncertain significance	0,001099	0,001857	Benign	Tolerated	58	Yes, 201132, 201139, 201140
122A	NLRP2	Chr19(GRCh38):g.54982758A>G	c.1060A>G	p.Ile354Val	HET		BS1, BP4, BP6	Likely benign	0,005026	0,007894	Benign	Tolerated	29	
123A	NLRP2	Chr19(GRCh38):g.54983044C>T	c.1346C>T	p.Thr449Met	HET	rs144271525	BP4	Uncertain significance	0,000182	0,000356	Benign	Tolerated	81	Yes, 201132, 201139, 201140

HET: heterozygous
cHET: compound heterozygous
HOM: homozygous
nfe: non-Finish European
NA: not available