UNIVERSITÀ DEGLI STUDI DI MILANO

DIP. DI BIOTECNOLOGIE MEDICHE E MEDICINA TRASLAZIONALE

DOTTORATO DI RICERCA IN
BIOTECNOLOGIE APPLICATE ALLE SCIENZE MEDICHE

XXVIII Ciclo

Settore Scientifico – Disciplinare MED/03

GENOME WIDE ANALYSIS IN A COHORT OF 46,XX PATIENTS AFFECTED BY AN EXTREME PHENOTYPE OF PRIMARY OVARIAN INSUFFICIENCY: AN EFFICIENT TOOL TO IDENTIFY NEW GENES INVOLVED IN OOCYTE MATURATION AND DIFFERENTIATION

Tutore: Prof.ssa Anna MAROZZI
Co-tutore: Prof.ssa Palma FINELLI

Tesi di Dottorato di Ricerca di:
Dott.ssa Ilaria BESTETTI
Matricola R10076

Anno Accademico 2014-2015

INDEX

ABSTRACT

Primary Ovarian Insufficiency (POI) is a heterogeneous group of disorders with an incidence of 1:10,000 women by age 20, 1:1,000 by age 30, 1:100 by age 40. POI describes the progression toward the cessation of ovarian function and can occur in the most serious form as primary amenorrhea (PA), with absent pubertal development and/or ovarian dysgenesis (OD), or in milder phenotype with post-pubertal onset and secondary amenorrhea (SA). Several are the etiological causes that may induce ovarian dysfunction, among which the genetic component is considered prevalent (as supported by the occurrence of families with more than one affected women and the existence of several idiopathic cases) but highly heterogeneous.

46,XX non-syndromic women showing the most severe phenotype, characterized by the absence of pubertal development with PA and OD, are very rare but the search for genetic variations in this extreme phenotype may be more effective in identifying novel pathogenic mechanisms. Hence, using high resolution array-CGH we searched for rare high-penetrant CNVs involving genes essential for ovarian function in a cohort of 67 46,XX non-syndromic patients affected by PA, namely 53 sporadic (79.1%) and 14 familial (20.9%) cases. 28 out of 67 women resulted positive to array-CGH analysis because having at least one rare "ovarian" CNV: a total number of 32 CNVs involving 37 ovarian genes was selected.

Population from Database of Genomic Variants (DGV) was used to evaluate the rarity of POI CNVs, but it does not match to the ideal set of controls for POI disease (neither age nor gender of DGV controls are known). Thus, to better understand the CNVs contribution in disease onset, the rare "ovarian" CNVs found in patients according to DGV were searched in an *ad hoc* control cohort, previously screened by array-CGH, consisting in 140 healthy women with normal reproductive life and physiological menopause. 28 out of 32 rare "ovarian" CNVs detected in patients were not found in the control group thus supporting their role in the POI's pathogenesis. Moreover, to evaluate the presence of a significant enrichment in ovarian genes in the POI group, array-CGH of the *ad hoc* control cohort were analyzed with the same approach adopted for patients cohort and 49 CNVs involving 54 ovarian genes were selected. Several statistical analyses were performed comparing patients' to controls' data and revealed no significant differences. Nevertheless, the CNVs found in the POI cohort containing ovarian genes are more harmful respect to the CNVs identified in the controls cohort.

The 37 genes perturbed or possibly perturbed by POI CNVs are implicated in several ovarian processes (e.g., regulation of cytoskeleton dynamics for oocytes asymmetric division, maintenance of oocytes genomic integrity, ovarian differentiation, follicular development, and meiotic resumption), thus supporting their involvement in POI etiology. Validation and characterization of selected CNVs, as well as the study of a possible gene perturbation at mRNA

level, was also crucial in order to perform a correct genotype-phenotype correlation and to propose new candidate genes for POI disease (e.g. *TP63*, *VLDLR*).

39 out of 67 women resulted negative for rare "ovarian" CNVs (58.2%) suggesting to combine different genomic approaches in order to increase the detection rate of the disorder. Hence, 17 out of 67 collected patients, were submitted to a preliminary WES analysis searching for rare SNVs in a total of 191 genes selected from array-CGH data, and literature regarding POI and ovary. The WES preliminary analysis allowed to confirm the importance of some array CGH new candidate genes in POI onset (e.g. *VLDLR*) and the complex heterogeneity of POI.

The combination of these molecular evidences, with major or minor contribution, might have been at the basis of POI supporting the existence of a disease genetic model characterized by oligogenic heterozygosity (i.e., the simultaneous presence in a single patient of multiple heterozygous quantitative variants/rare mutations, both de novo and/or inherited, affecting multiple genes).

The present approach using both array-CGH and WES techniques, resulted an efficient tool to identify rare variants (CNVs and SNVs) involving both genes already reported in POI, and new candidate genes with a role in oocyte maturation and differentiation. The results of this study are promising to expand the knowledge about the molecular pathways involved in POI pathogenesis and probably provide the basis for a more accurate genetic diagnosis of POI patients.

1. INTRODUCTION

1.1 Physiological aspects of ovarian growth and differentiation

The main function of the female gonad is the development of mature oocytes for future fertilization. Moreover it produces ovarian hormones allowing the development of female secondary sexual features and supporting pregnancy [1].
Oogenesis and folliculogenesis are the two biological processes that occur to ensure oocyte growth and differentiation.

1.1.1 Oogenesis

Oogenesis consists in the formation of haploid gametes from primordial germ cells (oogonia) through meiosis [2], which starts in the first three months during fetal life and ends only when a mature oocyte is fertilized [3]. The number of oogonia is fixed at birth and it is thought that no additional primary oocytes are created during life, while male spermatogenesisis a continuous process [4,5]. Female reproductive process differs from spermatogenesis in several features: it takes place in a rhythm called menstrual cycle and for every primordial germ cell provides just one functional gamete. The others cells originated from meiosis cellular divisions are little polar bodies that die fast [3,4].
In detail, during human embryogenesis, primordial germ cells migrate from the epithelium of the yolk sack to colonize the gonadal ridge. When primordial germ cells reach the genital ridge they lose their motility and in females become oogonia [6,7]. Germ cells start to proliferate by mitosis and develop in cluster of cells called germ line cysts or oocyte nests, which are connected by intercellular bridges [2]. At 11–12 weeks of gestation some of the oogonia end mitosis, enter the first stages of meiosis becoming **primary oocytes** stalling in profase I (diplotene) [8]. They are stopped in the first meiosis division till preovulatory stage that occurs shortly before pubertal age [4]. At 20^{th} week of gestational age, the number of primary oocytes arrested in meiosis achieve their maximum development, when approximately 7,000,000 have been created **(Fig.1)** [1,5]. Then, the rate of oogonial mitotic division declines [1]. Besides, the majority of primary oocytes die before birth and in childhood, through a physiological programmed cell death mechanism known as “atresia”. Conversely, the oocytes survived remain quiescent till pubertal age [9].
A female has about 400,000 primary oocytes at the beginning of pubertal age **(Fig.1)**, which constitutes the ovarian reserve (OR) for her entire life. Once reached puberty, menstrual cycle starts and about 20-25 primary oocytes restart growth and differentiation processes monthly to complete meiosis I division and produce a secondary haploid oocyte with a second polar body. Secondary oocyte arrests the cycle at metaphase II of meiotic division waiting for ovulation and

meiosis II process will end only after the fertilization of the mature oocyte, unless it will die by apoptosis and will never end meiosis [3].

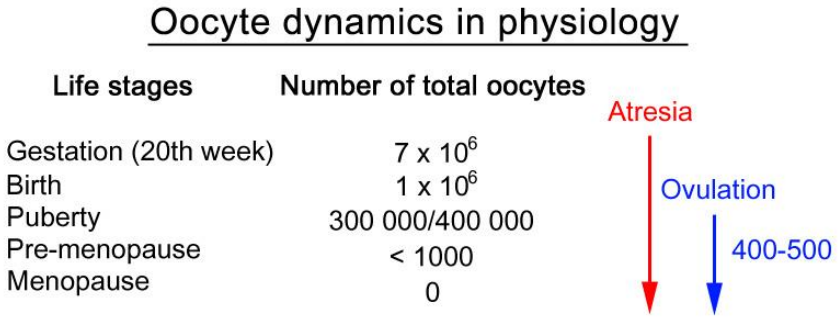

Fig.1. Oocyte dynamics in physiology. *Stages of female reproductive life are shown with the respective number of oocyte reserve. Atresia occurs from the 20th week of gestation, whereas ovulation starts from pubertal age; about 400-500 oocytes become dominant ovulating follicles. This image was adapted from Persani et al, 2010 [10].*

1.1.2 Folliculogenesis

Folliculogenesis is a physiological process that runs at the same time of oogenesis and it is regulated both by endocrine as well as paracrine factors [11,12,13].

It consists in the maturation of the **ovarian follicle**, formed by an oocyte wrapped by tightly packed layers of somatic cells. The main role of the ovarian follicle is oocyte support. Follicular development starts during fetal life in the internal part of the ovary immediately after the primary oocytes reach diplotene [2] and consists in the growth of a number of small primordial follicles into primary, secondary and then large preovulatory follicles that enter the menstrual cycle (**Fig.2**). Follicular maturation is a very long process, where the primary oocytes are incorporated in primordial follicles before birth and remain quiescent in the ovary, until recruitment into the growing pool every month during reproductive life [8,14,15].

As well as oocytes, only a few follicles survive to complete the differentiation process and ovulation, while immature follicles (more than 99% of follicles) will undergo atresia [1,16,17]. In the mammalian ovary, follicular development and atresia are closely regulated and seem to rely on a balance between cell death (inhibin, Bax, FasL, TNF-α, caspase, etc.) and survival-promoting factors (FSH, insulin-like growth factor-I, interleukin-1β, epidermal growth factor, Bcl-2, etc.) [9]. Although atresia can occur any time throughout follicular development, most follicles become atretic during the early antral stage of folliculogenesis after exposure of the granulosa cells to gonadotropin. At this time granulosa cells start to differentiate thus being more susceptible to apoptosis and becoming early and progressed atretic follicles [12,18,19].

The follicular development can be divided into two stages (**Fig.2**) based on the subsequently follicular maturative phases and on the involvement of ovarian, hypothalamic and pituitary hormones:

- ***gonadotropin independent or preantral phase*** (primordial, primary, secondary follicles included);
- ***gonadotropin dependent or antral phase*** (tertiary, graafian, preovulatory follicles included) [20,21].

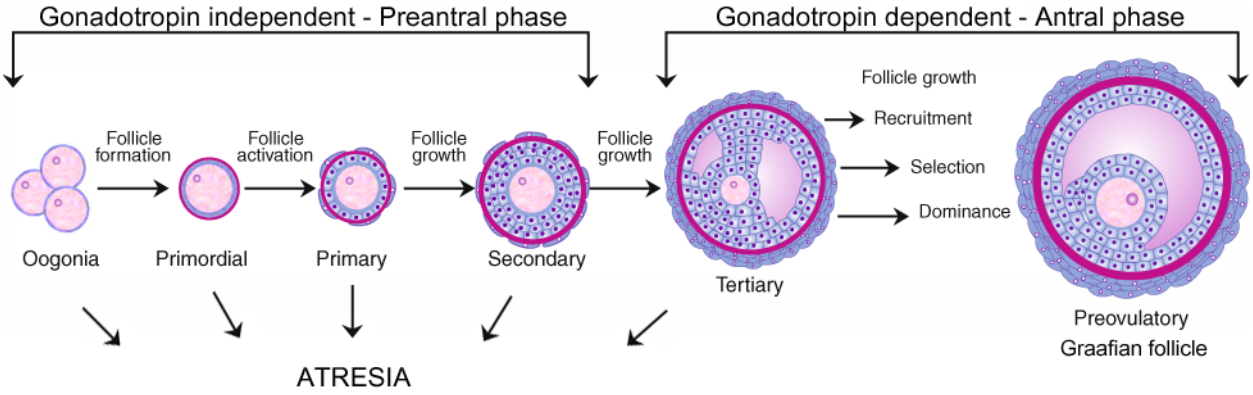

Fig.2. Schematic representation of follicular development. *Preantral and antral phases of folliculogenesis are depicted. Atresia occurs during all the growing and differentiation process. Oogonia develop from primordial germ cells and differentiate into mature oocytes. All the preantral follicles have a primary oocyte; the tertiary follicles are characterized by an antral cavity which contains follicular fluid. Graafian preovulatory follicles have more antral fluid and contain a secondary oocyte. This image was adapted from Araujo, Gastal et al.[22].*

1.1.2.1 Gonadotropin independent stage: from primordial to pre-antral follicles

Primordial follicles formation begins around midgestation and continues until just after birth with the recruitment of pregranulosa cells to each oogonia nest. In particular, the oocyte nests break apart into individual cells, lose their intercellular bridges and become enclosed in a layer of flattened pregranulosa cells, becoming the primordial follicles [2,8,12,23]. Oocytes not surrounded by granulosa cells enter apoptosis [24,25,26].

As aforementioned when mitosis stops, the oogonia enter the early stages of meiosis I as primary oocytes. These immature oocytes within the primordial follicles stall in meiosis I [8,27], and remain quiescent for months or years showing little to no biological activity [28]. During maturation, some of the primordial follicles leave the resting pool and are stimulated to grow and develop into primary follicles. The process by which primordial cells "wake up" is known as **initial recruitment** [1,8] and it is mediated by secretion and communication of various stimulatory, inhibitory hormones and growth factors between oocyte and granulosa cells

[29,30]. During primordial follicle development, the primary oocytes grow as well, but remain arrested in the prophase of meiosis I. For those follicles not recruited, the default pathway is to stay dormant [1].

Subsequently, the granulosa cells of primordial follicle start to change from a flat to a cuboidal structure and both the oocyte and the follicle grow significantly, increasing up to 0.1 mm in diameter: **primary follicles** are forming. During this development phase the oocyte genome is activated and several genes are transcribed [31] among which those encoding proteins ZP-1, ZP-2, ZP-3 are relevant for the extracellular matrix production called zona pellucida [32]. Granulosa cells send out some extensions through zona pellucida to form gap junctions with the oocyte membrane. To note, connexin 37 (Cx37) is a component of the ovarian gap junction with a critical role in folliculogenesis, ovulation and fertility [33]. Indeed, the communication between oocyte and granulosa cells is crucial to allow the passage of ions, metabolites, signaling factors [34,35] and it is responsible for the positive or negative regulation of all the follicular growth process. Moreover, primary follicles express receptors for follicle stimulating hormone (FSH), but they remain gonadotropin-independent until the antral stage [28,36]. Vascularization is not present in primary follicles [20].

The most important change that occurs in the development of a **secondary follicle** is the formation of a theca layer. This tissue, made of stroma cells, further differentiates in inner theca interna and outer theca externa [20]. Theca interna is characterized by a vascularization system and is able to produce steroid hormones through endocrine activity: i.e. androstenedione (intermediate of both estrogens and androgens), testosterone, progesterone. The outer layer (theca externa) is composed by smooth muscular cells innervated by autonomic nerves [37].
At this stage more than one layer of granulosa cells is present due to mitotic divisions. These cells are able to produce estrogens, converting steroid hormones produced by theca cells. Moreover at this time, each granulosa cell bears FSH receptors whereas theca cells start to express LH receptors [38].

The hallmark of a **tertiary pre-antral follicle** is the formation of an antrum through a process called cavitation. It begins to form at one pole of the oocyte with the accumulation between granulosa cells of a follicular fluid, which is a medium in which regulatory molecules are found. The ending of cavitation, corresponds to the formation of a Graafian dominant follicle which is mature for ovulation. Antral formation starts in a gonadotropin independent manner, but it is controlled by autocrine paracrine mechanism through growth factors as activin and kit ligand [20,39,40]. Meanwhile the follicle reaches the pre-antral stage, the oocyte has been enlarging considerably and has become competent enough to resume the meiotic process [41].

1.1.2.2 Gonadotropin dependent stage and menstrual cycle

A **Graafian follicle** can be defined as a family composed by heterogeneous large follicles characterized by a cavity (or antrum) containing a follicular fluid. The antral follicles maintain the same structural complexity even though dramatic changes occur in their size.
After pubertal onset, a small number of the large antral follicles can be rescued to continue the growth in a gonadotropins-dependent manner through a process called **cyclic recruitment** that overlaps with the follicular phase of the menstrual cycle [1,4]; indeed, it takes only 14 days for an antral follicle to become a dominant Graafian follicle [1,21].
At this stage, as a result of the increase in circulating FSH, only a limited number of follicles survive, and the fate of the others is to undergo atresia. Among the recruited follicles, one is considered as dominant and secrets high levels of estrogens and inhibins to suppress pituitary FSH release as a negative selection of the remaining follicles. At the same time, an increase in growth factors secretion allows a positive selection for the dominant follicle, thus ensuring its final growth and ovulation [1].

Once reached puberty, together with follicular development as explained above, a primary oocyte completes every month the first meiotic division and starts the second one arresting in metaphase II and becoming secondary oocyte. This cyclic process is called **menstrual cycle** and it is composed by:

- ovarian cycle (ovarian functional and structural changes)
- uterine cycle (uterine functional and structural changes)
- cycle changes in the secretion of ovarian, hypothalamic and pituitary hormones [4].

The **ovarian cycle** can be divided in two phases: follicular stage and luteal stage, each lasting about 14 days. The former starts at the beginning of menstruation and ends with ovulation, the latter matches with the other part of the cycle [4]. In particular, as mentioned before, in the follicular phase a pool of antral follicles is recruited and continues to grow; the dominant Graafian follicle that receives the major FSH surge ovulates in the 14^{th} day of menstrual cycle to release the mature oocyte for fertilization, whereas the remaining theca and granulosa cells become the corpus luteum [1]. This structure is named from a yellow lipid that accumulates in theca layers and it is important to prepare the uterus for a possible pregnancy ensuring a proper hormonal supply. Contrary it atrophies and becomes a scar tissue named corpus albicans [3].

Hormonal changes during menstrual cycle

Several hormones participate in a complex process of positive and negative feedback to regulate folliculogenesis (**Fig.3**):

- gonadotropin releasing hormone (GnRH) secreted by the hypothalamus;
- gonadotropins: follicle-stimulating hormone (FSH); luteinizing hormone (LH);
- estrogen and progesterone.

Early follicular stage is characterized by low levels of estrogen and progesterone in the plasma and represents the final stage of luteal phase as a consequence of luteum body degeneration. Since these two hormones tend to inhibit gonadotropin secretion from pituitary, plasmatic levels of FSH and LH start to increase (**Fig.3**). Furthermore, GnRH stimulates the release of FSH and LH from the anterior pituitary gland that will have a stimulatory effect on follicle growth during antral stage of folliculogenesis. In particular, FSH stimulates follicles growth and differentiation through binding FSH receptors on granulosa cells, whereas the theca cells of the growing oocytes show LH receptors. Thus, LH binding triggers the secretion of androgens that are converted via aromatization into estrogens by granulosa cells [42]. The estrogen surge and inhibin produced by granulosa cells of the growing follicles act in a negative feedback pathway on hypothalamus, inhibiting LH and FSH release. While LH levels tends to be stable, as a feedback mechanism, FSH levels decrease due to inhibin secretion. Besides, the reduction of FSH level is partially responsible of the atretic process of non dominant follicles [3,9].

In the late stage of follicular phase, as more estrogen is secreted, more LH receptors are expressed by the theca cells, triggering theca cells to produce more androgen that will become estrogen. This positive feedback loop causes LH to increase sharply. A peak level of LH is critical to ensure ovulation because causes nuclear maturation with the end of meiosis I. Moreover estrogen secretion decreases and granulosa cells start to produce progesterone and some enzyme able to disrupt the follicular wall for ovulation [8,43] (**Fig.3**).

After ovulation, the dominant follicle with no oocyte becomes a corpus luteum, which produces a huge amount of progesterone that inhibits GnRH, FSH and LH secretion and stimulates the uterus to prepare for a possible pregnancy. When it reaches the maturity starts to secrete also a little amount of estrogens (**Fig.3**). After about 10 days from ovulation, without the fertilization of the mature oocyte, the corpus luteum degenerates causing a drastic decrease of estrogen and progesterone that leads to menstruation [3].

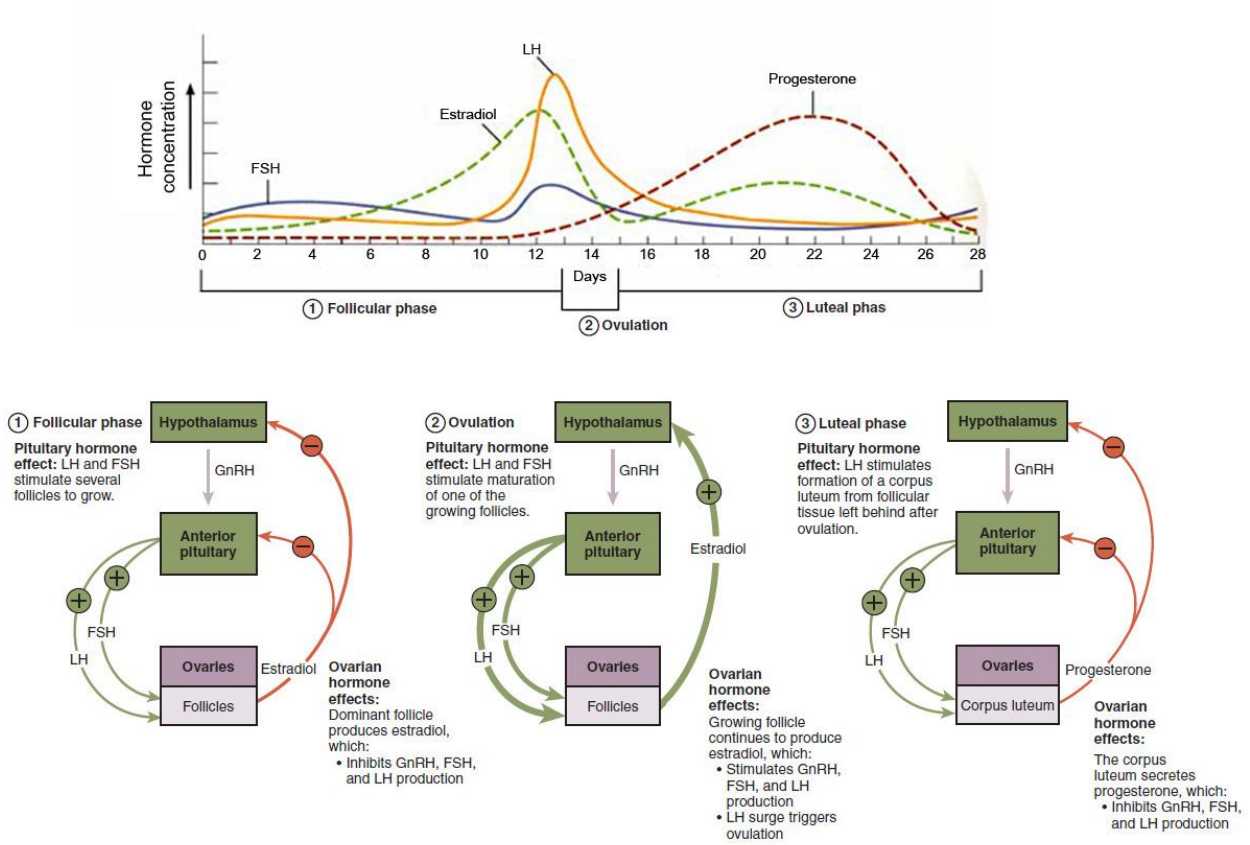

***Fig.3**. **Hormonal changes during menstrual cycle.** Above, changes in concentration of anterior pituitary (blue and yellow continuos lines) and ovarian hormones (green and red dashed lines) are shown during the three follicular phases. Below, hormones regulatory feedback loops acting during follicular phase (1), ovulation (2), and luteal phase (3).*

1.2 Primary Ovarian Insufficiency (POI)

1.2.1 Prevalence and clinical overview

The age of menopause in a woman is determined by both genetic and environmental factors. The average age at which this process occurs in Caucasian women is 50.7 years [44,45]. However, about 1% of women before 40 years experience a premature cessation of the ovarian reserve and suffer from a fertility defect named Primary Ovarian Insufficiency (POI) [46]. This condition, firstly described by Fuller Albright in 1942 [47,48], has an incidence of 1 in 10,000 women by age 20, 1 in 1,000 by age 30, 1 in 100 by age 40 and 1 in 10 by age 45 [10]. The prevalence of POI, however, varies by ethnicity, with women of oriental origin having a lower risk and African Americans a higher risk compared to Caucasian Americans [44].

According to the onset age, the disorder can manifest i) in the most serious form as primary amenorrhea (PA), with absent pubertal development and/or ovarian dysgenesis (a condition in which the ovaries are completely free of follicles), or ii) in milder phenotype with post-pubertal onset characterized by secondary amenorrhea (SA) and defect in folliculogenesis [46,49,50]. Based on evidence that the disorder has a long and heterogeneous clinical course, the term Primary Ovarian Insufficiency has been recently proposed to better describe the progression toward the cessation of ovarian function [51,52]. Indeed, POI is commonly adopted to indicate diagnostically similar conditions as hypergonadotropic hypogonadism, premature ovarian failure (POF) and ovarian dysgenesis [10,49].

Biochemically, the disorder is characterized by elevated serum gonadotropin levels as in the menopausal range (usually FSH over 40 IU/L, obtained at least 1 month apart) and serum estradiol levels less than 50 pg/mL, which indicate hypoestrogenism [53,54,55]. For this reason, compared with healthy women, patients with POI are more likely to experience long-term consequences of hypoestrogenism including osteoporosis, cardiovascular diseases (myocardial infarction and stroke), neurocognitive disorders (sleep disturbance and decreased mental concentration), loss of sexual desire or vaginal dryness. However, infertility is the major consequence of POI [10,49,53,56].

1.2.2 Diagnosis and management of POI

As mentioned before, POI is characterized by low levels of gonadal hormones (estrogens and inhibins), and high levels of gonadotropins (LH and FSH) before the age of 45. The diagnosis is confirmed if the FSH value is in the menopausal range (> 40 IU/L) in at least two serial measurements after 4-6 months of amenorrhea or primary amenorrhea manifestation.

Although standard diagnostic criteria for POI have yet to be established, another indicator to be evaluated is the anti-Mullerian hormone (AMH), which can help to assess the state of follicular senescence and it is considered a putative marker of the ovarian reserve even before menarche [57]. Ultrasonography of ovaries, together with age, are traditional indicators to evaluate ovarian aging. Furthermore it is advisable to collect a comprehensive family history and to perform physical examination to exclude other causes of POI: pregnancy, polycystic ovary syndrome, lifestyle habits, hypothalamic amenhorrea, hypothalamic or pituitary lesions, hyperprolactinemia, hypothyroidism, hyperthyroidism, autoimmune disorders, fragile X syndrome or intellectual disability [46,49]. In this regard, karyotype to determine X chromosomal abnormalities and *FMR1* premutation screening are usually carried out [54,58].

Considering that unexpected diagnosis of POI affects woman's physical and emotional well-being, management is crucial for this disorder. Therapeutic strategies include hormone replacement therapy (HRT), fertility management, psychological support and annual screening to assess women's health care. Long-term HRT (estrogen and progestin) is advised to provide relief from menopausal symptoms due to hypoestrogenism, to prevent osteoporosis, caridiovascular diseases and it is usually recommended to continue this treatment till the age of natural menopause [49]. Even though some POI patients show a milder phenotype and present clinical findings similar to menopause, almost 50% have some residual ovarian function, and 5-10% are able to accomplish pregnancy [56]. For this reason, fertility preservation is important to allow future childbearing. Particularly, infertility rescue includes ovarian hyperstimulation followed by oocyte or embryo cryopreservation, and ovarian tissue preservation, but the fertility treatment with the highest success rate is still the assisted reproductive technique using donor oocytes [49,56]. Besides medical treatment, professional and family support is essential [53].

1.2.3 Etiopathogenesis

Most of primordial follicles formed during embryonic development are lost in fetal and postnatal life by atresia, and out of them only about 400-500 generally ovulate before physiological menopause. The three alterantive mechanisms that might lead to POI can be i) an initial reduction in the primordial follicle pool, ii) an accelerated atresia or iii) an altered maturation of primordial follicles [10,56]. However, the etiological causes that may induce ovarian dysfunction are highly heterogeneous and include:

- **Chromosomal abnormalities and gene mutations** (see chapter 1.3);
- **Autoimmune diseases** (e.g. hypothyroidism, adrenal insufficiency, hypoparathyroidism, type 1 diabetes mellitus, pernicious anemia, myasthenia gravis, rheumatoid arthritis,

systemic lupus erythematosus, Addison's disease, autoimmune polyglandular syndrome and steroid cell autoimmunity primary ovarian insufficiency (SCA-POI), caused by autoimmune destruction of theca cells [49,54,56,59]);

- **Metabolic disorders** (e.g. galactose 1-phosphate uridylyltransferase deficiency (galactosemia), characterized by toxic accumulation of galactose during infancy, leading to ovarian, ocular, renal, hepatic and neurological defects; 17-OH hydroxylase deficiency (CYP17) as well as aromatase deficiency (CYP19), cause defects in the pathway involved in sex steroid biosynthesis [49,56]);
- **Infections** (e.g. mumps oophoritis, characterized by mononuclear cell infiltration into the theca layer of antral follicles; HIV);
- **Iatrogenic treatments** (e.g. ovarian surgery, radiotherapy or chemotherapy for several neoplastic diseases) accounting for about 25% of all forms;
- **Environmental/lifestyle factors** (e.g. repeated doses of the ovotoxic substance 4-vinylcyclohexene diepoxide (VCD), cigarette smoking, epilepsy [10,56]).

Nevertheless, in more than 50% of POI women the main causative event remains elusive, strongly suggesting a genetic origin of the disease. These cases are classified as idiopathic or spontaneous ovarian insufficiency [10,50,51,53,55,60].

1.3 Genetic causes of POI

In the pathogenesis of idiopathic POI patients, genetic causes are considered prevalent as supported by the occurrence of families with several affected women, corresponding to an incidence from 4 to 31%, depending on the studied population [10,60,61]. Dominant or recessive forms of inheritance have been demonstrated, but the most frequent form follows an X-linked maternal inheritance pattern with incomplete penetrance, indicating that X chromosome defects are an important cause of familial POI [59,62,63,64]. Furthermore, the presence in the same family of women with heterogeneous phenotypes indicates that primary ovarian insufficiency is a genetic disorder with high variability, supporting the idea of POI as a complex multigenic disease [62]. Indeed, the existence of mouse models with absent ovarian function (knockouts in more than 100 genes [65]) and of monogenic forms of POI indicates also the relevance of autosomal genes implicated in POI pathogenesis [54]. The genes involved have several biological functions as regulation of the hypothalamic-pituitary ovarian axis, regulation of oogenesis and folliculogenesis, coordination of endocrinal functions, that in turn cause reduction of the reserve of primordial follicles and accelerate atresia leading to POI phenotype. Even though a large number of genes have been found, with some understanding of their pathogenesis, the precise genetic mechanisms underlying POI are unclear [10,56].

1.3.1 Chromosomal aneuploidies and structural abnormalities

X chromosome abnormalities have been described in both familial and non-familial POI patients, and represent about 13% of POI cases, which make them one of the commonest genetic causes [55]. Almost all chromosomal defects are involved, i.e. aneuploidies and rearrangements.

Regarding chromosomal aneuploidies, ***X monosomy or Turner's syndrome***, corresponds to a syndromic condition of POI with a prevalence of about 1 in 2,500 female births [66]. Among the various clinical features it is associated with gonadal dysgenesis with primary amenorrhea, sexual infantilism, webbing of the neck, cubitus valgus, short stature and cognitive deficits. Infertility in 45,X patients is caused by oocyte loss in the early stages of meiotic prophase resulting in ovarian dysgenesis. However, about 10% of Turner patients reach menarche because of a mosaic condition of chromosome X (45,X/46,XX) [53,67].

POI may be also associated with ***triple X syndrome or trisomy X*** (47,XXX) with endocrine findings of hypergonadotropic hypogonadism. Although not present in all cases, 21 POI patients have been reported until 2010 [68,69,70]. Hence, POI is not an essential characteristic of ***X chromosome tetrasomy*** (48,XXXX), even if detected in a 16-years-old girl without ovaries [71].

Regarding structural abnormalities, the commonest is ***Xq isochromosome***, where the resulting chromosome contains structurally identical arms with the same gene content. In particular, the isochromosome for the long arm (q) is the most common X structural abnormality in POI patients [49]. X chromosome ***deletions*** (mainly within Xp11 and Xq13-q26) and ***inversions,*** undetectable by conventional karyotyping, are also found in women with primary or secondary amenorrhea as well as with balanced and unbalanced ***X/autosomal translocations*** [56,72,73,74]. Conversely, very few cases of POI are reported to be carriers of ***translocation within autosomes***, being up to date only three: two cases with a karyotype 46,XX,t(2;15)(q32.3;q13.3) and one with a translocation between chromosomes 13 and 14 [53,75,76].

Molecular cytogenetic analyses of POI patients bearing X-rearrangements allowed the identification of two "**critical regions**" for ovarian growth and development. The first is located on the long arm of the X chromosome, spanning from Xq13.3 to Xq27 [53,59,77,78] with the following two main subregions further being identified: region 1 (Xq13.3-Xq21.1, locus POF2) associated more frequently with translocations and region 2 (Xq26-q27, locus POF1) where deletions are more frequent.

The second critical region is located on the short arm of the X chromosome between Xp11.1 and Xp21: patients with deletion in this region display primary amenorrhea and gonadal dysgenesis [53,55,79]. Although these regions have been associated with POI phenotype, several studies emphasized the role of other autosomal genes in the etiopathogenesis of the disorder.

1.3.2 Genes associated with 46,XX POI

1.3.2.1 Syndromic POI

As previously mentioned for Turner Syndrome, POI in some patients is associated with syndromic conditions whose responsible genetic cause is usually identified. The main genes whose mutations cause some of the syndromes associated with POI are listed and discussed below: *GALT*, *GNAS*, *POLG*, *AIRE*, *EIF2B*, *ATM*, *STAR*, *CYP17A1*, *CYP19A1* and *FOXL2*.

- POI occurs in all women affected by Galactosemia (OMIM #230400), a hereditary disorder of galactose metabolism due to homozygous mutation in ***GALT*** gene leading to GALT enzyme (galactose-1-phosphatase uridyltransferase) deficiency. Considering that optimum galactose concentrations are essential for normal ovarian development and function, its toxic accumulation cause oocytes apoptosis. Moreover these patients present complications in other organs such as liver, kidney and heart [10,80].

- Also loss of function mutations in the ***GNAS*** gene, encoding a G protein involved in the hormonal regulation of adenylate cyclase, cause a form of hormone resistance named pseudohypoparathyroidism type 1a which leads to ovarian insufficiency [48].
- Mutations in ***POLG*** gene (OMIM #174763), an enzyme that replicates human mitochondrial DNA, cause the autosomal dominant or recessive form of progressive external ophthalmoplegia (PEO) characterized by ptosis, limited eye movements and by variable POI manifestations from PA to SA.
- Mutations in the ***AIRE*** (Autoimmune Regulator protein) gene cause a disease called Autoimmune polyglandular syndrome type I (APS1; OMIM #240300) or autoimmune-polyendocrinopathycandidiasis-ectodermal dystrophy (APECED), an autosomal recessive disorder characterized by the presence of Addison's disease, hypoparathyroidism, and chronic mucocutaneous candidiasis. More than half of the patients present POI due to an autoimmune response to steroidogenic enzymes and steroid ovarian cells, leading to ovarian depletion [81].
- Ovarian leukodystrophy characterized by degeneration of brain white matter and POI (OMIM #603896) was firstly described by Schiffmann *et al.* [82]. The genetic basis of the disease is represented by mutations in any of the five subunits of the eukaryotic translation initiation factor ***EIF2B***.
- Another syndromic condition of POI is ataxia teleangiectasia (AT; OMIM #208900), an autosomal recessive neurodegenerative disorder characterized by infertility, degeneration of the brain which controls movement and speech, ocular telangiectases, chromosome instability, immunodeficiency, and cancer predisposition. Mutations in the ***ATM*** gene, which encodes for a protein kinase implicated in cell cycle regulation, DNA damage and metabolism, generally result in the total loss of the protein and are the underlying cause of this syndrome [10].
- Furthermore, also genes implicated in the pathways of steroid hormone synthesis are associated to syndromic POI. For instance, the ***StAR*** gene (Steroidogenic Acute Regulatory) encodes a protein involved in the conversion of cholesterol into pregnenolone; mutations in this gene lead to lipoid congenital adrenal hyperplasia, characterized by cholesterol accumulation with damage to ovarian cells. As well, mutations in enzyme of the cytochrome P450 superfamily ***CYP17A1***, which converts pregnenolone and progesterone into androgens (dehydroepiandrosterone and androstenedione), cause congenital adrenal hyperplasia with primary amenorrhea. Moreover, the ***CYP19A1*** gene, encoding the aromatase enzyme, converts

androgens into estrogens and mutations are associated to aromatase deficiency syndrome [10,53].

- Finally, the ***FOXL2*** (Forkhead Box L2) gene, belonging to the forkhead transcription factors family, encodes a transcription regulatory protein involved in ovarian development and function. It can be considered both a syndromic and a non-syndromic POI gene (POF3, OMIM #608996): indeed, mutations in this gene are associated with blepharophimosis-ptosis-epicanthus inversus syndrome (BPES, OMIM #110100) characterized by dominant malformation of the eyelids, but recently, mutations in this gene leading to POI without any malformations have also been found [83,84]. Moreover, Foxl2 knockout murine ovaries are characterized by an altered granulosa cells differentiation leading to premature activation of primordial follicles and consequent atresia [85].

1.3.2.2 Non-syndromic POI

Contrary to what described so far, in other patients POI appears to be isolated and its cause is not usually identified [53]. Nevertheless, there are many causative or candidate genes responsible for non-syndromic POI (**Fig.4**). The following list includes the genes whose involvement in the disorder is supported by several experimental evidences: *FMR1 and FMR2*, *TGFβ* superfamily *(BMP15*, *GDF9*, *INHA)*, *FSHR*, *LHR*, *NR5A1*, *FOXO* family, *NOBOX*, *FIGLA*, *STAG3*, *DIAPH2*, *POF1B*, *HFM1*, *NANOS3*, *MCM8*, *MCM9* ,*SYCE1* [10,86].

- ***FMR1*** **and** ***FMR2***

The fragile X mental retardation 1 gene (***FMR1***; OMIM #309550) is located at Xq27.3 and is responsible for the fragile X syndrome, a form of X-linked mental retardation (FRAXA). Mutations in *FMR1* lead to the expansion of a polymorphic CGG repeat in the 5' untraslated region, which can become unstable. Based on the number of repeats, three allelic groups can be found: normal alleles (from 6 to 55 repeats), premutated alleles (from 55 to 200 repeats) and full mutated allele (› than 200 repeats). When a full mutation is present, the gene is silenced by methylation, resulting in the absence of the FMR protein, which causes intellectual disability. Contrary, the premutation of the gene represents the cause of spontaneous 46,XX POI (POF1, OMIM #311360) in about 7% of sporadic cases and 21% of familial POI [87]. The underlying molecular mechanisms of how the triplet expansion causes POI are unknown, although it is hypothesized that premutated alleles remaining unmethylated might lead to increase in *FMR1* transcript levels that in turn can have a toxic effect during the reproductive life. In the same

way, the accumulation of FMR protein may impair the expression of genes required for oocyte development [56,87,88].

FMR2 (AFF2) gene (Fragile Mental Retardation 2, FRAXE, OMIM #309548) is located at Xq28, 600 kb distal from *FMR1* and likewise it has a trinucleotide repeat (GCC) within the untraslated region of exon 1. *FMR2* has also normal, premutated and full mutated alleles; thus, even if the mechanism underlying the pathogenesis of POI is not well known, it has been proposed a mechanism similar to that mentioned for *FMR1*. Moreover, microdeletions in this gene are associated with about 1.5% of POI cases, a higher percentage compared to the general population [55,80,89].

- ***BMP15***

BMP15 (also named *GDF9b*) gene (POF4, OMIM #300510), located at Xp11.2 within the Xp critical region, belongs to the TGFβ superfamily and encodes a pre-pro-protein consisting of a signal peptide, a pro-region and a mature domain that can form homo- or hetero-dimers with related factors, such as *GDF9* (see below) [10,90]. The main roles of BMP15 include: follicle maturation since the primordial phases of folliculogenesis; promotion of oocyte developmental competence; and determination of the ovulation quota for establishing the final number of ovarian follicles [10,91,92].

Regarding animal models, *Bmp15* knockout female mice and sheep were subfertile, with reduced ovary size [93]. In humans, the first *BMP15* genetic alteration associated with POI was detected in 2004 in two Italian sisters with hypergonadotropic ovarian failure characterized by primary amenorrhea and ovarian dysgenesis, who inherited the mutation from the father. Particularly, the p.Y235C mutation generated aberrant high-molecular weight products that were shown to act as dominant negative factor decreasing *in vitro* the growth of granulosa cells after stimulation with wild-type BMP15 [94]. Afterwards, many other mutations in *BMP15* gene have been identified in women with primary ovarian insufficiency, each involving the pro-domain [95,96,97,98,99,100]. The *BMP15* role in female fertility was also emphasized in a publication by Castronovo *et al.* [67] in which 40 patients with Turner syndrome (TS) have been screened by array comparative genomic hybridization (array-CGH) aiming at discovering genes with a possible role in ovarian development. Interestingly, one of the six TS patients, who presented spontaneous menarche, was found to bear a duplication involving the entire *BMP15* gene. Thus, the double dose of the genes might have preserved the correct amount of the protein needed for follicular functionality and pubertal development in the TS patient [67].

Nevertheless, the contribution of *BMP15* in the pathogenesis of the disorder is still uncertain in some cases because studies failed to find association of the gene with POI; moreover, *BMP15* mutations were also found at low percentage in control populations. Thus the overall findings merge in suggesting that mutations in the gene might confer predisposition to POI, considered as a multigenic disorder [10,55].

- ***GDF9***

Among the TGFβ family members, the growth differentiation factor 9 (*GDF9*, OMIM #601918) is located on chromosome 5 (5q31.1) and is homologous of *BMP15*. It encodes an oocyte-secreted growth factor with a pivotal role in the progression of folliculogenesis, mainly in the early stages of this process as proved by the presence in the human ovary of abundant expression levels of *GDF9* mRNA and protein [101]. According to animal models, homozygous *Gdf9* null female mice are infertile and characterized by significantly smaller ovaries with follicular growth blocked at the primary stages [102].

Therefore, *GDF9* can be considered a candidate gene for human POI [10,56]. Indeed, several studies have been performed for mutational screening of this gene in different populations, starting from the study of Takebayashi *et al.*, in which no mutation was found in 15 Japanese women with POI [103]. However, further screening performed on a more extensive number of patients revealed several point mutations probably causative of the disorder. Variations are all heterozygous and affects exclusively the pro-region (like the homologous *BMP15*) with a prevalence of 1.4% [98,104,105,106].

- ***INHA***

The inhibin (OMIM #147380), is a dimeric glycoprotein mainly produced in the gonads and constituted by two subunits a and b encoded by *INHA* and *INHB* genes respectively. These two genes present a different mode of action during the ovarian cycle: the mRNA levels of the first (*INHA*) increase in the middle of the cycle, whereas those of *INHB* increase in the middle of the follicular phase [107]. They play a role in the inhibition of the hypothalamic-pituitary-gonadal axis, regulating the secretion of FSH and allowing the ovulation of one mature oocyte. Particularly, given its pivotal role in regulating pituitary hormones secretion, *INHA* has become another candidate gene for mutational studies in humans [10,108]. Nevertheless several studies performed on different populations reported *INHA* variations, the association with POI never reached the statistical significance [109,110,111]. Thus, it is likely that variations in this gene may only confer susceptibility to develop POI [112].

- ***FSHR and LHCGR***

As part of the hypothalamo-pituitary-ovarian axis, FSH (Ovarian Dysgenesis 1, ODG1, OMIM #233300) and LH (OMIM #152790) glycoprotein hormone receptors, together with their binding hormones (FSH and LH) are pivotal for normal ovarian function. *FSHR* gene, mapping to 2p16.3 and encoding FSH receptor, is crucial for recruitment of ovarian follicles and follicular maturation from and beyond the preantral stage. Conversely LHR gene, which maps to 2p16.3 close to FSHR locus and encoding LH receptor, plays an important role in the maintenance of progesterone production by corpus luteum and promotes ovulation throughout oocyte maturation. Loss of function mutations involving these receptors cause gonadotropin resistance with hypergonadotropic hypogonadism and gonadal dysgenesis. Similar to the human condition, *Fshr* knockout mice are characterized by absent pubertal development, high gonadotropin and low estrogen levels, respectively, whereas *Lhcgr* knockout mice have small ovaries, and no preovulatory follicles [113,114].

The first mutation in *FSHR* gene related to POI etiology was reported by Aittomaki *et al.* [115], and up to date nine additional mutations have been described [86,116]. The mutant receptor shows an altered folding which affects its trafficking to the plasma membrane. When complete FSH resistance is found, the patients have no pubertal development and primary amenorrhea; on the other hand, partial forms of resistance are characterized by post-pubertal POI and secondary amenorrhea [10,115]. Regarding *LHR* gene and POI etiology, the first causative homozygous nonsense mutation was described by Latronico *et al.* [117], although biallelic inactivating variants of this gene are very rare in 46,XX POI women. Indeed, they represent a particular form of POI disease characterized by LH levels higher than those of FSH [86].

- ***NR5A1***

NR5A1 (POF7, OMIM #612964) gene, also termed steroidogenic factor 1 (SF1), is located at 11q13 and encodes a nuclear receptor essential for gonadal development. In particular, it is a key transcriptional regulator of genes involved in ovarian steroidogenesis, including *STAR*, *CYP17A1*, *CYP19A1*, *LH/CGR*, and *INHA* [10]. Its pivotal role in the ovary is also emphasized by the *Sf1* knockout mice, which are sterile, have fewer follicles, lack corpora lutea and have hemorrhagic cysts [55,118]. Mutations were firstly associated in patients with sex developmental disorders (46,XY DSD), but recently more than 19 mutations associated with gonadal dysgenesis with PA or SA, were detected in families affected by both 46,XY DSD and 46,XX POI, and families with isolated POI, establishing its involvement in the onset of the disorder [119].

- ***FOXO* family**

The family of forkhead transcription factors comprises over 100 members involved in several developmental processes. Similar to ***FOXL2*** gene cited before, *FOXO3a* (OMIM #602681), *FOXO1a* (OMIM #136533), and *FOXO4* (OMIM #300033) have been shown to play a key role in ovarian development and function. Regarding *FOXO3a* gene, located at 6q21, several studies revealed that in oocytes of knockout mice, the follicles prematurely activated result in an earlier reserve depletion and infertility. Furthermore, constitutive expression of Foxo3a determines a significant reduction in *BMP15* expression, suggesting a negative regulatory effect on this factor [120,121]. Thus, keeping into account that the ovarian mouse model phenotype well recapitulates the phenotype of POI patients, it is reasonable to consider *FOXO3a* a candidate gene for the disorder.

- ***NOBOX***

NOBOX gene (POF5, OMIM #611548), mapping at 7q35, is a homeobox gene preferentially expressed in oocytes, mainly from the primordial follicle stage through metaphase II, encoding a transcription factor that regulates oocyte-specific genes [122]. It is essential for folliculogenesis as demonstrated by *Nobox* knockout mice which are infertile and with atrophic ovaries. Moreover, follicles do not differentiate and become fibrous tissue in a manner similar to non-syndromic POI in human [123]. Mutational screening of the gene in several cohorts of POI women affected by primary or secondary amenorrhea, revealed many variations not reported in female controls [124,125,126].

- ***FIGLA***

FIGLA gene (POF6, OMIM #612310), located at 2p13.3, encodes a basic helix-loop-helix transcription factor that regulates the expression of zona pellucida genes. It is expressed during every ovarian follicular stage and in mature oocyte. Hu *et al.* [127] found that *FIGLA* expression in ovary is involved in activating genes required for oogenesis: effectively female knockout mice show rapid oocytes loss after birth and no primordial follicles formation. The ovarian phenotype noticed in mice is highly suggestive that this gene may be involved in human POI. This hypothesis has been supported by screening for mutations in *FIGLA* gene in 100 Chinese women with primary ovarian insufficiency as this study identified some variants not reported in female controls [10,128].

- ***DIAPH2 and POFIB***

As aforementioned, Xq13.3-21.1 is designated as critical region for premature ovarian failure 2 (POF2), clustering several breakpoints of translocation found in many POI patients [79,129].

One of the genes interrupted by the rearrangements involving this or the surrounding region, is diaphanous related formin 2, *DIAPH2* (POF2A, OMIM #300511). The gene is located at Xq21.33 and encodes a protein associated with cytoskeleton and involved in metaphase chromosome alignment during oogenesis. Its mutations may lead to alteration in the normal follicle cell proliferation, but its role in the etiology of the disorder is still unclear [130].

Despite no causal relationship between mutations of *DIAPH2* and POI phenotype has been reported, recent evidences describe overexpression of *DIAPH2* due to a position effect (mediated by a balanced X;1 translocation) in a girl with short stature, ovarian dysgenesis, and hirsutism [74]. The other gene mapping at Xq21 is *POF1B* (POF2B, OMIM #300604) associated with non-muscle myosin, which plays a pivotal role in cell division. Its role in POI etiology has been proved by Lacombe *et al.* who identified in 5 affected members of a Lebanese family a causative homozygous point mutation [131].

- ***HFM1***

This gene (POF9, OMIM #615724), located at 1p22.2, is an ATP-dependent DNA helicase exclusively expressed in ovary and testis with a role in genome integrity [132]. *Hfm1*-null mice ovaries, due to alteration in the formation of synaptonemal complex, showed a significant reduction in size, an increase in stromal cells, and a reduced number of follicles [133]. A recent study has identified recessive mutations in 2 sisters and an unrelated woman with POI, supporting the role of *HFM1* in ovarian failure also in human [134].

- ***STAG3***

The *STAG3* gene (POF8, OMIM #615723), located at 7q22.1, encodes a subunit of cohesin, a large protein complex that plays an important role in proper pairing and segregation of chromosomes during meiosis. It is the most recently reported gene, whose homozygous point mutations have been associated to non-syndromic POI (PA) by two studies (Caburet *et al.* [135] and Le Quesne Stabej *et al.* [136]). Moreover, the ovaries of *Stag3*-null female mice presented with a distinct lack of oocytes and ovarian follicles, indicating severe and very early ovarian dysgenesis [86,135].

- ***NANOS3***

The gene encodes a CCHC zinc finger RNA interacting protein (OMIM #608229), recently reported to be an important gene for germ cell development, controlling their migration and survival in the first stage of development. The knockout mouse model of *Nanos3* had reduced gonad size and was infertile, in keeping with the relationship of low levels of mRNA with a decreased number of germ cells [137]. Two pathogenic mutations, a missense variant in a Chinese POF woman, and a homozygous variant in two Brazilian PA sisters, have been reported

so far highlighting a role of the gene in the etiogenesis of the disorder, mainly mediating a protective effect against apoptosis in primordial germ cells [138,139].

- ***MCM8 and MCM9***

The proteins encoded by these genes (*MCM8*, POF10, OMIM #612885; *MCM9*, ODG4, OMIM #616185) are two highly conserved mini-chromosome maintenance proteins (MCM), which are essential for mediating Double Strand Break repair during gametogenesis, and for replication fork maintenance. Evidences about their role in POI pathogenesis are reported in two studies in which whole exome sequencing analysis of consanguineous Saudi Arabian and Kurdish families disclosed two homozygous mutations in *MCM8* and *MCM9*, respectively [140,141]. In addition, *Mcm8* null female mice are sterile and have arrested primary follicles, whereas *Mcm9* null females are sterile as well and ovaries lack oocytes completely [142].

- ***SYCE1***

The synaptonemal complex central element protein 1 gene encodes a component of the synaptonemal complex which is important for maintaining paired chromosome homologues for crossover during meiosis [143,144,145]. Only one homozygous mutation has been described so far in human by de Vries *et al.* [146], namely in two Muslim Arab POI sisters from a consanguineous family. Considering that female mice *Syce1-/-* are infertile and had small ovaries with a decreased number of follicles [147], the mutation has been strongly suggested to be the causative factor in the POI sisters.

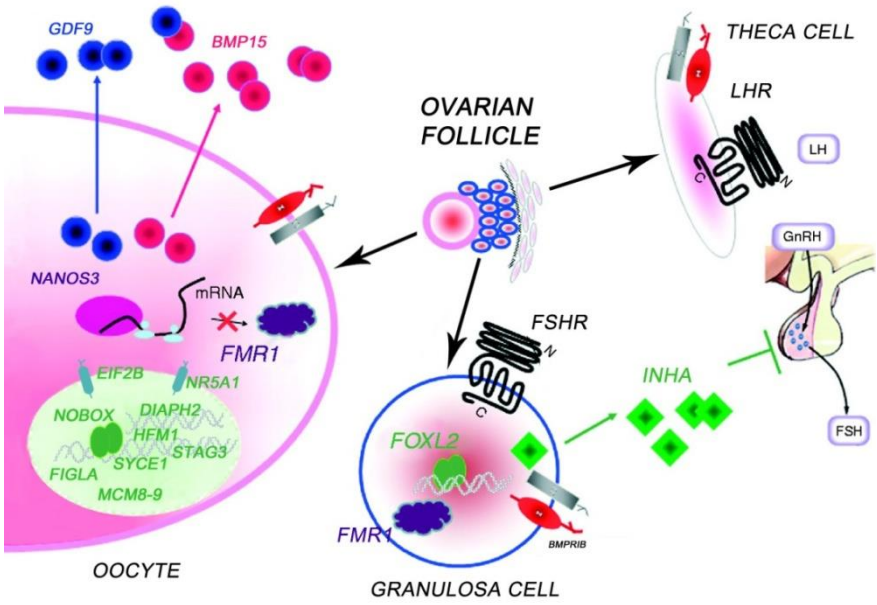

Fig.4. Schematic view of the principal genes known to be involved in POI pathogenesis. *Enlargement of oocyte, granulosa and theca cells is shown to highlight the site of expression of each ovarian gene. The image was modified from Persani et al., 2010 [10].*

1.4 High throughput genome wide analyses in POI

Whereas POI is considered idiopathic in 50-80% of cases, it is clear that some genes responsible of the POI phenotype are not yet known and the molecular pathways underlying the complex mechanism of fertility in women still need to be elucidated in order to achieve a more accurate genetic characterization of the patients, formulate a proper diagnosis, and ensure the correct therapeutic intervention.

The best approach to understand and discover the molecular basis of a complex and heterogeneous disorder is a genome wide investigation. As a matter of fact, the genetic of POI has been recently explored through:

- Genome Wide Association studies (GWAS);
- array Comparative Genomic Hybridization (array CGH) looking for associated POI common CNVs;
- rare SNVs and CNVs detection via Next Generation Sequencing (NGS) and array CGH, respectively;

Regarding the investigation for **common variants** possibly causal or risk factor for POI, in several studies employing GWAS evaluated the frequency of polymorphisms in unrelated affected individuals and healthy controls, discovering common variants with a high significant association with POI (e.g. in *ADAMTS19*, *IGF2R*, *PTHB1* genes) [148,149,150]. Moreover, Qin *et al.* conducted in 2012 the first largest-scale GWAS, investigating 391 Chinese Han patients with POI (32 PA patients were included because of the early SA of the other 359 women) and 895 Chinese female controls. No candidate genes emerged but several SNPs with a high association were found at 8q22.3 [151].

In the same way, Aboura and collaborators [152] published in 2009 the first study searching for common CNVs possibly associated with the disorder by array CGH. The project included a cohort of 99 46,XX Caucasian POI women (66 SA and 33 PA patients) and reported eight statistically significant common CNVs with different frequency compared to a control group. The new putative candidate genes associated to POI and identified in the CNVs were involved in reproductive disease (*DNAH5* and *NAIP*), in reproductive endocrinology (*DUSP22* and *NUPR1*), and in folliculogenesis (*AKT1*). Nevertheless, this finding was not confirmed by a recent publication by Castronovo *et al.* [67], in which the frequency of common CNVs found in 34 Turner syndrome (TS) patients was compared with that found in a control population of 55

healthy women as no enrichment of these common CNVs was detected in TS patients in comparison to controls.

In 2011, Liao and collaborators [153], following the previous approach, analyzed 15 Chinese primary amenorrea patients through array-CGH to discover common CNVs. The study allowed to evidence in four patients a microdeletion at 17q21.31-21.32 containing the *NSF* gene (N-ethylmalemide-sensitive factor), encoding a factor involved in female reproductive secretory pathway.

Processing these patients by high throughput technologies allowed to shed light not only on genes and pathways possibly related to POI etiology but also on genes and pathways involved in ovarian-related homeostasis conditions and ovarian aging [154]. This concept is supported by the recent paper by Day *et al.* [155] on the GWAS screening of 70,000 women between 40-60 years old to identify common and low-frequency variants associated with age at natural menopause. In addition to findings on SNPs strongly associated to DNA Damage Response pathways, the statistical analysis performed allowed to identify an enrichment in known POI genes, thus highlighting that some genes involved in physiological menopause are also responsible, if disrupted, for primary ovarian insufficiency.

However, common variants might have only a susceptibility role in the onset of the disease, and can explain only a small proportion of the genetic basis of POI. Recently there is an increasing interest towards the identification of **rare and highly penetrant variants** (CNVs and SNVs) able to play an important role in POI pathogenesis. To this purpose, array CGH technology as well as NGS are the two main approaches recently used.

1.4.1 array CGH and NGS technologies searching for rare variants

CNVs are at the basis of genetic variability in the population and their key role in modulating the expression of the genes included has been widely demonstrated for pathological conditions including schizophrenia, autism, attention deficit hyperactivity disorder, etc. [156,157,158].

Up to now, two studies have been published searching for rare CNVs in cohorts of 46,XX POI patients [159,160]. In the first [160], the analysis of 30 patients with PA and 44 with SA by a low resolution array-CGH platform (105A), revealed 44 rare CNVs containing genes involved in meiosis, DNA repair and folliculogenesis. In the most recent study by Norling *et al.* [159], custom 1M high resolution array CGH enriched for 78 genes involved in ovarian development was used to analyze 17 PA and 9 SA women. Eleven rare CNVs have been indentified, involving *GDF9*,

three *NOBOX*-binding elements, and other candidate genes (e.g. *DNAH6*, *TSPYL6*, *SMARCC1*, *CSPG5*, *ZFR2*).

Studies using whole genome SNP array [161,162] and array CGH targeted analysis of the X chromosome [163,164,165] have also been reported, allowing the detection of several novel genes involved in ovarian development, both on autosomes and on X chromosome.

However, many of the patients analyzed by array approaches turned out to be negative for any rare variant involving genes possibly related to POI: for this purpose the application of NGS might be a valuable complementary tool to discover the molecular defects of the pathology. Based on the present literature, whole exome sequencing (WES) analysis has been performed in a few familial and sporadic cases. Particularly, in two POI cases from consanguineous families (PA patients), SNVs implicated in ovarian function, namely in *STAG3* and *SYCE1*, were revealed respectively [135,136,146]. Katari *et al* [116], recently analyzed by array-CGH+SNP and WES two Indian sisters with PA born to consanguineous parents and reported a homozygous mutation in the *FSHR* gene. Actually in a publication by Fonseca *et al.* [166] a targeted-sequencing method of 70 candidate genes performed in twelve unrelated Colombian women affected by non-syndromic POI (SA patients) allowed the identification of novel mutations in *ADAMTS19* and *BMPR2*, probably related to POI onset [166]. Furthermore Qin *et al.* performed WES in three related women with SA and found a missense mutation in *PGBD3* [167], encoding a protein with a role in DNA repair and transcription regulation. To validate exome data and detect independent mutations, they further performed Sanger sequencing of *PGBD3* in 432 sporadic POI patients and 400 matched control females. This analysis revealed another missense mutation and a nonsense mutation in two SA patients [167].

Based on the above breakthroughs, the whole genome high throughput approaches could unveil the genetic heterogeneity of POI and represent powerful tools for future diagnostic and prognostic purposes aimed at the analysis of a subset of POI candidate genes.

2. MATERIALS AND METHODS

2.1 Patients

Genomic DNA (gDNA) from 67 women affected by the most severe POI phenotype (primary amenorrhea, PA) were collected. From the biochemical and endocrinological standpoint, the patients presented hypergonadotropic hypogonadism with high levels of gonadotropins (FSH › 40 IU/L). Conventional karyotyping resulted 46,XX for all the women and the genetic tests usually performed for POI diagnosis (e.g. genetic screening of low XO mosaicism, *FMR1* premutation [49]), were negative. Previous ovarian surgery, or exposure to chemotherapy or radiotherapy were not present in any patients.

The cohort of 67 POI patients includes 53 sporadic (79.1%) and 14 familial (20.9%) cases. Among familial cases six couples of sisters, and two patients with a relative affected by SA were included in the analyzed cohort.

In order to detect copy number variants (CNVs), gDNA from all patients was used to perform array CGH analysis (Agilent Technology). Two kits with different resolutions were used, specifically:

- 9 women by 244K Human Genome CGH Microarray, spanning approximately 236,381 probes with an average spatial resolution of 8.9 kb (7.4 kb in gene-enriched genomic regions);
- 58 patients by 400K Human Genome CGH Microarray, spanning approximately 411,056 probes with an average spatial resolution of 5.3 kb (4.6 kb in RefSeq genes).

In the event of detection of rare CNVs, where possible, patients' parents were analyzed to characterize the origin of the unbalanced microrearrangement (*de novo* or inherited).

2.2 Molecular karyotyping by array-based Comparative Genomic Hybridization (array CGH) analysis

2.2.1 Extraction of genomic DNA (gDNA) from peripheral blood

For gDNA extraction from fresh or frozen (-80°C) peripheral blood collected in anticoagulant (EDTA) tubes, the *GenElute Blood Genomic DNA Kit (Sigma)* was used following the manufacturer's instructions. Briefly:

- collect 500 µl of whole blood in a 2 ml tube;
- add 50 µl of Proteinase K;
- add 40 µl of RNase A Solution and incubate for 2 minutes at room temperature;
- add 550 µl of Lysis Solution C, then vortex thoroughly for 15 seconds and incubate at 55°C for 10 minutes;
- add 550 µl of ethanol (95-100%) to the lysate and mix thoroughly by vortexing for 15 seconds;
- in the meantime, add 550 µl of the Column Preparation Solution to each pre-assembled GenElute Miniprep Binding Column, then centrifuge at 12,000 X *g* for 1 minute and discard the flow-through liquid;
- transfer 600 µl of the sample into the treated column, centrifuge at 6,500 X *g* for 1 minute and discard the flow-through liquid;
- repeat this step until the entire sample was loaded into the column, discard the collection tube containing the flow-through liquid and place the binding column in a new 2 ml tube;
- add 500 µl of Prewash Solution and centrifuge at 6,500 X *g* for 1 minute. Discard the collection tube containing the flow-through liquid and place the binding column in a new 2 ml tube;
- add 500 µl of Wash Solution and centrifuge for 3 minutes at maximum speed (12,000–16,000 X *g*) to dry the binding column;
- place the binding column in a new 2 ml tube and let it dry under a safety hood for 10 minutes;
- add 50 µl of Elution Solution directly into the centre of the binding column and incubate for 5 minutes at room temperature, then centrifuge at 6,500 X *g* for 2 minutes.
- The concentration and quality of the gDNA are determined by agarose gel electrophoresis and by a spectrophotometric analysis by using the NanoDrop ND-1000 UV-VIS Spectrophotometer.

2.2.2 gDNA restriction digestion (*SureTag DNA Labeling Kit, Agilent*)

Test and reference DNA must be processed separately. The amount of DNA required, which must be the same for test and reference, depends on the slide used as indicated in the table below.

	array 244K	array 400K
DNA (µg)	0.5-3 µg	0.5-1.5 µg

For 244K and 400K arrays:

- add the amount of gDNA to nuclease-free tube and add nuclease-free water up to a volume of 20.2 µl;
- prepare for each sample the following Digestion Master Mix:

- 2 µl of nuclease-free water
- 2.6 µl of 10X Restriction Enzyme Buffer
- 0.2 µl of BSA
- 0.5 µl of AluI (10 U/µl)
- 0.5 µl of RsaI (10 U/µl)

- add 5.8 µL of Digestion Master Mix to each reaction tube containing the gDNA to make a total volume of 26 µl;
- incubate at 37°C for 2 hours;
- incubate at 65°C for 20 minutes to inactivate the restriction enzymes;
- move the sample tubes to ice;
- assess the quality of the digestion by agarose gel electrophoresis run (Agarose 0.8% in TAE 1X and Gel Red Nucleic Acid Stain) loading 2 µl of digested gDNA. The majority of the digested products should be between 200 bp and 500 bp in length.

2.2.3 Fluorescent Labeling of gDNA (*SureTag DNA Labeling Kit, Agilent*)

For labeling reaction fluorescent cyanines (Cy3 and Cy5) are used. The reaction must be performed in the dark. The test sample is labeled with Cy3 and the reference with Cy5.

For 244K and 400K arrays:

- add 5 µl of Random Primers to each reaction tube containing 24 µl of digested gDNA;
- incubate at 95-100°C for 5 minutes;
- move to ice and incubate on ice for 5 minutes;
- prepare for each sample the following Labeling Master Mix:

- 2 µl of nuclease-free water

- 10 µl of 5X Reaction Buffer
- 5 µl of 10X dNTPs
- 3 µl of Cy3-dUTP (test DNA) or Cy5-dUTP (reference DNA)
- 1 µl of Exo-Klenov enzyme

- add 21 µl of Labeling Master Mix to digested gDNA to make a total volume of 50 µL;
- incubate at 37°C for 2 hours;
- incubate at 65°C for 10 minutes to inactivate the enzyme, then move to ice.

2.2.4 Clean-up of labeled gDNA

- Add 430 µl of 1X TE (pH 8, Promega) to each sample and transfer into a 2 ml collection tube with a purification column;
- spin for 10 minutes at 14,000 X *g*, discard the flow-through and place the columns back in the collection tubes;
- add 480 µl of 1X TE to each column and spin for 10 minutes at 14,000 X *g*;
- invert the filter into a fresh 2 ml tube, spin for 1 minute at 1,000 X *g* to collect purified sample;
- measure the volume of each eluate;
- add 1X TE to bring the sample volume to 80.5 µl (244K) or 41 µl (400K);
- take 1.5 µl of each sample to determine the yield and specific activity by using the NanoDrop ND-1000 UV-VIS Spectrophotometer.

2.2.5 Hybridization with Oligo aCGH Hybridization Kits (Agilent)

The hybridization reaction must be performed in the dark:
- combine test and reference sample;
- add the reagents as described in the table below to prepare the Hybridization Master Mix:

	244K	**400K**
Labeled gDNA	158 µl	79 µl
Human Cot-1 DNA (1mg/ml)	50 µl	25 µl
Agilent 10X Blocking agent	52 µl	26 µl
Agilent 2X Hybridization buffer	260 µl	130µl
Total volume	520 µl	260 µl

- incubate at 95°C for 5 minutes and subsequently at 37°C for 30 minutes;
- dispense hybridization sample mixture slowly onto the gasket well, previously loaded into the Agilent Hybridization Chamber;

- put a microarray slide onto the gasket, with the numeric barcode side facing up and the Agilent-labeled barcode facing down, and close the Agilent Hybridization Chamber;
- place the assembled chamber into a preheated oven at 67°C and leave for 24-48 hours.

2.2.6 Microarray wash with Oligo aCGH Wash Buffer 1 and 2 (Agilent)

All steps must be performed in the dark and under fume hood:
- disassemble the Agilent Hybridization Chamber;
- dry the microarray-gasket sandwich submerging in Wash Buffer 1 for the disassembly;
- put the microarray slide into the slide rack in the slide-staining dish containing Wash Buffer 1 and keep stirring for 5 minutes in a magnetic stirrer;
- wash the slide in Wash Buffer 2 pre-warmed at 37°C for 1 minute;
- submerge the slide in Acetonitrile for 1 minute;
- transfer slide rack in Stabilization and Drying Solution, and stir for 30 seconds;
- slowly remove the slide rack trying to minimize droplets on the slides and put in the slide holder for the scanning with Microarray Scanner (Agilent).

2.2.7 Microarray Scanning and Analysis

After scanning is completed, Feature Extraction software extracts data from the scanned microarray image .tif and translated into a file .txt. CytoGenomics 3.0 software associates each oligo probe on the slide with the corresponding position in the genome and the log ratios, allowing to identify aberrations.

2.3 CNVs classification, analysis of the gene content, and evaluation of ovarian genes enrichment

Once a CNV was detected, the first step was to understand whether it was a rare variant or a CNV already reported in healthy controls according to the Database of Genomic Variants (DGV, http://projects.tcag.ca/variation/). In the case of rare CNVs, namely never reported in their full extension in the DGV or reported at very low frequency (« 1%), array CGH analysis was extended to the parents (only when available), to determine whether the rearrangement is de novo or inherited. Of note, considering the Primary Ovarian Insufficiency disease, which shows an oligo-/polygenic genetic etiology, it cannot be assumed a priori that an inherited CNV is benign *per se*. Indeed, the combination of several variants in different loci, both *de novo* and/or inherited, often leads to the onset of the disease, which is not present in the parents as they do not share the same combination of variants identified in the children. Besides we cannot exclude the possibility that any of the "common" CNVs detected in our series may have contributed to POI susceptibility.

2.3.1 Analysis of the gene content of the identified rare CNVs

The analysis of the gene content of rare CNVs was carried out investigating public databases *ad hoc*, useful for the evaluation of possible pathogenetic significance of the imbalances identified. In detail:

- the UCSC Genome Browser (http://genome.ucsc.edu/cgi-bin/hgGateway), NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/gene), GeneCards (http://www.genecards.org/) which are useful to collect information about the genes involved in CNVs (i.e. gene function, molecular structure, presence of different isoforms, expression in different tissues, pathways in which they are involved, etc.);
- the Decipher database (http://decipher.sanger.ac.uk/) and ISCA Consortium database (www.iscaconsortium.org), which collect pathogenetic CNVs reported to date as well as the clinical description of the related patients;
- the OMIM database (http://www.ncbi.nlm.nih.gov/omim; Online Mendelian Inheritance in Man), which provides information about all Mendelian disorders described to date and about more than 12000 genes, focusing particularly on genotype-phenotype correlation;

- PubMed (http://www.ncbi.nlm.nih.gov/pubmed), which is useful to examine the international medical literature;
- the Ovarian Kaleidoscope Database (OKdb) (http://okdb.appliedbioinfo.net), a public database which provides information regarding the known biological function, the expression pattern, the mutations identified in human or animal model and a collection of all the literature of the genes known to play a role in the ovary.

Thus, based on the gene content, the rare CNVs identified were classified as possibly related and unrelated to POI pathogenesis. Moreover, for the genes possibly related to POI, an association score to the disease from 1 (low) to 4 (high) was established. In addition, their classification in four functional categories was performed (A, B, C, D).

To discover interactions between the proteins encoded by target genes several bioinformatic tools were interrogated:

- STRING (http://string-db.org/), medium confidence 0.400, without textmining;
- ToppGene (https://toppgene.cchmc.org/);
- geneMania (http://www.genemania.org/).

2.3.2 Analysis of an *ad hoc* control group

To find a real enrichment in potentially pathogenic ovarian CNVs in POI women, a comparison between array CGH data of patients and controls was performed. *Ad hoc* control group composed by 140 healthy women with normal 46,XX karyotype who had a normal reproductive life and reached physiological menopause was used. Array CGH 244K and 400K previously performed in our laboratory on 140 control women were analyzed and CNVs classification was carried out as in POI patients. Moreover, a POI association score from 1 (low) to 4 (high) to the ovarian genes found in controls was assigned.

A statistical comparison to investigate significant differences between array CGH data of patients and controls was performed using the *chi-square test*. *Fisher's exact test* was used to evaluate whether the prevalence of the common CNVs was significantly different between POI and controls. P-value < 5% assigns a statistically significant test.

2.4 Validation and molecular characterization of rare ovarian CNVs

Once a rare ovarian CNV has been identified, a validation following the flow chart described below has been performed (**Fig.5**).

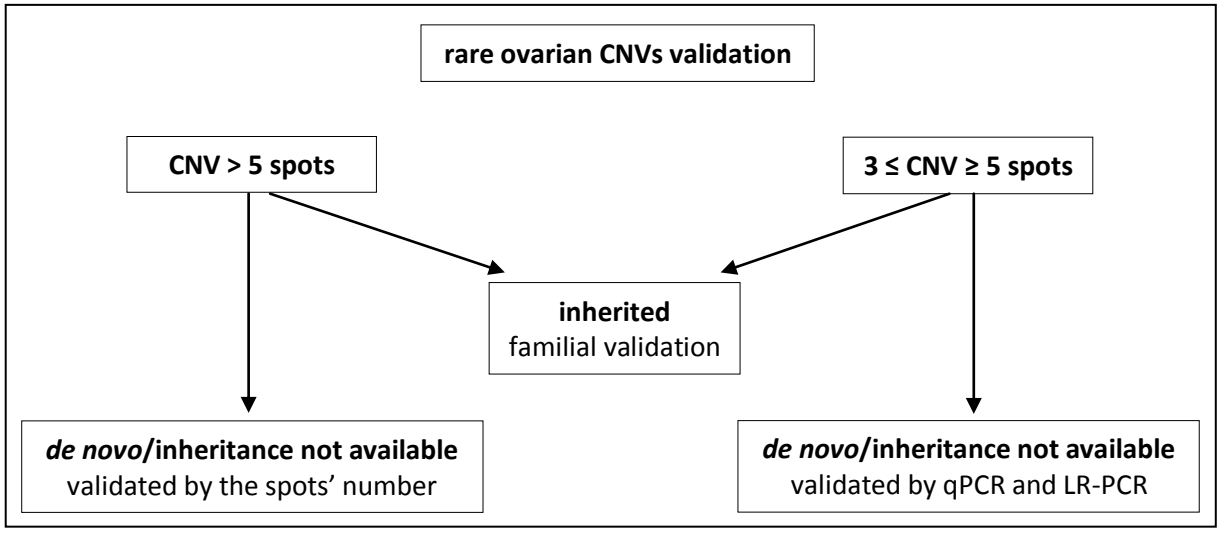

Fig.5. *Flow-chart used to validate the rare ovarian CNVs.*

2.4.1 Validation and molecular characterization of rare ovarian CNVs by qPCR

To verify that rare "ovarian" CNVs covered by a number of spots between 3 and 5 were not fakes, experiments of Real Time PCR on gDNA (qPCR) using the method of SYBR® Green were carried out to monitor the fluorescence during the amplification reaction. Moreover qPCR was used to characterize the breakpoints of some rare CNVs identified in patients POI 9, 11 and 46 and to investigate in patients POI 9, 11, 17, 60, 62 and some controls (C_F38, C_F40, C_F77, C_F99, C_F103) the "common" rearrangement at 10q26.3 involving *SYCE1* gene previously associated to the disorder.

Specific primers for the genomic regions to be investigated were picked by the software Primer3 (http://frodo.wi.mit.edu/primer3/) (**Tab.1**). Particularly, the qPCR primers need different features compared to the primers normally used in a standard PCR reaction. Before proceeding with their order, the following variables have to be modify in the input command:

Product Size Ranges (n. of bases) →*60-100*

Primer size (n. of bases) →*Min 18; Opt 20; Max 22*

Primer Tm (°C) →*Min 59; Opt 60; Max 63*

Product Tm (°C) →*Min 65; Opt 80; Max 90*

Max Self Complementarity →*4*

Tab.1. List of validated and characterized CNVs and qPCR primers used.

ID POI	Gain/Loss	qPCR Probe	Primers
7	Loss in 9p21.3	1	Primer Forward: AGTCCCAGGAATCAAGGAATC Primer Reverse:GAAAAGAGCTTCACAGGCAGTT Amplicon: 77 bp
		2	Primer Forward: GGCAAATGACAAACACACAAA Primer Reverse:TGCTCCAACATCAGACATCC Amplicon: 100 bp
		3	Primer Forward: CGAACCCTGTAACCTGTTGG Primer Reverse:CGCCTGTTTCACCCTATGAA Amplicon: 84 bp
9-11	Loss in 5p13.1	1	Primer Forward: AGGGGCTTTTGCATACCAC Primer Reverse:aactCAACACTGCCCTCTGC Amplicon: 66 bp
		2	Primer Forward: GGTTGCTGACATTTTGATGC Primer Reverse:AATGGAAGGGTAAGCTGTTCTG Amplicon: 87 bp
		3	Primer Forward: CACATCAAGGCTCCGAATCT Primer Reverse:CATGAATTGACTGGGCATAAAG Amplicon: 69 bp
		4	Primer Forward: GGCTTGGCACAGACTTTCTC Primer Reverse:TTGTGGCTATTTCCCAGTCC Amplicon: 63 bp
		5	Primer Forward: CTCCTACCCCACCACCACTA Primer Reverse:CCAAAGTCAGGAAATGAGTTGA Amplicon: 74 bp
		6	Primer Forward: AATAGCTTAATCCTGGGGAAGG Primer Reverse:TATTTGGAATGCCCTTGGTG Amplicon: 98 bp
18	Loss in 14q22.2	1	Primer Forward: AAGTCTCCAAAAATCGTGCAA Primer Reverse:TTCCATACCTGGGTCCATAAA Amplicon: 71 bp
		2	Primer Forward: TTGAAGGCAGAAGCCCTGT Primer Reverse:CAACAACACAGCACAAGATGC Amplicon: 68 bp
		3	Primer Forward: TGTTGCTGTTATCTTTGCATGG Primer Reverse:TCCATTGTTCAGTGCATTCC Amplicon: 70 bp
46	Loss in 15q25.2	1	Primer Forward: TCCTGCCCTTATTTTCAGCA Primer Reverse:CCTTCTTTCCCGTTGACAGT Amplicon: 98 bp
		2	Primer Forward: CAAAGGTGAGGGAAATGGTC Primer Reverse:AAGGAGGAGGAGGAGGATTTT Amplicon: 72 bp
		3	Primer Forward: GGCTCACCAGGAATAAACGA Primer Reverse: CTTTCCGACACAGGAAGGAG Amplicon: 61 bp
		4	Primer Forward: TGGGATTTGGGGTTTTAGGT Primer Reverse: TGAGACAACTCTGGAGCCTTT Amplicon: 76 bp
		5	Primer Forward: AGCCTGTCTTTCCCTAGCATC Primer Reverse: GGCCGTAGAGATGGGGTACT Amplicon: 61 bp
		6	Primer Forward: ACGCATTCGGTAGTCCAATC Primer Reverse: TGACTGTTGTGTGCCCAAGT Amplicon: 88 bp
		7	Primer Forward: AAGCTCAACCCCAGTCAGTC Primer Reverse: GGGATTATTTTGCCACCCTAA Amplicon: 84 bp

		8	Primer Forward: CATCCGTAAGTGTGGCTGTG Primer Reverse: TAAAATAGCCCTGCCCCAGA Amplicon: 93 bp
		9	Primer Forward: ACTACCAGGCCACCCAAAG Primer Reverse: CTGGCTGAGTTTCGCTTTTT Amplicon: 64 bp
		10	Primer Forward: CAACGTGCTCTGGGAATACA Primer Reverse:TTTGGGGACTCATCACTGGT Amplicon: 78 bp
		11	Primer Forward: CCTGTGCCTTCCAGAATAGG Primer Reverse:TCAACTACTGGGCTGTGTCG Amplicon: 72 bp
		12	Primer Forward: CCTCTAGGAATGAGGCAGCA Primer Reverse: CCGAGGCTTTTTCATGTCA Amplicon: 72 bp
		13	Primer Forward: CTCGCGCCACACTACCTG Primer Reverse:GCTTCACACCTTCGCTTGG Amplicon: 70 bp
		14	Primer Forward: ACCCCGGCAGCTTATGAA Primer Reverse:ACACAGGTGCTCCGCATC Amplicon: 74 bp
		15	Primer Forward: GTCCTCTTTGGACCGGAGTT Primer Reverse: TCCCTTGCCCTATTTCTGTG Amplicon: 90 bp
		16	Primer Forward: TCCTCTCCTGGCACTTAGGA Primer Reverse:CCCTCTCCCCTTAGCCATTA Amplicon: 72 bp
		17	Primer Forward: TCGTTTCCTCCTCTCTGACC Primer Reverse: TGGGCACTATTCTGGGTTTC Amplicon: 65 bp
55	Loss in Xq21.33	1	Primer Forward: CAGGCTCTTACAAATGGAGCA Primer Reverse: GCGGAGTGGAGGGACATATT Amplicon: 98 bp
		2	Primer Forward: CCACCTTGGGAATTTTGTGT Primer Reverse: AGATATTGGGGGCACATCAT Amplicon: 60 bp
		3	Primer Forward: GTTCTTGTGGCGATGCTTTT Primer Reverse: GCCCACTGCTGGAACACTA Amplicon: 64 bp
62	Gain in 20q13.2	1	Primer Forward: TAGAAAGCAAAGCGGGAGAG Primer Reverse:CACCCCCACCCAACTCTT Amplicon: 66 bp
		2	Primer Forward: TTCAGCTTCTCAACTCCAGGT Primer Reverse: AAGTCGTAACCAGCCCAATG Amplicon: 69 bp
		3	Primer Forward: AAAACAACCAAAAATGGGAAAA Primer Reverse: GAGCCCGCTGGGAGTCTA Amplicon: 72 bp
		4	Primer Forward: GCTGGTCTCAAATGCTCGAT Primer Reverse: AAACACTGGATTGGCTGAGTG Amplicon: 82 bp
66	Loss in 2p14	1	Primer Forward: GCGTCAGTAATAGCCATCAAAA Primer Reverse: TTGCACCCAGATATTCAGGA Amplicon: 67 bp
		2	Primer Forward: ATGGCAGATGAGGAAAATGC Primer Reverse: GGCTCTGTGTTGCTGAGGTT Amplicon: 71 bp
		3	Primer Forward: GAGGAGGGGGAGCAGTAAAT Primer Reverse: ATGGAGAGTCAGAGCCTGGA Amplicon: 60 bp
9-11-17-	Gain/Loss in	1	Primer Forward: TGCAGCCAGTTAGGTGAAGA

60-62	10q26.3		Primer Reverse:GCAGGGATGTTATCTAGGGTCA Amplicon: 78 bp
		2	Primer Forward: TGCCACATCTCACTGTAGGG Primer Reverse:GCAGGTGCTCTGTCATGTTC Amplicon: 89 bp
		3	Primer Forward: CCCCACTGTCAAGAACCAGT Primer Reverse:TCCTCTCCCTAGAACCAGCA Amplicon: 67 bp
		4	Primer Forward: CTGCCTCCCACCACTACCT Primer Reverse:AGGTCCCTGACATCGAAGG Amplicon: 81 bp
		5	Primer Forward: CAGAGGCTTTCGGGAGAAC Primer Reverse:TCTTCCCCTGGCTCTTCAC Amplicon: 85 bp

Each experiment was performed in triplicate on patients and two controls gDNA. For each couple of primers Mix 1 and Mix 2 has been prepared:

Mix 1	
SYBR Green PCR Master Mix (Applied Biosystem) 2X	7.5 µl
Primer Forward (10 pmol/µl)	0.45 µl
Primer Reverse (10 pmol/µl)	0.45 µl
H_2O	3.6 µl
Total volume	12 µl

Mix 2	
DNA	[10 ng]
H_2O	up to 3 µl
Total volume	3 µl

Each qPCR reaction, performed in a final volume of 15 µl, is composed by 12 µl of Mix 1 and 3 µl of Mix 2. The reactions were incubated using the ABI Prism 7900 (Applied Biosystem) with the SDS software v2.3. The PCR conditions used are the following:

50 °C	2 min	
95 °C	10 min	
95 °C	15 sec	x 40 cycles
60 °C	1 min	

The experiment is analyzed by calculating the "threshold cycle" (Ct) of each reaction using the RQ Manager 1.2 software (Applied Biosystem). The gDNA copy number in the regions of interest in patients and controls was determined by the $2^{-\Delta\Delta Ct}$ relative quantification method [168] using *PCNT* and *FGFR3* as normalizing genes (**Tab.2**). In this study the values with standard deviation over 0.5 were excluded and the results were analyzed by two-tailed Student's *t* test in order to investigate the statistical significance of the DNA copy number in patients compared to controls ($p< 0.05$).

Tab.2. List of the normalizing genes qPCR primers.

Gene	Primers
PCNT	Primer Forward: TCCAGAACATTCCTTGACAGAG Primer Reverse: GTACCCCTCCCAATCTTTGC Amplicon: 68 bp
FGFR3	Primer Forward: TAGGGTACTTTGGGGCACGA Primer Reverse: GCGTTCTGACTTCCACCAG Amplicon: 65 bp

2.4.2 Validation and molecular characterization of some rare ovarian CNVs by Long Range PCR and Sanger sequencing

Long Range PCR

In order to validate and to characterize the breakpoint junctions of the rare ovarian CNVs identified in patients POI 3, 8, 10, 44 and 46 we designed specific primer pairs using Primer3 software (**Tab.3**). The Long Range PCR (LR-PCR) primers need different features compared to the primers normally used in a standard PCR reaction:

Primer size (n. of bases)→*Min 22; Opt 28; Max 34*

Primer Tm (°C) →*Min 58; Opt 65; Max 80*

The melting temperature (Tm) of each primer, the respective annealing temperature (Ta) and the PCR conditions, were obtained by the software Optimase™ ProtocolWriter (www.mutationdiscovery.com).

Tab.3. List of LR-PCR primers used for CNVs validation and characterization.

ID POI	Gene	LR-PCR Probe	Primers	PCR Conditions
3	*RYR3*	1	Primer Forward: CCCTGAACCTTGTTTCTTTCTGTCTAT Primer Reverse: AGAGGGAGATAACATCTTACCAGATCA Amplicon expected: ~ 6-7 kb	Tm = 59.9°C Tm = 58.4°C Ta = 62.2°C
8-10	*TP63*	1	Primer Forward: AGGGACATTACTTCTCTTGCAATCACT Primer Reverse: GACAACAGAATGATGAAGGGATTAGAC Amplicon expected: ~ 7-8 kb	Tm = 58.4°C Tm = 59.9°C Ta = 62.2°C
44	*VLDLR*	1	Primer Forward: CATATCATATATAGGGGAGTTGGCTTA Primer Reverse:ATTCCTGATCTTAGGGTGAGAGCAGTC Amplicon expected: ~ 6-8 kb	Tm = 58.4°C Tm = 61.3 °C Ta = 62.9 °C
46	*CPEB1*	1	Primer Forward: TAGCCATGATCATCAACTAGGACAACA Primer Reverse: CTTTTTCATGTCATGGGTTCCTCAG Amplicon expected: ~ 3-4 kb	Tm = 58.4°C Tm = 57.8°C Ta = 61.1°C

Another test tube containing only the reaction mix with no DNA was prepared as a negative control (C-). LR-PCR reactions has been performed using the *TaKaRa LA* Taq™ kit (TaKaRa), mixing the following reagents in a 0.2 ml tube:

10X LA PCR™ Buffer II (Mg^{2+} free)	5 µl
dNTPs Mixture (2.5 mM)	8 µl
Forward primer (25pmol/µl)	0.5 µl
Reverse primer (25pmol/µl)	0.5 µl
DNA	~100 ng
TaKaRa LA Taq™ (5U/µl)	0.5 µl
H_2O	up to 50 µl

The reactions are incubated in the iCycler thermocycler (Bio-Rad), with the following conditions (the time of elongation at 68°C from 1 to 8 minutes depends on length of the amplicon expected):

94°C	3 min	
94°C	30 sec	x 35 cycles
Ta °C (**Tab.3**)	30 sec	
68°C	1-8 min	
72°C	10 min	
4°C	∞	

At the end of the reaction, an agarose gel electrophoresis run (Agarose 0.8% in TAE 1X and Gel Red Nucleic Acid Stain) is carried out to assess the real presence of the amplification product, loading 5µl of reaction product mixed with 5 µl of Gel Loading Buffer 2X and a marker for nucleic acid size determination (StoSIkb DNA Ladder, GeneSpin). The PCR products will serve as a template for the subsequent sequencing reaction.

Purification and concentration of PCR products

Before going on with the sequencing reaction, it is necessary to purify and concentrate the DNA obtained from the PCR reaction using the *illustra GFX PCR DNA and Gel Band Purification Kit* (GE Haelthcare). This method uses the silica membrane contained in GFX MicroSpin™ column, which binds selectively the DNA. The procedure is the following:

- add 250 µl of Capture buffer type 3 to up to 50 µl of sample and mix thoroughly;
- check that the Capture buffer type 3-sample mix is yellow or pale orange in color to verify if the pH is optimal (pH ≤ 7.5) for the DNA binding to the silica membrane;
- for each purification, place one GFX MicroSpin column into one Collection tube;
- load the Capture buffer type 3-sample mix into the column and Collection tube;
- spin the assembled column and Collection tube at 16,000 X *g* for 30 seconds;

- discard the flow through by emptying the Collection tube, then place the GFX MicroSpin column back inside the same Collection tube;
- add 500 µl of Wash buffer type 1 to the column and spin at 16,000 X *g* for 60 seconds;
- discard the Collection tube and transfer the column to a fresh DNase-free 1.5 ml tube;
- add 10-50 µl of Elution buffer type 6 to the center of the membrane in the column;
- incubate at room temperature for 60 seconds;
- spin the assembled column and 1.5 ml tube at 16,000 X *g* for 60 seconds to recover the purified DNA. Store the purified DNA at -20°C.

Sanger sequencing reaction

The sequencing reactions were performed in a total volume of 20 µl using the *BigDye® Terminator v3.1 Cycle Sequencing Kit* (Applied Biosystem). The reaction is prepared in a 0.2 ml tube mixing the following reagents:

5X Sequencing Buffer	3.5 µl
PreMix v3.1	1.6 µl
Primer Forward or Reverse (3 pmol/µl) (**Tab.3,4**)	2 µl
Purified PCR sample	7 µl
H_2O	up to 20 µl

The reactions are incubated in the iCycler thermocycler (Bio-Rad), with the following conditions:

95°C	1 min	
95°C	10 sec	x 35 cycles
55°C	6 sec	
60°C	3 min	
4°C	∞	

Because of method's limits, other specific primers were chosen in order to read more than 600-700 bp of target sequence (**Tab.4**).

Tab.4. List of primers used for Sanger sequencing of the CNVs breakpoint junction.

Gene	Sequencing Probe	Primers
RYR3	1	LR-PCR probe 1
	2	Primer Forward: TTTAGTAGAGGCAGCATTTCATCATGT Primer Reverse: AGCAAATCCTCTACCCAGCTTTATTTGT
TP63	1	LR-PCR probe 1
	2	Primer Forward: ATGGGTGAGGAAAAACTCCA Primer Reverse: CAAGACCCTGCCCAGTCTAC

	3	Primer Forward: ACGCCCACTAATTTTTGCAC Primer Reverse: CAAAAATATTGGAAAGCCACTC
	4	Primer Forward: CACAAAAGGGCCAAGAGTGT Primer Reverse: GCAAAGCTAAAATGCTTCATCTG
VLDLR	1	LR-PCR probe 1
	2	Primer Forward: GCAGCTGAAGTCTCTTTTTCTTCCTCT Primer Reverse:CAGAGATTTGTCCTTTTGAATAAACATGG
CPEB1	1	LR-PCR probe 1
	2	Primer Forward: GCACCTACTTTGCAGACTTCG Primer Reverse:TTTGGGGACTCATCACTGGT

Purification of the sequencing product

The purification of the reaction products was carried out by using *CENTRI-SEP Columns kit* (Applied Biosystems). This step allows to remove the excess of DyeDeoxy™ terminators from completed DNA sequencing reactions. The procedure is the following:

- gently tap the column and ensure that the dry gel has settled in the bottom of the spin column;
- remove the top column cap and reconstitute the column by adding 800 µl of nuclease-free water, then replace the column cap and hydrate the gel by shaking and inverting the column or vortexing briefly;
- allow at least 30 minutes of room temperature hydration time before using the columns;
- remove air bubbles from the column gel by inverting the column;
- remove the top column cap and the column end stopper from the bottom, then transfer the column into a 2 ml wash tube;
- spin the assembled column and wash tube at 750 X g for 2 minutes to remove interstitial fluid;
- discard the wash tube and the interstitial fluid, then transfer the column into a new 1.5 ml collection tube;
- load 20 µL of sequencing reaction onto the center of the gel bed, without disturbing the gel surface;
- spin the assembled column and collection tube at 750 X g for 2 minutes, to collect the sample.

The purified sample can be stored at -20°C.

Sequencing data processing

The ABI PRISM 3500 Genetic Analyzer, used in the laboratory, exploits capillary electrophoresis as mechanism analysis. These devices use a polymer contained in a capillary that acts as a sieve to detect DNA molecules. The result of this analysis is a sequence represented by a series of colored peaks corresponding to each nucleotide: the electropherogram. Each sequence is analyzed by the ChromasPro program (version 1.5).

2.5 Gene expression evaluation on peripheral blood, commercial fetus and ovary RNAs

2.5.1 RNA extraction from peripheral blood

RNA from peripheral blood was collected using the *Tempus™ Blood RNA Tube* and isolated using the *Tempus™ Spin RNA Isolation Kit* (Applied Biosystem) according to manufacturer instructions:

- draw 3 ml of blood directly into Tempus Blood RNA Tube, which contains 6 ml of Stabilizing Reagent that lyses blood cells and inactivates cellular RNases, to make a total volume of 9 ml,
- shake the tube vigorously to ensure that the Stabilizing Reagent makes uniform contact with the sample; it is possible to store Tempus tubes containing stabilized samples for seven days at 4°C;
- put the contents of the tube into a clean 50 ml Falcon™ tube;
- add 3 ml of PBS 1X to each sample to bring the total volume to 12 ml;
- replace the cap on the tube and vortex vigorously for at least 30 seconds;
- centrifuge the tube at 4°C at 3,000 X *g* for 30 minutes;
- carefully pour off the supernatant and leave the tube inverted on absorbent paper for 1 to 2 minutes;
- add 400 µL of RNA Purification Resuspension Solution into the tube, then vortex briefly to resuspend the RNA pellet;
- keep the resuspended RNA on ice while preparing for the next steps;
- insert the filter into a waste collection tube and pre-wet the filtration membrane by adding 100 µl of RNA Purification Wash Solution 1;
- transfer the resuspended RNA (~400 µl) into the purification filter and centrifuge at 16000 X *g* for 30 seconds;
- remove the purification filter, discard the supernatant and re-insert the purification filter into the waste tube;
- add 500 µl of RNA Purification Wash Solution 1 into the purification filter, then centrifuge at 16000 X *g* for 30 seconds;
- remove the purification filter, discard the supernatant and re-insert the purification filter into the waste tube;
- add 500 µl of RNA Purification Wash Solution 2 into the purification filter, then centrifuge at 16000 X *g* for 60 seconds;
- remove the purification filter, discard the supernatant and re-insert the purification filter into the waste tube;

- perform a DNase treatment:

a) add 100 µl of AbsoluteRNA Wash Solution into the purification filter, then incubate at room temperature for 15 minutes;

b) add 500 µl of RNA Purification Wash Solution 2 into the purification filter, incubate at room temperature for 5 minutes, then centrifuge at 16000 X *g* for 30 seconds;

c) remove the purification filter, discard the supernatant and re-insert the purification filter into the waste tube.

- Add 500 µl of RNA Purification Wash Solution 2 into the purification filter, then centrifuge at 16000 X *g* for 30 seconds;

- remove the purification filter, discard the supernatant and re-insert the purification filter into the waste tube;

- centrifuge at 16000 X *g* for 30 seconds to dry the membrane;

- transfer the purification filter to a new collection tube;

- add 100 µl of Nucleic Acid Purification Elution Solution into the purification filter, incubate at 70°C for 2 minutes, then centrifuge at 16000 X *g* for 30 seconds;

- pipet the collected RNA eluate back into the purification filter and centrifuge at maximum speed (16,000 -18,000 X *g*) for 2 minutes;

- discard the purification filter, then transfer approximately 90 µl of the RNA eluate to a new collection tube. Extracted RNA can be stored at – 80°C for long-term storage.

The concentration and quality of the RNA are determined by a spectrophotometric analysis using the NanoDrop ND-1000 UV-VIS Spectrophotometer.

2.5.2 Reverse Transcription

To obtain cDNA the *High Capacity cDNA Reverse Transcription Kit* was used (Applied Biosystem). For each sample the following reaction mix using 500 ng of RNA is prepared in a 0.2 ml tube:

10X RT Buffer	2 µl
dNTP Mix 100 mM	0.8 µl
RT Random Primers 10X	2 µl
RNase Inhibitor	1 µl
MultiScribeTM Reverse Trascriptase	1 µl
RNA	500 ng
DEPC H_2O	up to 20 µl

The reaction is incubated in the iCycler thermocycler (Bio-Rad) and the reverse transcription conditions used are the following:

25°C	10 min
37°C	2 hours
85°C	5 min
4°C	∞

At the end of the reaction the cDNA is stored at -20°C.

2.5.3 Polymerase Chain Reaction (PCR)

With the purpose to assess the expression of the 37 genes (and 1 non-coding RNA) involved in the rare CNVs on RNA from peripheral blood, commercial fetus and ovary (*Human Fetus, Whole Poly A+ RNA* and *Human Ovary Poly A+ RNA*, Clontech), PCR reactions on cDNA has been performed by the *AmpliTaq Gold® kit* (Applied Biosystems).

The forward and reverse primers used were specific for the exon junctions in order to amplify only the cDNAs. The software Primer3 standard PCR conditions are reported below (**Tab.5**):

Primer size (n. of bases)→*Min 18; Opt 20; Max 27*

Primer Tm (°C) →*Min 57; Opt 6; Max 63*

The melting temperature (Tm) of each primer, the respective annealing temperature (Ta) and the PCR conditions, have been obtained by Optimase™ ProtocolWriter.

Tab.5. List of primers used for the PCR reaction on cDNA.

AMFR	Primer Forward Primer Reverse Amplicon	GCCCTCTTCGAGTGAGTGAG GATCCTTGCAGAGCTGAACC 180 bp	Tm = 58.3°C Tm = 56.3°C Ta = 60.3°C
BMP4	Primer Forward Primer Reverse Amplicon	TGAGCCTTTCCAGCAAGTTT GCGCAGCCCAAACATC 288 bp	Tm = 52.2°C Tm = 51.8°C Ta = 55.0°C
BNC1	Primer Forward Primer Reverse Amplicon	CCAGGTGGAGATTGTCCAGT CTCCAAAACGAAGGAACTGC 284 bp	Tm = 56.3°C Tm = 54.2°C Ta = 58.3°C
BRDT	Primer Forward Primer Reverse Amplicon	ATGCGAAGGCTTCAGAATGT TGCTCTTCTTGTGGCATCTG 151 bp	Tm = 52.2°C Tm = 54.2°C Ta = 56.2°C
BTBD1	Primer Forward Primer Reverse Amplicon	CAAAAGCACAATGGATGCAA CCTGCTGCAAATTCCTCAAT 252 bp	Tm = 50.1°C Tm = 52.2°C Ta = 54.2°C
CAPZA1	Primer Forward Primer Reverse Amplicon	TCGGATGAGGAGAAGGTACG AGGCGTGAACTGATCCATGT 168 bp	Tm = 56.3°C Tm = 54.2°C Ta = 58.3°C
CPEB1	Primer Forward Primer Reverse Amplicon	CCAAAGACCCCTTCAGCATA ATACCCTTTGGGAGGACACC 263 bp	Tm = 54.2°C Tm = 56.3°C Ta = 58.3°C

CYP2E1	Primer Forward	CGTGGAAATGGAGAAGGAAA	Tm = 52.2°C
	Primer Reverse	GGTGATGAACCGCTGAATCT	Tm = 54.2°C
	Amplicon	277 bp	Ta = 56.2°C
DDX1	Primer Forward	GAGGAATTGATATCCACGGTGT	Tm = 55.1°C
	Primer Reverse	CAGCCTCCATCTTCCTTGAG	Tm = 56.3°C
	Amplicon	217 bp	Ta = 58.7°C
DIAPH2	Primer Forward	ATGTGCGTGACCGAATTACA	Tm = 52.2°C
	Primer Reverse	CCGGATTGCTGGTTAAAGAA	Tm = 52.2°C
	Amplicon	423 bp	Ta = 55.2°C
DIAPH2-AS1	Primer Forward	CAGCCGGACTGTCTTGAAAC	Tm = 56.3°C
	Primer Reverse	TTGCATCGAGTTGTTATGTCTTC	Tm = 53.7°C
	Amplicon	300 bp	Ta = 58.0°C
EEF2	Primer Forward	AGCGAGGACAAGGACAAAGA	Tm = 54.2°C
	Primer Reverse	CCGAGAAGACTCGTCCAAAG	Tm = 56.3°C
	Amplicon	283 bp	Ta = 58.3°C
ELAVL2	Primer Forward	GAGCTTGGGATATGGCTTTG	Tm = 54.2°C
	Primer Reverse	GCTTCTTCTGCCTCAATTCG	Tm = 54.2°C
	Amplicon	312 bp	Ta = 57.2°C
FZD6	Primer Forward	TCGCAAATCTGGAATGTTCA	Tm = 50.1°C
	Primer Reverse	ATGCCTTGGACACCAAAATC	Tm = 52.2°C
	Amplicon	302 bp	Ta = 54.2°C
GRIA1	Primer Forward	CTGTGAATCAGAACGCCTCA	Tm = 54.2°C
	Primer Reverse	TTGGGTCTCTTCCAGTCCAC	Tm = 56.3°C
	Amplicon	252 bp	Ta = 58.3°C
HBE1	Primer Forward	GAATGTGGAAGAGGCTGGAG	Tm = 56.3°C
	Primer Reverse	ACCATCACGTTACCCAGGAG	Tm = 56.3°C
	Amplicon	279 bp	Ta = 59.3°C
IMMP2L	Primer Forward	TGCAGCCTTCTTTGAATCCT	Tm = 52.2°C
	Primer Reverse	CCGGTTTTTGTGTCCTATGG	Tm = 54.2°C
	Amplicon	197 bp	Ta = 56.2°C
LRIG2	Primer Forward	CCGCTATCCTGGATTTCAGT	Tm = 54.2°C
	Primer Reverse	CTGGAGTGCCTGTGCATTTA	Tm = 54.2°C
	Amplicon	194 bp	Ta = 57.2°C
MAP2K2	Primer Forward	AGGAAGCTGATCCACCTTGA	Tm = 54.2°C
	Primer Reverse	ATGTTGGAGGGCTTCACATC	Tm = 54.2°C
	Amplicon	302 bp	Ta = 57.2°C
MOV10	Primer Forward	GACCCCCAGTTCATAACCAA	Tm = 54.2°C
	Primer Reverse	TTACGCCGTGAAAGATGATG	Tm = 52.2°C
	Amplicon	184 bp	Ta = 56.2°C
MSL3	Primer Forward	TCTTAAAAGGCCTCCCCACT	Tm = 54.2°C
	Primer Reverse	GTTGGTCTGGCATGGAAGTT	Tm = 54.2°C
	Amplicon	281 bp	Ta = 57.2°C
MYCN	Primer Forward	CCCTGAGCGATTCAGATGAT	Tm = 54.2°C
	Primer Reverse	AGGATCAGCTCGCTGGACT	Tm = 55.9°C
	Amplicon	184 bp	Ta = 58.1°C
NRP1	Primer Forward	GAGAGAACAAGGTGTTCATGAGG	Tm = 57.2°C
	Primer Reverse	CGTGGAGAGAGCTGGAAAAG	Tm = 56.3°C
	Amplicon	161 bp	Ta = 59.8°C
PARD3	Primer Forward	ATTAATGATGGCGACCTTCG	Tm = 52.2°C
	Primer Reverse	AGGGCTAAAACGGCTTGAAT	Tm = 52.2°C
	Amplicon	174 bp	Ta = 55.2°C
PCSK5	Primer Forward	GCCAGCAAGTACGGATTCAT	Tm = 54.2°C
	Primer Reverse	ATCCCTCTTTGTCCGCTTTT	Tm = 52.2°C
	Amplicon	186 bp	Ta = 56.2°C
PRKAA1	Primer Forward	CCTCAAGCTTTTCAGGCATC	Tm = 54.2°C
	Primer Reverse	CGCCGACTTTCTTTTTCATC	Tm = 52.2°C
	Amplicon	198 bp	Ta = 56.2°C
RHOC	Primer Forward	AAGGATCAGTTTCCGGAGGT	Tm = 54.2°C
	Primer Reverse	ATGTTTTCCAGGCTGTCAGG	Tm = 54.2°C
	Amplicon	206 bp	Ta = 57.2°C

RREB1	Primer Forward	CAGAAGGAACCAGGAAACGA	Tm = 54.2°C
	Primer Reverse	TCAGTGACTTTCCGCAGATG	Tm = 54.2°C
	Amplicon	161 bp	Ta = 57.2°C
RUNX2	Primer Forward	GACAGCCCCAACTTCCTGT	Tm = 55.9°C
	Primer Reverse	CCGGAGCTCAGCAGAATAAT	Tm = 54.2°C
	Amplicon	159 bp	Ta = 58.1°C
RYR3	Primer Forward	CACAGGAGAAGCCTGTTGGT	Tm = 56.3°C
	Primer Reverse	GTGCTGGGAATCATTCTGGT	Tm = 54.2°C
	Amplicon	310 bp	Ta = 58.3°C
SH3GL3	Primer Forward	CTTACCTGGACCTCCCAACA	Tm = 56.3°C
	Primer Reverse	CATTGCCAAAGGTGGAGTCT	Tm = 54.2°C
	Amplicon	313 bp	Ta = 58.3°C
SIRT6	Primer Forward	GAGAGCTGAGGGACACCATC	Tm = 58.3°C
	Primer Reverse	CATGACCTCGTCAACGTAGC	Tm = 56.3°C
	Amplicon	251 bp	Ta = 60.3°C
SLC5A5	Primer Forward	GCGGGACTTTGCAGTACATT	Tm = 54.2°C
	Primer Reverse	GTCCAGACCACAGCCTTCAT	Tm = 56.3°C
	Amplicon	175 bp	Ta = 58.3°C
SLMAP	Primer Forward	TCAGAAATTGAGGCAAAGCA	Tm = 50.1°C
	Primer Reverse	AACCTTTGCTGCTGCTTGAT	Tm = 52.2°C
	Amplicon	312 bp	Ta = 54.2°C
SMYD3	Primer Forward	CAGGCAATCATAAGCAGCAA	Tm = 52.2°C
	Primer Reverse	TCTTCATTGCTTGGGGAAAC	Tm = 52.2°C
	Amplicon	247 bp	Ta = 55.2°C
SUPT3H	Primer Forward	TGGATCTGGCTCTTCTTGTG	Tm = 54.2°C
	Primer Reverse	TGTGAGCCTCAACACCACAG	Tm = 56.3°C
	Amplicon	135 bp	Ta = 58.3°C
SYCE1	Primer Forward	TTGCCAGGAAAAGGAAAGTG	Tm = 52.2°C
	Primer Reverse	TGATGCTTCACGTCTTCCAG	Tm = 54.2°C
	Amplicon	291 bp	Ta = 56.2°C
TM6SF1	Primer Forward	ACGCAATTTCAAGAGCCCTA	Tm = 52.2°C
	Primer Reverse	AACGATAGGCCAAGAGCTGA	Tm = 54.2°C
	Amplicon	310 bp	Ta = 56.2°C
TP63	Primer Forward	CCACCTTCGATGCTCTCTCT	Tm = 56.3°C
	Primer Reverse	GGATAACAGCTCCCTGAGGA	Tm = 56.3°C
	Amplicon	216 bp	Ta = 59.3°C
TSHZ2	Primer Forward	ACTCCAGCTGGGGATACAGA	Tm = 56.3°C
	Primer Reverse	TGGTAGCTGAAGCAGCCTTT	Tm = 54.2°C
	Amplicon	252 bp	Ta = 58.3°C
VLDLR	Primer Forward	CCAATTCCAGTGCACAAATG	Tm = 52.2°C
	Primer Reverse	TGAACCATCTTCGCAGTCAG	Tm = 54.2°C
	Amplicon	196 bp	Ta = 56.2°C
WHAMM	Primer Forward	GGCAACCATGTTCTTCCAGT	Tm = 54.2°C
	Primer Reverse	CGCATCCTGTTCCATTTCTT	Tm = 52.2°C
	Amplicon	262 bp	Ta = 56.2°C
XPO1	Primer Forward	CAGCAAAGAATGGCTCAAGA	Tm = 52.2°C
	Primer Reverse	TATTCCTTCGCACTGGTTCC	Tm = 54.2°C
	Amplicon	183 bp	Ta = 56.2°C
ZNF217	Primer Forward	GGAAGGTGGTTCTGAAGACG	Tm = 56.3°C
	Primer Reverse	CGCACTGTCAGCGGTTAATA	Tm = 54.2°C
	Amplicon	346 bp	Ta = 58.3°C

Another test tube containing only the reaction mix with no cDNA has been prepared as a negative control (*C-*). PCR reactions are set up in 0.2 ml sterile tubes mixing the following reagents:

10X PCR Gold Buffer	2.5 µl
25 mM $MgCl_2$ (2.5 mM)	2 µl
dNTPs (2 mM)	2.5 µl
Primer Forward (25 pmol/µl)	1 µl
Primer Reverse (25 pmol/µl)	1 µl
AmpliTaq Gold DNA Polymerase (5U/ µl)	0.25 µl
cDNA template	1 µl
H_2O	up to 25 µl

The reactions are incubated in the iCycler thermocycler (Bio-Rad), the time and temperature conditions used are the following:

95°C	5 min	
95°C	30 sec	
Ta °C (**Tab.5**)	30 sec	x 35 cycles
72°C	30 sec	
72°C	5 min	
4°C	∞	

At the end of the reaction, purification of PCR product is performed as previously described in paragraph 2.4.2. An agarose gel electrophoresis run (Agarose 2% in TAE 1X and Gel Red Nucleic Acid Stain) is then carried out to assess the real presence of the amplification product, loading 5µl of reaction product mixed with 5 µl of Gel Loading Buffer 2X and a marker for nucleic acid size determination (ΦX174 DNA/HaeIII 50 µg/µl, Invitrogen).

2.6 Transcript relative quantification analysis by Real Time RT-PCR

Only for 4 out of 67 patients (POI 6, 9, 10, 11), RNA samples were available. RNA extraction and retrotranscription reactions were carried out as explained in paragraph 2.5.1 and 2.5.2. To evaluate the presence of a quantitative alteration in the transcript levels of the interested genes *DIAPH2*, *TP63*, and *PRKAA1*, involved in rare CNVs, Real Time RT-PCR experiments were carried out using the universal protocol *TaqMan* (Applied Biosystem). Particularly, to investigate the transcript amount of the genes previously cited and of the genes used as *housekeeping* (*TBP*, *GUSB*, *RPLP0*, *HMBS*), commercial TaqMan probes, recommended for these genes and cDNA specific, were selected from the on-line shop at www.appliedbiosystems.com **(Tab.6 and 7)**.

Tab.6. List of the TaqMan assays of the target genes used for RT-qPCR analyses. The rearrangements involving the genes are summarized and the housekeeping genes used are shown.

Gene	Position	ID Assay	Amplicon lenght	Housekeeping gene
DIAPH2 *(NM_007309)* *Intronic gain between ex 26-27*	ex 15-16	Hs00246501_m1	62 bp	*TBP, GUSB*, RPLP0*
	ex 26-27	Hs00963595_g1	65 bp	*TBP, GUSB*, RPLP0*
TP63 *(NM_003722)* *Intragenic gain from ex 2 to 10*	ex 7-8	Hs00978343_m1	85 bp	*TBP, GUSB, RPLP0**
	ex 11-12	Hs00978347_m1	76 bp	*TBP, GUSB, RPLP0**
PRKAA1 *(NM_206907)* *Intronic loss between ex 1-2*	ex 1-2	Hs01562307_mH	60 bp	*GUSB, RPLP0, HMBS*
	ex 9-10	Hs01562315_m1	118 bp	*GUSB, RPLP0, HMBS*

*the normalization graphical view is shown in results.

Tab.7. List of the TaqMan assays of the housekeeping genes used for the RT-qPCR analyses.

Gene	ID Assay	Amplicon lenght
HMBS	Hs00609297_m1	64 bp
GUSB	Hs00939627_m1	96 bp
RPLP0	Hs99999902_m1	105 bp
TBP	Hs00427620_m1	91 bp

Two TaqMan assays were chosen for each genes, one "outside" the rearrangements, and the other "inside" them **(Tab.6)**. The experiments were performed on the patients and 10 healthy controls (both male and female) without any CNVs involving the target genes; one

retrotranscription of about 500 ng for each sample was made and the reactions were performed in triplicate.

The reactions have been set up in a 96-plate, preparing a Mix 1 for each genes, and a Mix 2 for each retrotranscription in a final volume of 10 µl:

Mix 1		Mix 2	
Master Mix (Applied Biosystem) 2X	5 µl	DNA	2.5 µl
Assay TaqMan 20X	0.5 µl	H_2O	1 µl
H_2O	1 µl	Total volume	3.5 µl
Total volume	6.5 µl		

Each experiment was replicated three times using different *housekeeping* genes. The details are shown in **Tab.6**. The reactions were performed using the ABI Prism 7900 (Applied Biosystem) with the SDS software v2.3, and the PCR conditions used are the following:

50 °C	2 min	
95 °C	10 min	
95 °C	15 sec	x 40 cycles
60 °C	1 min	

At the end of Real Time RT-PCR, the experiment is analyzed calculating the "threshold cycles" of each PCR by the program RQ Manager 1.2 (Applied Biosystem). For each assay the transcript amount of the target genes in POI women and in healthy controls was determined using the relative quantification method of $2^{-\Delta\Delta Ct}$, and the housekeeping genes as normalizers.

In these analyses the standard deviation values greater than 0.5 were excluded and the results obtained were analyzed by Student t-test with the aim to investigate a statistically significant value relative to the transcript amount of the patients compared to controls ($p< 0.05$).

2.6.1 *VLDLR*: position effect assessment

RT-qPCR with Syber Green assays were performed in POI 44 using *GAPDH* as normalizing gene. PCR and Sanger sequencing were used for heterozygosity analysis on *VLDLR*. Luciferase assay using pGL3 vector (*SV40* promoter) was carried out both on HeLa and COV434 cell lines, whereas pGL3b vector (*VLDLR* promoter) was transfected only in COV434 cell line. The analyses were performed thanks to a collaboration with Dott.ssa Caslini Cecilia, Department of Medical Biotechnology and Translational Medicine.

2.7 Whole Exome Sequencing preliminary analysis

Thanks to a collaboration with the Genetics of common disorders Lab. at San Raffaele Hospital, WES analysis was carried out for 17 out of 67 POI women. The raw data (Fastq files) obtained were processed through a in-house pipeline. Briefly, using BWA software, the Fastq data of each sample were concatenated, indexed, and aligned to a reference genome (assembling hg19) generating a SAM file. Then, using SAMtools, the data were sorted, indexed and PCR duplicated were excluded; finally total BAM files were created. Quality control and estimation of bad coverage were investigated using FastQC and BEDtools, respectively. Variant calling of each exome analysis was performed using GATK [169], and VCF files were created.

The preliminary analysis performed consists in a filtering based on 191 genes composed by:

- 41 genes derived from array CGH results;
- 135 genes implicated in non-syndromic form of POI, or in syndromic POI, or in ovarian growth and development. Namely, the selection of most of them was performed considering the supplementary tables S14, and S15 of the paper by Day Felix [155], which includes genes associated to POI and ovarian genes, respectively;
- 15 genes known to be interactors of *VLDLR*, *TP63*, *CPEB1* (data obtained from STRING analysis).

A coordinate BED file related to the target genes was created. Table browser of UCSC was used and only exonic sequences were considered.

VCF files of each patients were intersected with the BED file created using BEDtools and a new targeted VCF file was obtained. Then, annotation of the variants was carried out using wANNOVAR (http://wannovar.usc.edu/index.php) on-line software and a CSV file for each patient was created.

To prioritize rare variants with potentially pathogenetic role in POI, several filtering steps were carried out as shown in **Fig.6.** In Step 1, irrelevant variants were excluded, namely synonymous and variants mapping in repeated regions (e.g. locus AR), whereas in Step 2 variants covered by more than 20 reads were considered. Based on the information provided by the 1000 Genomes Project (http://1000genomes.org/) and ExAC Variant Server (http://exac.broadinstitute.org/), variants falling above > 0.5% in the 1000G Project (Step 3) were considered not rare and were filtered out as well as ExAC variants exceeding 5% (Step 4).

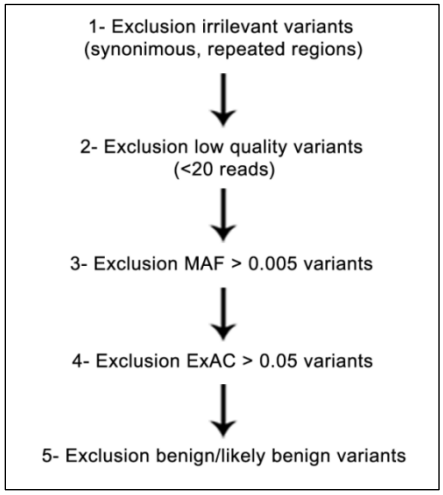

Fig.6. *Flowchart used for WES data analysis.*

The obtained variants were then analyzed using several databases to establish their potential pathogenicity:

- dbSNP (http://www.ncbi.nlm.nih.gov/SNP/),
- Ensembl (http://www.ensembl.org/index.html),
- SIFT (http://sift.bii.a-star.edu.sg/)[170],
- PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/)[171],
- MutationTaster (http://www.mutationtaster.org/)[172],
- ClinVar (http://www.ncbi.nlm.nih.gov/clinvar/),
- HGMD (http://www.hgmd.cf.ac.uk/ac/index.php).

The variants with almost two of the following features were excluded because considered benign/likely benign (Step 5).

- unreported in HGMD,
- reported as benign, probably non pathogenic in ClinVar (when available),
- predicted benign/likely benign in almost two out of the three prediction softwares used.

2.7.1 Sanger sequencing validation (in progress)

As final step, the direct sequencing using the classical Sanger method was performed to validate the variants identified (16 SNVs have been analyzed till now; **Tab.8**). PCR and Sanger sequencing reactions, as well as purification steps have been carried out as described in paragraph 2.4.2.

Tab.8. List of primers used for the PCR and sequencing reactions.

ID POI	Gene	Primers	PCR Conditions
49	*SIRT6*	Primer Forward: ACCATGCACCCTCTGTCC Primer Reverse: CTTGGCCTTCACCCTTTTG	Tm = 55.4°C Tm = 53.7°C

		Amplicon: 350 bp	Ta =57.6°C
50	*POF1B*	Primer Forward:TTCCTAAAATATGTCCCATTACACAA Primer Reverse: AGATTTTAATTTGGGACTTTCTGTTC Amplicon: 270 bp	Tm = 53.3°C Tm = 53.3°C Ta =56.3°C
	RAD51C	Primer Forward: TCTGTTGCCTTGGGGAGTAT Primer Reverse: GGTCTCAGATGGGCACAAAT Amplicon: 370 bp	Tm = 54.2°C Tm = 54.2°C Ta =57.2°C
	WHAMM	Primer Forward: TTGAGAACATGAAGGTTTCTTTT Primer Reverse: TGGTAAACCCAAACAAAGTACAAA Amplicon: 379 bp	Tm = 50.1°C Tm = 52.4°C Ta =54.3°C
54	*RREB1*	Primer Forward: TGAACTGACTTCTTTGTTAGATCACAT Primer Reverse: GGGAAAACACACGTTAAGAGC Amplicon: 323 bp	Tm = 55.2°C Tm = 54.7°C Ta =58.0°C
	PCSK5	Primer Forward:CCTGACCACCATGTTTTTCC Primer Reverse:AGTCGGATGGCATCTCTCAC Amplicon: 282 bp	Tm = 54.2°C Tm = 56.3°C Ta =58.3°C
55	*HFM1*	Primer Forward: CCTCATCATTAAATATTGTCAGAATC Primer Reverse: GAGTGTGGCAGTGCAGACAT Amplicon: 288 bp	Tm = 53.3°C Tm = 56.3°C Ta =57.8°C
	ADAMTS19	Primer Forward: ATCTCCATGTATGGGCCTTG Primer Reverse: CGGGGCATTAACCAAAACTA Amplicon: 320 bp	Tm = 54.2°C Tm = 52.2°C Ta =56.2°C
	SH3GL3	Primer Froward: GCTAGGTCTTTGCTGCTGGT Primer Reverse: GAAGGGGGAGATGGAGAGAT Amplicon: 286 bp	Tm = 56.3°C Tm = 56.3°C Ta=59.3°C
	RELN	Primer Forward: TGTGATGCCATGATAGGAAGA Primer Reverse: AGGTGGAGCAATCCAGATGA Amplicon: 177 bp	Tm = 52.7°C Tm = 54.2°C Ta =56.5°C
	POU5F1	Primer Forward: CGCCGTATGAGTTCTGTGG Primer Reverse: AGGTGACCACTTCCCCATC Amplicon: 288 bp	Tm = 55.9°C Tm = 55.9°C Ta =58.9°C
57	*RYR3*	PrimerForward: CCAACTTCTCTACATCCTCAACC Primer Reverse: CACCAACCCCCTATTTTCAG Amplicon: 354 bp	Tm = 57.2°C Tm = 54.2°C Ta =58.7°C
60	*FMR1*	Primer Forward: ACCTGCTGTTCACCAGATCC Primer Reverse: CAGTGTTTTGCACAGGCTAGA Amplicon: 300 bp	Tm = 56.3°C Tm = 54.7°C Ta =58.5°C
63	*CYP19A1*	Primer Forward: CACATTGCATTTGGAGCAAC Primer Reverse: AAAAGGCACATTCATAGACAAAAA Amplicon: 324 bp	Tm = 52.2°C Tm =50.7°C Ta =54.5°C
65	*GALT*	Primer Forward: GGGTTTCTTGGCTGAGTCTG Primer Reverse: TGCTAAGGCCTCCTAGCAAGT Amplicon: 386 bp	Tm = 56.3°C Tm = 56.6°C Ta =59.5°C
67	*VLDLR*	Primer Forward: TTCTAGCATGGCATGTTCAGTT Primer Reverse: AGAGAAACAAAGAAGAGAGATTCCA Amplicon: 290 bp	Tm = 53.2°C Tm = 54.8°C Ta =57.0°C

Furthermore this approach will be necessary to: i) cover the regions with low capture efficiency or bad sequencing coverage (< 20); ii) confirm they are not fakes emerged from sequencing or analysis irregularities.

3. RESULTS

3.1 Identification and classification of CNVs

A cohort of 67 POI women was analyzed by high resolution array CGH to identify CNVs possibly implicated in POI pathogenesis. In 58 out of 67 patients (86.6%), rare CNVs (one or more) not yet reported in healthy subjects according to the DGV were detected. The group consists of:

- 46 sporadic (79.3%) and 12 familial (20.7%) cases, including five sisters pairs, and two patients with a relative affected with POF (not included in the analyzed cohort).
- 7 patients analyzed by array CGH 244K and 51 by higher resolution array CGH 400K.

The detection rate for rare CNVs using the genome wide analysis was approximately 77.8% and 87.9% for the patients processed with 244K and 400K, respectively.
Overall 101 rare CNVs were detected, ranging from 5 kb to 1.7 Mb, including 45 gains (44.6%) and 56 losses (55.4%). Among them, 4 out of 101, 2 duplications and 2 deletions, were homozygous CNVs. The inheritance for 81 CNVs (80.2%) is unknown. None of the remaining 20 CNVs was *de novo*, being all inherited, 12 from the father (60%) and 7 from the mother (35%). In one patient the homozygous CNV was inherited from both the consanguineous parents (5%).
A detailed list of the identified rare CNVs is shown in **Tab.9**.

Furthermore, high-resolution array CGH analysis allowed to identify in a subset of the POI cohort a few CNVs already reported in the DGV. Nevertheless these are "common" CNVs containing genes previously associated with the disorder (*NAIP*, *DUSP22*, *SYCE1*) [152,161], that were selected as putative loci for POI susceptibility. In detail:

- 5 CNVs involving *NAIP* at 5q13.2 (1 duplication and 4 deletions),
- 16 CNVs involving *DUSP22* at 6p25.2 (5 duplications and 11 deletions),
- 5 CNVs involving *SYCE1*, *CYP2E1* at 10q26.3 (4 duplications and 1 deletion) (**Tab.10**).

Tab.9. Rare CNVs identified in the POI cohort by means of array CGH analysis

POI ID	Chromosomal Band	Gain/ Loss	Het/ Hom	Size	N° of genes	Gene(s)	Nucleotide position[#]	Inheritance
Array-CGH 244K								
1	4q32.1	Gain	het	43 kb	1	*GLRB*	chr4:158085813-158129097	NA
	7q11.22	Loss	het	28 kb	1	*CALN1*	chr7:71826130-71854571	NA
3	15q14	Loss	het	26 kb	1	*RYR3*	chr15:34125784-34151995	NA
4	11q22.3	Loss	het	221 kb	-	*(GRIA4)*	chr11:105126944-105347752	NA
6	Xq21.33	Gain	het	53 kb	2	*DIAPH2, DIAPH2-AS1*	chrX:96774546-96827670	Pat
7	6q24.1	Loss	het	93 kb	-	-	chr6:140576639-140669614	NA
	9p21.3	Loss	het	68 kb	-	*(ELAVL2)*	chr9:24242825-24310533	NA
	9p21.3	Loss	het	79 kb	-	-	chr9:24426310-24505824	NA
8*	2q24.3	Loss	het	6.5 kb	1	*XIRP2*	chr2:168116199-168122776	Pat
	2q32.3	Loss	het	20.5 kb	1	*SDPR*	chr2:192679673-192700138	Pat
	3q28	Gain	het	158.5 kb	1	*TP63*	chr3:189427139-189585621	Pat
	9p22.2	Loss	het	59 kb	1	*CNTLN*	chr9:17298850-17357754	Pat
9*	2q37.2	Loss	het	13 kb	1	*AGAP1*	chr2:236489321-236502202	Mat
	5p13.1	Loss	het	6.5 kb	1	*PRKAA1*	chr5:40779232-40785703	Pat
Array-CGH 400K								
10*	2q24.3	Loss	het	6.5 kb	1	*XIRP2*	chr2:168116199-168122776	Pat
	3q28	Gain	het	161 kb	1	*TP63*	chr3:189427139-189588477	Pat
	7p14.1	Gain	het	144 kb	1	*CDK13*	chr7:39999692-40144263	Mat
	9p22.2	Loss	het	69 kb	1	*CNTLN*	chr9:17295644-17364986	Pat
11*	2q37.2	Loss	het	17 kb	1	*AGAP1*	chr2:236484897-236502202	Mat
	5p13.1	Loss	het	6.5 kb	1	*PRKAA1*	chr5:40779232-40785703	Pat
	18q21.33	Loss	het	9 kb	-	*(ZCCHC2, TNFRSF11A)*	chr18:60098813-60108560	Pat
	Xp22.33	Gain	het	59 kb	3	*ASMTL, P2RY8, CRLF2*	chrX:1553947-1612474	Pat
12	4p15.1	Loss	het	231 kb	-	*(PCDH7)*	chr4:31744797-31975862	NA

Tab.9. Continued

	6p25.1-24.3	Gain	het	36 kb	1	*RREB1*	chr6:7075538-7111544	NA
13	1q44	Gain	het	523.5 kb	1	*SMYD3*	chr1:246168808-246692300	NA
	1q44	Gain	het	321 kb	5	*AHCTF1, ZNF695, ZNF670, ZNF669, ZNF124*	chr1:247088226-247409060	NA
	11p15.4	Gain	het	42 kb	3	*HBG2, HBE1, OR51B5*	chr11:5454730-5496463	NA
14	Xp11.22	Loss	hom	36 kb	1	*BMP15*	chrX:50653777-50689335	Mat, Pat[§]
15*	15q26.1	Loss	het	58 kb	1	*ABHD2*	chr15:89595324-89652892	Mat
16*	8q21.3	Gain	het	25 kb	1	*TMEM64*	chr8:91704940-91730316	Mat[¥]
17	9p13.3	Loss	het	12 kb	1	*KIF24*	chr9:34280912-34293051	NA
18	14q22.2	Loss	het	20 kb	-	*(BMP4)*	chr14:54195768-54215968	NA
19	Xq28	Loss	het	106 kb	-	*(GABRA3, GABRQ)*	chrX:151638056-151743864	NA
20*	7p21.3	Loss	het	19 kb	-	-	chr7:10795079-10814265	Mat
21*	7q22.3	Loss	het	24 kb	-	-	chr7:107262480-107286241	Mat
23	8q13.3	Loss	het	31 kb	1	*XKR9*	chr8:71613753-71645051	NA
	9q32	Loss	het	13 kb	-	-	chr9:115698459-115711373	NA
24	21q21.1	Loss	het	32 kb	1	*NCAM2*	chr21:22418118-22450219	NA
25	1p13.2	Gain	het	91 kb	-	*(MOV10, RHOC, PPM1J, FAM19A3,CAPZA1, SLC16A1, LRIG2)*	chr1:113339006-113430121	NA
	1q31.3	Gain	het	310 kb	-	-	chr1:194012929-194322620	NA
26	5q34	Gain	het	772.5 kb	2	*GABRB2, ATP10B*	chr5:160019212-160791695	NA
27	15q21.1	Gain	het	41 kb	1	*SORD*	chr15:45319178-45360400	NA
	Xp11.21	Loss	het	5 kb	-	-	chrX:56093719-56098999	NA
	Xq13.3	Loss	het	54 kb	-	-	chrX:74772380-74826319	NA
28	2p22.3p22.2	Gain	hom	23 kb	-	*(CRIM1)*	chr2:35967990-35990844	NA
	17q12	Loss	het	16 kb	1	*ASIC2*	chr17:32130833-32146606	NA
29	5q22.2	Loss	het	13 kb	1	*APC*	chr5:112179040-112192379	NA
30	4q26	Gain	het	22 kb	-	*(NDST4)*	chr4:115711838-115734119	NA
31	3p23	Gain	het	447 kb	1	*GADL1*	chr3:30914366-31361404	NA
33	8p23.1	Loss	het	101 kb	1	*ERI1*	chr8:8790358-8891761	NA

Tab.9. Continued

	9q21.13	Gain	het	542 kb	1	*PCSK5*	chr9:78117874-78659948	NA
	12q12	Gain	het	72.5 kb	2	*GXYLT1, YAF2*	chr12:42521259-42593754	NA
34	9p21.3	Loss	het	17 kb	-	-	chr9:21678040-21695476	NA
35	1q32.1	Gain	het	67 kb	1	*GPR37L1*	chr1:202027010-202093734	NA
	21q21.1	Loss	het	5 kb	1	*CHODL*	chr21:19349121-19354449	NA
38*	19p13.11	Gain	het	61 kb	1	*SLC5A5*	chr19:17981993-18043097	NA
39*	2p24.3	Gain	het	1.1 Mb	3	*NBAS, DDX1, MYCN*	chr2:15091702-16172000	NA
	19p13.11	Gain	het	61 kb	1	*SLC5A5*	chr19:17981993-18043097	NA
40	2q21.2	Loss	het	401 kb	1	*NCKAP5*	chr2:133725047-134126372	NA
	2q32.1	Gain	hom	193 kb	2	*DUSP19, NUP35, (NCKAP1)*	chr2:183919175-184112630	NA
41*	19p12	Loss	het	429.5 kb	3	*ZNF675, ZNF681, RPSAP58*	chr19:23624728-24054225	NA
42	2p15	Loss	het	36 kb	1	*(XPO1)*	chr2:61795359-61831792	NA
43	1p34.1	Gain	het	279 kb	4	*GPBP1L1, TMEM69, IPP, MAST2*	chr1:46116441-46395389	NA
	Xp22.11	Loss	het	254 kb	1	*PTCHD1*	chrX:23209046-23463336	NA
44	9p24.2	Loss	het	14.5 kb	1	*VLDLR*	chr9:2654203-2668811	NA
45	3q11.2	Loss	het	134 kb	-	-	chr3:94689602-94823543	NA
	10p11.22	Gain	het	97.5 kb	-	*(PARD3, NRP1)*	chr10:34134992-34232568	NA
46*	15q25.2	Loss	hom	9 kb	1	*CPEB1*	chr15:83305533-83314278	NA
	16q12.2	Gain	het	25 kb	1	*AMFR*	chr16:56436830-56461987	NA
47	-	-	-	-	-	-	-	-
48	3p14.2	Gain	het	1 Mb	1	*FHIT*	chr3:59180080-60204731	NA
	5q33.1	Loss	het	58 kb	-	*(GRIA1)*	chr5:152263564-152321761	NA
50	2q24.1	Loss	het	108 kb	1	*CCDC148*	chr2:159142708-159250404	NA
	Xp22.31	Gain	het	1.7 Mb	3	*STS, HDHD1, PNPLA4*	chrX:6456036-8133172	NA
52	2p22.3	Gain	het	7 kb	-	-	chr2:35410756-35418046	NA
	4p15.31	Loss	het	6 kb	1	*KCNIP4*	chr4:20748440-20754466	NA
	7q31.33	Loss	het	18 kb	-	-	chr7:125886329-125904559	NA
	10q26.2	Loss	het	15 kb	1	*ADAM12*	chr10:127997316-128012550	NA
	20q13.2	Loss	het	114 kb	-	*(TSHZ2)*	chr20:51258881-51372398	NA

Tab.9. Continued

53	6q27	Gain	het	257 kb	3	*FAM120B*, *PSMB1*, *TBP*	chr6:170609029-170866001	NA
	8q22.3	Loss	het	28 kb	1	*FZD6*	chr8:104306192-104334440	NA
54	3p14.3	Gain	het	118 kb	2	*DENND6A*, *SLMAP*	chr3:57671881-57790035	NA
	15q25.2	Loss	het	1.6 Mb	15	*CPEB1*, *FAM103A1 C15orf40*, *EFTUD1P1*, *AP3B2*, *SCARNA15*, *FSD2*, *WHAMM*, *HOMER2*, *BTBD1*, *TM6SF1*, *HDGFRP3*, *BNC1*, *SH3GL3*, *ADAMTSL3*	chr15:83214012-84812693	NA
	21q11.2	Loss	het	10 kb	1	*ABCC13*	chr21:15666538-15676565	NA
55	4q22.1	Gain	het	258 kb	1	*CCSER1*	chr4:91157649-91415956	NA
	8p22	Loss	het	86 kb	1	*DLC1*	chr8:13034651-13120326	NA
	22q13.33	Gain	het	87 kb	2	*ARSA*, *SHANK3*	chr22:51059148-51146462	NA
	Xq21.33	Loss	het	26 kb	-	*(DIAPH2)*	chrX:97984949-98010783	NA
56	8p21.2	Gain	het	41 kb	2	*NKX3-1*, *NKX2-6*	chr8:23528632-23569976	NA
57	1q23.1	Loss	het	295 kb	1	*CD1E*	chr1:158323371-158352947	NA
58	3q26.31	Gain	het	25 kb	1	*NAALADL2*	chr3:175345151-175370252	NA
59	6q24.1	Loss	het	28 kb	-	-	chr6:140376204-140403723	NA
	13q22.2	Loss	het	13 kb	1	*LMO7*	chr13:76256435-76269258	NA
60	Xp22.2	Gain	het	747 kb	3	*ARHGAP6*, *AMELX*, *MSL3*	chrX:11249664-11996670	NA
61	11q25	Gain	het	147 kb	-	-	chr11:134623571-134770989	NA
62	20q13.2	Gain	het	21.5 kb	1	*ZNF217*, *(TSHZ2)*	chr20:52163628-52185266	NA
63	19p13.3	Gain	het	885 kb	32	*TJP3*, *APBA3*, *MRPL54*, *RAX2*, *MATK*, *ZFR2*, *ATCAY*, *NMRK2*, *DAPK3*, *SNORD37*, *EEF2*, *PIAS4*, *ZBTB7A*, *MAP2K2*, *CREB3L3*, *SIRT6*, *ANKRD24*, *EBI3*, *CCDC94*, *SHD*, *TMIGD2*, *FSD1*, *STAP2*, *MPND*, *SH3GL1*, *UBXN6*, *CHAF1A*, *HDGFRP2*, *PLIN4*, *PLIN5*, *LRG1*, *SEMA6B*	chr19:3718839-4604407	NA
64	13q31.2	Gain	het	41 kb	-	-	chr13:88282085-88323161	NA
65	10p15.3	Gain	het	108 kb	2	*WDR37*, *ADARB2*	chr10:1175097-1282960	NA

Tab.9. Continued

66	2p14	Loss	het	18 kb	1	*MEIS1*	chr2:66714797-66732973	NA
	3q21.3	Gain	het	102.5 kb	2	*MGLL, KBTBD12*	chr3:127533289-127635789	NA
67	7q33	Gain	het	11 kb	1	*CALD1*	chr7:134525335-134536040	NA

[#]Physical position of the identified rare CNVs based on the UCSC Genome Browser, hg19, released February 2009.
*****Familial cases: 8-10, 9-11, 15-16, 20-21, 38-39, couple of sisters; 41 and 46 present a relative with SA who are not included in the analyzed cohort.
[$]Parents of POI 14 are first cousins.
[¥]The CNV is present also in the healthy sister.
() Genes potentially perturbed by rare CNVs through a position effect.

Tab.10. Common CNVs identified in the POI cohort by means of array CGH analysis

POI ID	Gain/Loss	Size	Genes/Locus	Nucleotide position[#]	POI ID	Gain/Loss	Size	Genes/Locus	Nucleotide position[#]
1[$]	Loss	578 kb	*NAIP*	chr5:69767948-70345611	**39***	Loss	119 kb	*DUSP22*	chr6:259881-378956
2[$]	Loss	92 kb	*DUSP22*	chr6:259528-293493	**42**	Loss	119 kb	*DUSP22*	chr6:259881-378956
4[$]	Loss	681 kb	*NAIP*	chr5:69741318-70422356	**44**	Gain	80 kb	*DUSP22*	chr6:299363-378956
6*[$]	Loss	92 kb	*DUSP22*	chr6:283968-375949	**48**	Loss	350 kb	*NAIP*	chr5:70308101-70657747
9*[$]	Gain	96 kb	*CYP2E1, SYCE1*	chr10:135281682-135377532	**50**	Gain	80 kb	*DUSP22*	chr6:299363-378956
11*	Gain	97 kb	*CYP2E1, SYCE1*	chr10:135281682-135378761	**52**	Gain	1.8 Mb	*NAIP*	chr5:68849594-70657747
13	Gain	119 kb	*DUSP22*	chr6:259881-378956	**53**	Loss	119 kb	*DUSP22*	chr6:259881-378956
17	Gain	126 kb	*CYP2E1, SYCE1*	chr10:135252327-135378761	**56**	Loss	119 kb	*DUSP22*	chr6:259881-378956
20*	Loss	1.8 Mb	*NAIP*	chr5:68818173-70657747	**57**	Loss	119 kb	*DUSP22*	chr6:259881-378956
	Loss	130 kb	*DUSP22*	chr6:259881-389482	**59**	Gain	80 kb	*DUSP22*	chr6:299363-378956
26	Gain	43 kb	*DUSP22*	chr6:259881-303298	**60**	Loss	126 kb	*CYP2E1, SYCE1*	chr10:135252327-135378761
33	Loss	119 kb	*DUSP22*	chr6:259881-378956	**62**	Gain	123 kb	*CYP2E1, SYCE1*	chr10:135281682-135404523
36*	Loss	99 kb	*DUSP22*	chr6:279667-378956	**65**	Loss	389 kb	*DUSP22*	chr6:259881-389482

[#]Physical position of the identified rare CNVs based on the UCSC Genome Browser, hg19, released February 2009.
*****Familial cases.
[$]POI patients analyzed by array CGH 244K.

3.2 Analysis of the gene content and validation of the ovarian rare CNVs

Based on the gene content, 33 rare CNVs were selected as possibly related to POI pathogenesis in the group of 58 patients positive to the array CGH analysis. In 10 CNVs qPCR or Long-Range PCR validation was performed and only 1 CNV could not be confirmed (**Tab.11**).

Tab.11. List of the ovarian rare CNVs with their molecular validation.

POI ID	Chromosomal Band	Type of rearrangement	Size	Ovarian Gene(s)	CNV validation
3	15q14	Intragenic deletion	26 kb	*RYR3*	LR-PCR
6	Xq21.33	Intronic duplication	53 kb	*DIAPH2, DIAPH2-AS1*	Familial
7	9p21.3	Position effect	68 kb	*ELAVL2*	qPCR
8*	3q28	Intragenic duplication	158.5 kb	*TP63*	Familial
9*	5p13.1	Intronic deletion	6.5 kb	*PRKAA1*	Familial, qPCR
10*	3q28	Intragenic duplication	161 kb	*TP63*	Familial
11*	5p13.1	Intronic deletion	6.5 kb	*PRKAA1*	Familial, qPCR
12	6p25.1-24.3	Partial gene duplication	36 kb	*RREB1*	spots › 5
13	1q44	Partial gene duplication	523.5 kb	*SMYD3*	spots › 5
	11p15.4	Intronic duplication	42 kb	*HBE1*	spots › 5
14	Xp11.22	Complete gene deletion	36 kb	*BMP15*	Familial
18	14q22.2	Position effect	20 kb	*BMP4*	qPCR
25	1p13.2	Position effect	91 kb	*MOV10, RHOC, CAPZA1, LRIG2*	spots › 5
33	9q21.13	Partial gene duplication	542 kb	*PCSK5*	spots › 5
38*	19p13.11	Complete gene duplication	61 kb	*SLC5A5*	spots › 5
39*	2p24.3	Complete gene duplication	1.08 Mb	*DDX1, MYCN*	spots › 5
	19p13.11	Complete gene duplication	61 kb	*SLC5A5*	spots › 5
42	2p15	Position effect	36 kb	*XPO1*	spots › 5
44	9p24.2	Partial gene deletion	14.5 kb	*VLDLR*	LR-PCR
45	10p11.22	Position effect	97.5 kb	*PARD3, NRP1*	spots › 5
46*	15q25.2	Intronic deletion[#]	9 kb	*CPEB1*	qPCR, LR-PCR
	16q12.2	Partial gene duplication	25 kb	*AMFR*	spots › 5
48	5q33.1	Position effect	58 kb	*GRIA1*	spots › 5
50	Xp22.31	Complete gene duplication	1.7 Mb	*STS*	spots › 5
52	20q13.2	Position effect	114 kb	*TSHZ2*	spots › 5
53	8q22.3	Partial gene deletion	28 kb	*FZD6*	spots › 5
54	3p14.3	Partial gene duplication	118 kb	*SLMAP*	spots › 5
	15q25.2	Complete gene deletion	1.6 Mb	*CPEB1, WHAMM, BTBD1, BNC1, SH3GL3*	spots › 5
55	Xq21.33	Position effect	26 kb	*DIAPH2*	qPCR
60	Xp22.2	Complete gene duplication	747 kb	*MSL3*	spots › 5
62	20q13.2	Position effect	21.5 kb	*TSHZ2*	qPCR
63	19p13.3	Complete gene duplication	885 kb	*EEF2, MAP2K2, SIRT6*	spots › 5
66	2p14	Intronic deletion	18 kb	*MEIS1*	-

***** Familial cases; [#] Homozygous deletion; -, not validated.

Overall, 32 ovarian CNVs were selected in 28 POI patients attesting a detection rate of 31.7%. These CNVs include 18 duplications (56.2%) and 14 deletions (43.8%), ranging from 6.5 kb to 1.7 Mb. According to the type of rearrangement affecting the involved the genes, the CNVs were classified as shown in **Fig.7**.

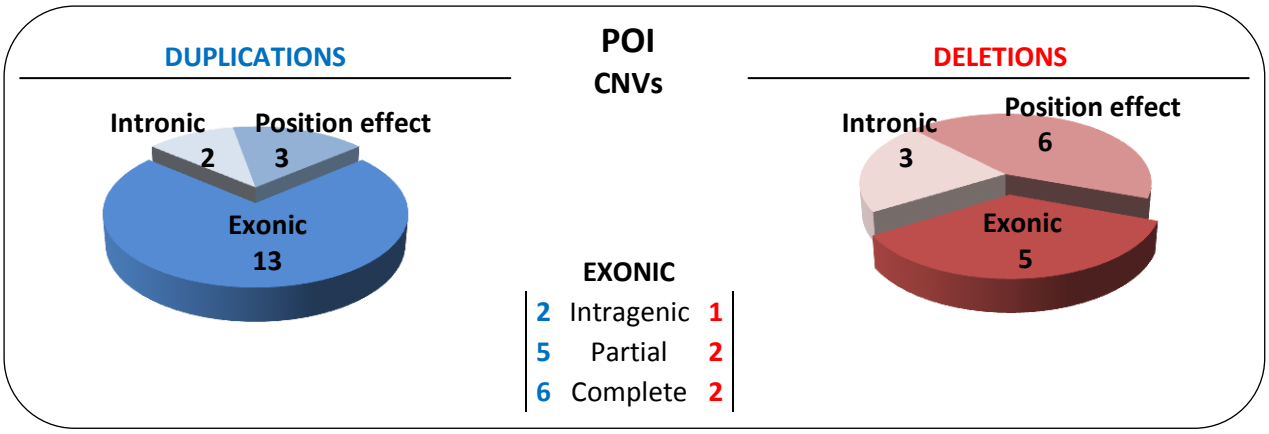

Fig.7. Distribution of the 32 ovarian CNVs. *13 gains and 5 losses involve the genes coding region resulting in intragenic, partial and complete deletion/duplication; 2 gains and 3 losses involve intronic sequences; 9 CNVs (3 gains and 6 losses) are located in desert genomic regions that may result in a possible position effect on the adjacent genes.*

Three POI patients (POI 39, 46, 54) bear more than one rare CNV.

The analysis of the CNVs gene content using *ad hoc* databases revealed a total of 37 genes whose alteration might play a role in POI etiology. Based on genes expression and function in the ovary, literature evidences, presence of animal models (i.e, knock-out mice), and experimental studies in affected women (i.e., identification of mutations and/or rare CNVs involving the gene), an increased association score to POI was assigned to each gene (**Tab.12**).

Tab.12. Criteria used for assigning genes POI association score.

	Known expression and/or function	Literature evidences	Relevance in animal	Relevance in Human
1 - Low	+	-	-	-
2 - Medium	+	+	-	-
3 - Moderate	+	+	+	-
4 - High	+	+	+	+

The percentage and the number of the genes in each score set is shown below and in **Fig.8** respectively.

1 – Low → 45.9% 2 – Medium → 29.7% 3 – Moderate → 24.3% 4 – High → 8.1%

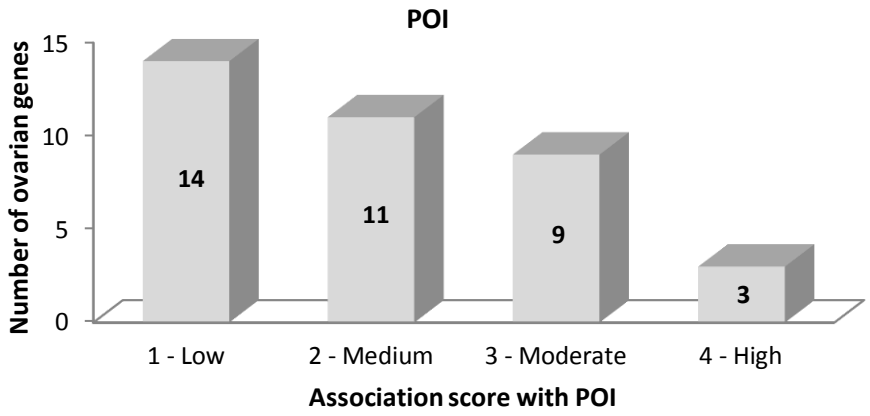

Fig.8. Distribution of the 37 genes potentially related to POI according to the association score. *The number of the genes included is shown.*

Through bioinformatic analysis the genes were classified in four functional categories which contribute to ensure the proper ovarian development and homeostasis (**Fig.9**):

A: genes involved in ovarian metabolism and homeostasis (19%);

B: genes with a role in the intracellular trafficking, and in the regulation of cytoskeleton dynamics for oocyte asymmetric division (21.6%);

C: genes involved in transcription and translation regulation, chromatin remodelling, maintenance of oocyte genomic integrity (35.1%);

D: genes that play a role in ovarian differentiation, follicular development, and meiotic resumption (24.3%).

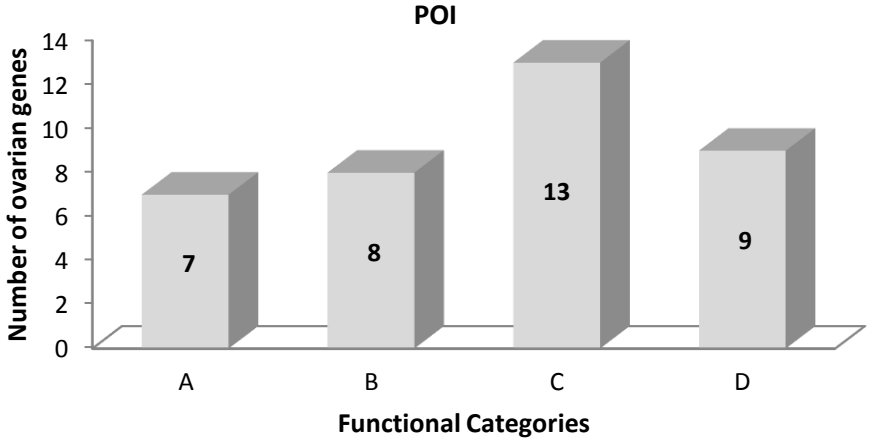

Fig.9. Distribution of the 37 ovarian genes according to the functional class they belong. *The number of genes is shown in each category.*

These data are summarized in **Tab.13** where a detailed list of the 37 selected genes with their general function and role in the ovary is reported. For each gene the identification number of the Ovarian Kaleidoscope database submission (OKdb) is also shown.

Tab.13. Detailed list of the genes potentially perturbed by the identified rare CNVs and possibly implicated in POI pathogenesis.

Gene	POI ID	General Function (from public databases: UCSC, GeneCards)	Role in the ovary	POI score	Functional Category	OKdb #	References
RYR3	3	*Ryanodine receptor 3.* The protein encoded by this gene is a receptor which releases calcium from intracellular storage, contributes to cellular calcium ion homeostasis and plays a role in cellular calcium signaling.	Ryanodine receptors are expressed on the oocyte surface of mature bovine (expression increases in metaphase I) and anphibia. Changes in the levels of calcium are essential for the activation and the development of oocytes.	2	A	-	[173,174]
DIAPH2	6, 55	*Diaphanous-related formin 2.* The gene encodes a protein that in Drosophila plays a role in actin cytoskeleton organisation both in mitosis and meiosis. In human, it could be involved in oogenesis and in the regulation of endosome dynamics.	It plays a role in the development and normal function of the ovaries. Defects have been linked to premature ovarian failure 2A.	4	B	37	[129,175,176]
ELAVL2	7	*ELAV (embryonic lethal, abnormal vision, Drosophila)-like 2.* The encoded protein is a neural-specific RNA-binding protein that is known to bind to several 3' UTRs, including its own.	It plays a role during oocyte maturation and early embriogenesis in bovine and mouse.	2	D	3896	[177,178]
TP63	8, 10	*Tumor protein p63.* It encodes a member of the p53 family of transcription factors and acts as a sequence specific DNA binding transcriptional activator or repressor. Mutations are associated with several autosomal dominant syndromes: ectodermal dysplasia and cleft lip/palate syndrome 3 (EEC3); split-hand/foot malformation 4 (SHFM4); ankyloblepharon-ectodermal defects-cleft lip/palate; ADULT syndrome (acro-dermato-ungual-lacrimal-tooth); limb-mammary syndrome; Rap-Hodgkin syndrome (RHS); and orofacial cleft 8.	It plays a role in the maintainance of oogonia genomic integrity, may be implicated in the meiosis and cell cycle control of germ cells in the mouse ovary. It is a main regulator to protect the fidelity of female germ cells during meiotic arrest.	3	C	1885	[179,180,181,182]

Tab.13. Continued

PRKAA1	9, 11	*Protein kinase, AMP-activated, alpha 1 catalytic subunit.* The encoded protein belongs to the ser/thr protein kinase family. It is the catalytic subunit of the 5'-prime-AMP-activated protein kinase (AMPK). It protects cells from stresses that cause ATP depletion by switching off ATP-consuming biosynthetic pathways.	Murine and porcine animal models show the involvement of AMPK in oocytes meiotic resumption. The process is transiently blocked by PRKA activators in a dose-dependent manner. Moreover, the absence of AMPK modifies murine oocyte quality through energy processes and oocyte/somatic cell communication.	3	D	818	[183,184,185, 186,187,188]
RREB1	12	*Ras responsive element binding protein 1.* It encodes a zinc finger transcription factor that binds to RAS-responsive elements (RREs) of gene promoters. It has been shown to increase expression of calcitonin, which may be involved in Ras/Raf-mediated cell differentiation.	Evidences from drosophila oocytes: link between human homolog RREB1 and regulation of cell adesion and migration.	1	B	-	[189]
SMYD3	13	*SET and MYND domain containing 3.* It encodes a histone methyltransferase which specifically methylates Lys- 4 and Lys-5 of histone H3, inducing di- and tri-methylation, but no monomethylation. Plays an important role in transcriptional activation as a member of an RNA polymerase complex.	Regulates the expression of transcriptional factors during bovine oocyte maturation and early embryonic development.	2	C	4348	[190]
HBE1	13	*Hemoglobin, epsilon 1.* HBE is normally expressed in the embryonic yolk sac: two epsilon chains together with two zeta chains (an alpha-like globin) constitute the embryonic hemoglobin Hb Gower I; two epsilon chains together with two alpha chains form the embryonic Hb Gower II. Both embryonic hemoglobins are normally supplanted by fetal, and later, adult hemoglobin.	It is associated with alterations in the development of the mouse gonads and thus might contribute to sexual dysgenesis syndromes.	1	D	2986	[191]
BMP15	14	*Bone morphogenetic protein 15.* The encoded protein is a member of the BMP family which is part of the transforming growth factor-beta superfamily (TGF-β). The TGF-β superfamily includes large families of growth and differentiation factors.	The protein may be involved in oocyte maturation and follicular development as a homodimer or by forming heterodimers with a related protein, Gdf9. Several *BMP15* mutations are involved in the pathogenesis of hypergonadotropic ovarian failure in humans.	4	D	8	[94,96,97,98]

Tab.13. Continued

BMP4	18	*Bone morphogenetic protein 4*. The encoded protein is a member of the BMP family which is part of the transforming growth factor-beta superfamily (TGF-β). It induces cartilage and bone formation and acts in mesoderm induction, tooth development, limb formation and fracture repair.	*BMP4* plays a role in follicular development and differentiation acting as a transition factor from primordial to primary mouse follicle. Knock-out mice are infertile. *BMP4* gene defects may not represent a risk factor in the development of POI among 99 Chinese Han women analyzed.	3	D	976	[192,193,194, 195,196]
MOV10	25	*Moloney leukemia virus 10, homolog (mouse)*. It encodes a probable RNA helicase. Required for RNA-mediated gene silencing by the RNA-induced silencing complex (RISC).	The gene plays a role as a telomerase associated protein and its mRNA was found highly expressed in human testis and ovary.	1	C	4113	[197]
RHOC	25	*Ras homolog family member C*. It encodes a member of the Rho family of small GTPases. It is thought to be important in cell locomotion. Regulates a signal transduction pathway linking plasma membrane receptors to the assembly of focal adhesions and actin stress fibers. It serves as a microtubule-dependent signal required for the myosin contractile ring formation during cell cycle cytokinesis.	*RHOC* mRNA is expressed during mouse oocyte development from the primordial to large antral follicular stages.	1	B	4750	[198]
CAPZA1	25	*Capping protein (actin filament) muscle Z-line, alpha 1*. It encodes a member of the F-actin capping protein alpha subunit family. The protein regulates growth of the actin filament by capping their barbed end.	*CAPZA1* mRNA is expressed during mouse oocyte development from the primordial to large antral follicular stages.	1	B	4660	[198]
LRIG2	25	*Leucine-rich repeats and immunoglobulin-like domains 2*. The encoded protein promotes epidermal growth factor signalling, resulting in increased proliferation.	*LRIG2* mRNA is expressed during mouse oocyte development from the primordial to large antral follicular stages.	1	C	2477	[198]
PCSK5	33	*Proprotein Convertase, Subtilisin/kexin-type, 5*. The encoded protein belongs to the subtilisin-like proprotein convertase family. The members of this family process latent precursor proteins into their biologically active products.	The gene plays a role in the processing of mouse ovarian inhibin subunits during folliculogenesis and its enzymatic product has been suggested to be an important regulator of inhibin and activin bioavailability.	3	A	1517	[199,200]

Tab.13. Continued

SLC5A5	38, 39	*Solute carrier family 5 (sodium iodide symporter)*. It encodes a member of the sodium glucose cotransporter family responsible for the uptake of iodine in tissues such as the thyroid and lactating breast tissue. Gene mutations are associated with autosomal recessive thyroid dyshormonogenesis 1, leading to congenital hypothyroidism.	*SLC5A5* mRNA is expressed during mouse oocyte development from the primordial to large antral follicular stages.	1	A	1764	[198]
DDX1	39	*DEAD (Asp-Glu-Ala-Asp) box polypeptide 1*. It encodes a putative RNA helicase implicated in a number of cellular processes involving alteration of RNA secondary structure such as translation initiation, nuclear and mitochondrial splicing, and ribosome and spliceosome assembly. Based on their distribution patterns, some members of this family are believed to be involved in embryogenesis, spermatogenesis, and cellular growth and division.	It plays a role in fertility and development, and localizes in Drosophila oocyte during oogenesis.	2	C	-	[201]
MYCN	39	*v-myc myelocytomatosis viral related oncogene, neuroblastoma derived*. It is a member of the MYC family and encodes a protein with a basic helix-loop-helix (bHLH) domain. Gene amplification is associated with a variety of tumors, most notably neuroblastoma.	*MYCN* mRNA is expressed during mouse oocyte development from the primordial to large antral follicular stages.	1	D	4089	[198]
XPO1	42	*Exportin 1*. The encoded protein mediates leucine-rich nuclear export signal (NES)-dependent protein transport. It is involved in the control of several cellular processes by controlling the localization of cyclin B, MAPK, and MAPKAP kinase 2.	*XPO1* mRNA is expressed during mouse oocyte development from the primordial to large antral follicular stages.	1	B	4425	[198]
VLDLR	44	*Very low density lipoprotein receptor*. It encodes a lipoprotein receptor that is a member of the LDLR family and plays important roles in VLDL-triglyceride metabolism and in the reelin signaling pathway.	*VLDLR* plays an important role in cholesterol and lipoprotein endocitosis in ovarian follicle. Several studies on bovine and chicken ovaries indicate its role during follicular development. Hens lacking a functional receptor are sterile. It is also associated with egg performance in duck.	3	A	2159	[202,203,204,205]

Tab.13. Continued

PARD3	45	*Partitioning defective 3 homolog*. It encodes a member of the PARD protein family and interacts with other PARD family members and other proteins; it plays a role in asymmetrical cell division and directs polarized cell growth.	It plays a role in establishing asymmetry division in mouse and Xenopus oocytes and in defining the future site of polar body emission.	3	B	4194	[206,207]
NRP1	45	*Neuropilin 1, vascular endothelial cell growth factor*. The encoded protein plays a role in several different types of signaling pathways that control cell migration. It is involved in the development of cardiovascular system, angiogenesis, the formation of certain neuronal circuits and in organogenesis outside the nervous system.	The *NRP1* gene is expressed in the granulosa and theca cells in the bovine ovary, suggesting its involvement in follicle development in the cow.	2	C	1582	[208]
CPEB1	46, 54	*Cytoplasmic polyadenylation element binding protein 1*. It encodes a member of the cytoplasmic polyadenylation element (CPE) binding protein family. This highly conserved protein binds to a specific RNA sequence called the CPE found in the 3' UTR of some mRNAs.	*CPEB1* is an important regulator of translation in oocytes and neurons. It plays a role in asymmetric division of Xenopus oocytes and is involved in Cyclin B translation and meiotic resumption in porcine oocytes. Cpeb1 controls mouse germ cell differentiation by regulating the formation of the synaptonemal complex: KO mouse are infertile. Microdeletion in the gene was found in a patient with POI through high-resolution SNP array analysis.	4	C	4491	[161,209,210, 211]
AMFR	46	*Autocrine motility factor receptor*. It encodes a receptor whose ligand, autocrine motility factor, is a tumor motility-stimulating protein secreted by tumor cells. The receptor is a member of the E3 ubiquitin ligase family of proteins, catalyzes ubiquitination and endoplasmic reticulum-associated degradation of specific proteins.	*AMFR* mRNA is expressed during mouse oocyte development from the primordial to large antral follicular stages.	1	A	1354	[198]

Tab.13. Continued

GRIA1	48	*Glutamate receptor, ionotropic, AMPA 1.* Glutamate receptors are the predominant excitatory neurotransmitter receptors in the mammalian brain and are activated in a variety of normal neurophysiologic processes. They are heteromeric protein complexes with multiple subunits, each possessing transmembrane regions, and all arranged to form a ligand-gated ion channel.	*GRIA1* is a critical mediator of ovulation in cow and it might be a useful target for reproductive therapy.	2	D	4841	[212,213]
STS	50	*Steroid sulfatase.* The encoded protein catalyzes the conversion of sulfated steroid precursors to estrogens during pregnancy. It is found in the endoplasmic reticulum, where it acts as a homodimer.	*STS* increases StAR (steroidogenic acute regulatory) protein expression level and stimulates steroid production. It is also expressed in human cumulus cells and its mRNA levels are controlled by FSH.	2	A	1166	[214,215]
TSHZ2	52, 62	*Teashirt zinc finger homeobox 2.* It encodes a probable transcriptional regulator involved in developmental processes.	*TSHZ2* mRNA is expressed during mouse oocyte development from the primordial to large antral follicular stages.	1	C	4602	[198]
FZD6	53	*Frizzled family receptor 6.* The encoded protein functions as a negative regulator of the canonical Wnt/beta-catenin signaling cascade, thereby inhibiting the processes that trigger oncogenic transformation, cell proliferation, and inhibition of apoptosis.	*FZD6* mRNA is regulated by FSH during dominant follicle selection in bovine oocyte.	2	D	2204	[216]
SLMAP	54	*Sarcolemma associated protein.* It encodes a component of a conserved striatin-interacting phosphatase and kinase complex. Striatin family complexes participate in a variety of cellular processes including signaling, cell cycle control, cell migration, Golgi assembly, and apoptosis.	*SLMAP* mRNA is expressed during mouse oocyte development from the primordial to large antral follicular stages.	1	C	4945	[198]
WHAMM	54	*WAS protein homolog associated with actin, golgi membranes and microtubules.* The encoded protein acts as a nucleation-promoting factor (NPF) that stimulates actin polymerization both at the Golgi apparatus and along tubular membranes. It participates in vesicle transport between the reticulum endoplasmic and the Golgi complex.	*WHAMM* is required for meiotic spindle migration and asymmetric cytokinesis in mouse oocytes.	3	B	4788	[217]

Tab.13. Continued

BTBD1	54	*BTB (POZ) domain containing 1 (BTBD1)*. It encodes a probable substrate-specific adapter of an E3 ubiquitin-protein ligase complex which mediates the ubiquitination and subsequent proteasomal degradation of target proteins.	Danio rerio homolog may have an important function in oogenesis and (or) early zebrafish development.	1	A	-	[218]
BNC1	54	*Basonuclin 1*. The encoded protein is a zinc finger protein present in the basal cell layer of the epidermis and in hair follicles. It is also found in abundance in the germ cells of testis and ovary. It is thought to play a regulatory role in keratinocyte proliferation and it may also be a regulator for rRNA transcription. Basonuclin is expressed in cells that are able to undergo division but are not necessarily in the division cycle; it is not found in terminally differentiated cells.	It may play a role in the differentiation of mouse oocytes. It increases transcription of the ribosomal RNA genes during mouse oogenesis. Moreover, mouse with null mutation is sub-fertil.	3	C	101	[219,220]
SH3GL3	54	*SH3-domain GRB2-like 3 (Endophilin A3)*. SH3 domain is an evolutionarily conserved, 50- to 60-amino acid module, carried by intracellular proteins involved in the transduction of signals for cell polarization, motility, enzymatic activation, and transcriptional regulation. Implicated in endocytosis. May recruit other proteins to membranes with high curvature.	The endophilins differentially contribute to oocyte endocytosis and development in chicken. Duplication of the gene was found in Norling *et al.*, 2014 by means of custom array-CGH analysis in a cohort of POI patients but an association with the disorder should be further investigated.	2	B	1731	[159,221]
MSL3	60	*Male specific lethal 3 homolog (Drosophila)*. The encoded protein may be involved in chromatin remodeling and transcriptional regulation. May have a role in X inactivation and it has been found as part of a complex that is responsible for histone H4 lysine-16 acetylation.	*MSL3* may be a candidate epigenetic reprogramming factor in human, rhesus monkey, and mouse oocyte.	2	C	5012	[222]
EEF2	63	*Eucariotic tanslation elongation factor 2*. It encodes a member of the GTP-binding translation elongation factor family. The protein, an essential factor for protein synthesis, promotes the GTP-dependent translocation of the nascent protein chain from the A-site to the P-site of the ribosome.	*EEF2* mRNA is expressed during mouse oocyte development from the primordial to large antral follicular stages.	1	C	306	[198]

Tab.13. Continued

MAP2K2	63	*Mitogen-activated protein kinase kinase 2 (MEK2).* The encoded protein is a dual specificity protein kinase that belongs to the MAP kinase kinase family. This kinase is known to catalyze the concomitant phosphorylation of a threonine and a tyrosine residue in a Thr-Glu-Tyr sequence located in MAP kinases.	*MEK1/2* regulates spindle formation and meiotic resumption during mouse oocyte development.	3	D	-	[223,224,225]
SIRT6	63	*Sirtuin 6.* It encodes a member of the sirtuin family of proteins, homologs to the yeast Sir2 protein which is a NAD-dependent protein deacetylase.	*SIRT6* plays a role in caloric restriction and high-fat diet which are important for oocyte development and to preserve the germ cells reserve in the mouse. Furthermore, SIRT6 loss leads to genomic instability and aging-like phenotype in mice.	2	C	4470	[226,227]

Transcriptional expression of all the genes was then investigated on ovarian mRNAs. Both total fetal and blood mRNAs were analyzed as different control tissues. The experimental results are shown in **Tab.14**.

Tab.14. Results from the mRNAs analysis: +, expressed; -, not expressed.

	mRNA expression				mRNA expression		
	Ovary	Fetal	Blood		Ovary	Fetal	Blood
AMFR	+	+	+	*NRP1*	+	+	+
BMP15	+	+	+	*PARD3*	+	+	+
BMP4	+	+	+	*PCSK5*	+	+	+
BNC1	+	+	-	*PRKAA1*	+	+	+
BTBD1	+	+	+	*RBSG1*	+	+	+
CAPZA1	+	+	+	*RHOC*	+	+	+
CPEB1	+	+	+	*RREB1*	+	+	+
DDX1	+	+	+	*RYR3*	+	+	+
DIAPH2	+	+	+	*SH3GL3*	+	+	-
EEF2	+	+	+	*SIRT6*	+	+	+
ELAVL2	+	+	-	*SLC5A5*	+	+	+
FZD6	+	+	+	*SLMAP*	+	+	+
GRIA1	+	+	-	*SMYD3*	+	+	+
HBE1	+	+	+	*STS*	+	+	+
LRIG2	+	+	+	*TP63*	+	+	+
MAP2K2	+	+	+	*TSHZ2*	+	+	-
MOV10	+	+	+	*VLDLR*	+	+	+
MSL3	+	+	+	*WHAMM*	+	+	+
MYCN	+	+	+	*XPO1*	+	+	+

To find some functional relationship between the proteins encoded by array-CGH ovarian genes and the already known POI genes (see introduction), bioinformatic search was performed using STRING database (http://string-db.org/). The analysis revealed a link for 11 out of the 32 proteins (Ddx1, Syce1, Xpo1, Eef2, Amfr, Cyp2e1, Smyd3, Prkaa1, Bmp15, Msl3, Sts) that is commented in the discussion.

3.3 Ovarian CNVs enrichment evaluation

3.3.1 Array CGH and CNVs gene content analyses of an *ad hoc* control group

All the CNVs detected were selected as rare according to the DGV. A limitation of this database is the lack of several informations regarding the healthy controls used: mainly it is unknown their gender and age. Moreover, considering that several pathologies have a late onset (e.g. secondary amenorrhea), it is possible to find these patients inside the DGV controls group. For this reason, it is crucial to compare the obtained data with a cohort of *ad hoc* controls.

For Primary Ovarian Insufficiency patients, the *ad hoc* controls consist in healthy women with normal 46,XX karyotype who had a normal reproductive life and reached physiological menopause.

Array CGH 244K and 400K previously performed on 140 control women were thus analyzed with the aim to find a real enrichment in potentially pathogenic ovarian CNVs in the POI patients. The analysis of the CNVs and their gene content in the control group was performed with the same method used for patients' analysis.

In 44 out of 140 controls we found 49 ovarian rare CNVs possibly implicated in POI etiopathogenesis, 25 duplications (51%) and 24 deletions (49%), ranging from 8 kb to 1.8 Mb. According to the type of rearrangement that involves the detected genes, the CNVs were classified as shown in **Fig.10**.

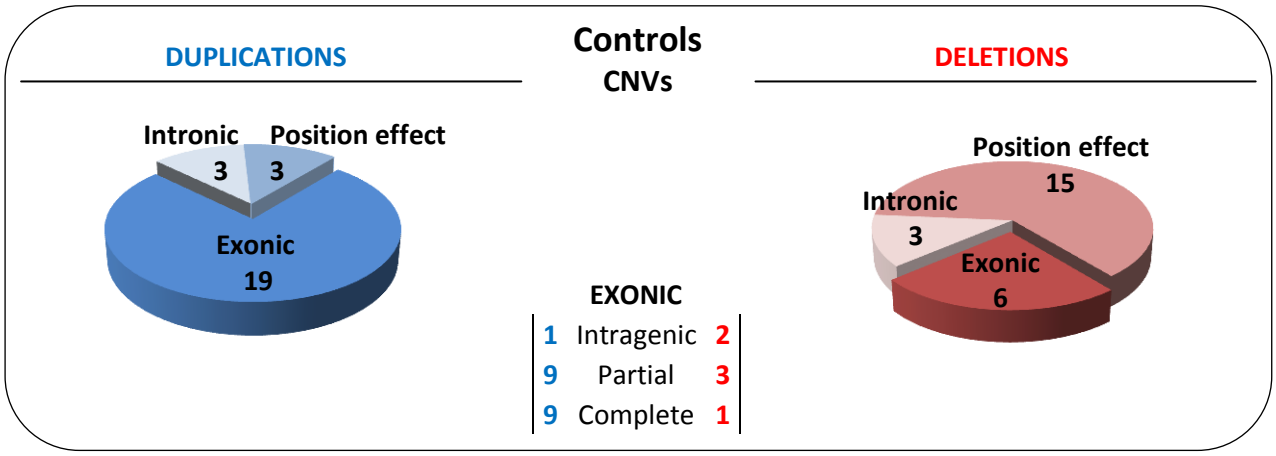

Fig.10. Distribution of the 49 ovarian CNVs in control women. *19 gains and 6 losses involve the genes coding region resulting in intragenic, partial and complete deletion/duplication; 3 gains and 3 losses involve intronic sequences, whereas 18 CNVs (3 gains and 15 losses) are located in desert genomic regions that may result in a possible position effect on the adjacent genes.*

Subsequently the analysis of the CNVs gene content in the control group revealed a total of 54 genes whose alteration might play a role in POI etiology. As well as patients, a POI score was attributed to each gene as shown in **Fig.11**.

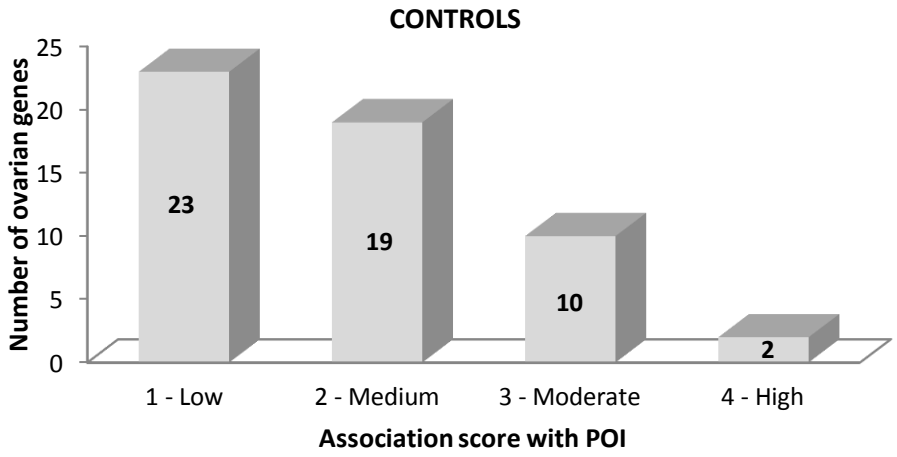

Fig.11. Distribution of the 54 genes included in rare CNVs present in the POI control cohort and potentially related to POI according to the association score. *The number of the genes included is shown.*

Furthermore, as well as in the POI cohort, "common" CNVs containing genes previously associated to the disorder (*NAIP*, *DUSP22*, *SYCE1*) [152,161], were identified:

- 6 CNVs involving *NAIP* at 5q13.2 (3 duplication and 3 deletions),
- 30 CNVs involving *DUSP22* at 6p25.2 (19 duplications and 11 deletions),
- 9 CNVs involving *SYCE1*, *CYP2E1* at 10q26.3 (9 duplications and 0 deletion).

3.3.2 Comparison between POI and *ad hoc* controls analyses

Based on the results obtained from the analyses performed, 4 CNVs involving the same genes were detected in both patients and controls (**Tab.15**).

Tab.15. CNVs in patients and controls involving the same genes.

ID	Chromosome Position(hg19)	G/L	Size	Gene and rearrangement
POI 25	chr1:113339006-113430121	Gain	91 kb	*MOV10*, position effect
CTRL_F122	chr1:113239072-113474655	Gain	236 kb	*MOV10*, partial duplication
POI 42	chr2:61795359-61831792	Loss	36 kb	*XPO1*, position effect
CTRL_F107	chr2:61877860-61888495	Gain	11 kb	*XPO1*, position effect
POI 50	chrX:6456036-8133172	Gain	1.7 Mb	*STS*, complete duplication
CTRL_F105	chrX:6449644-8133172	Gain	1.7 Mb	*STS*, complete duplication
POI 60	chrX:11249664-11996670	Gain	747 kb	*MSL3*, complete duplication
CTRL_F129	chrX:11686237-12187337	Gain	501 kb	*MSL3*, complete duplication

These results prompted us to further evaluate whether there was a real enrichment both in ovarian CNVs and in ovarian genes in the POI women. A preliminary estimation was performed

as shown (**Fig.12**) dividing each value for the total number of POI patients (67) and the *ad hoc* control group (140) analyzed.

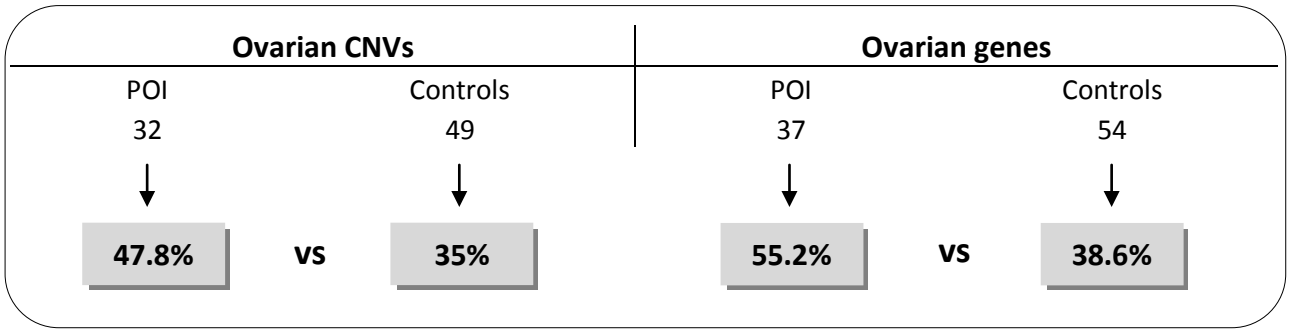

Fig.12. Results obtained from the preliminary evaluation. *The number of CNVs and genes with their respective percentage is shown for both POI and controls.*

As shown in the pre-analysis there seems to be a slight enrichment in favor of the group of patients. Thus, to underline eventual statistically significant differences between POI and *ad hoc* controls data, *chi-square* test was performed comparing the following values:

- number of people positive and negative to the array CGH analysis,
- number of CNVs involving exonic and non-exonic sequences,
- number of genes with a higher (3,4) and lower (1,2) POI association score.

None of the three analyses reached a statistical significance ($p < 0.5$) as shown in **Tab.16**.

Tab.16. Results of the *chi-square* test.

People	**Observed Frequencies**				**Expected Frequencies**		
		POI	Controls	Total		POI	Controls
	array-CGH Positive	28	44	72	array-CGH Positive	23.304	48.696
	array-CGH Negative	39	96	135	array-CGH Negative	43.696	91.304
	Total	67	140	207			
					***chi-square* test**	**0.143**	***p*>0.05**

CNVs	**Observed Frequencies**				**Expected Frequencies**		
		POI	Controls	Total		POI	Controls
	Exonic CNVs	18	25	43	Exonic CNVs	16.988	26.012
	Non-exonic CNVs	14	24	38	Non-exonic CNVs	15.012	22.988
	Totale	32	49	81			
					***chi-square* test**	**0.645**	***p*>0.05**

Genes	**Observed Frequencies**				**Expected Frequencies**		
		POI	Controls	Total		POI	Controls
	gene score 3,4	12	12	24	gene score 3,4	9.758	14.242
	gene score 1,2	25	42	67	gene score 1,2	27.242	39.758
	Total	37	54	91			
					***chi-square* test**	**0.278**	***p*>0.05**

After, dividing the total count of "common" CNVs detected in the two groups by the total number of POI and controls, a comparative evaluation was performed and no differences were identified (**Fig.13**). Fisher's exact test confirmation was also obtained (data not shown).

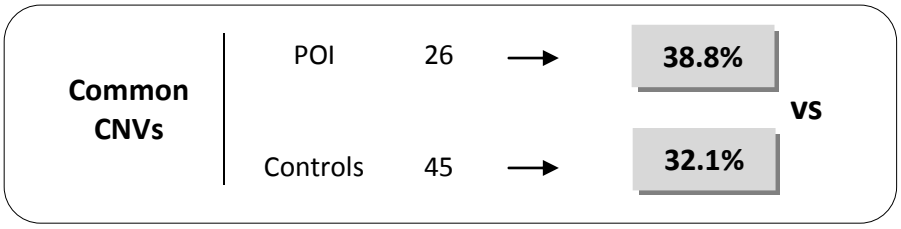

Fig.13. Results obtained from the evaluation of common CNVs. *The CNVs number with the respective percentage is shown for both POI and controls.*

Nevertheless, when one compares the number of rearrangements with the most probably deleterious effect on the genes with the highest association POI score (3,4), a difference can be observed. In detail, POI patients present 7 "deleterious" CNVs compared to the only one found in controls **Tab.17**.

Tab.17. List of the highest association POI score genes with the respective rearrangement.

POI			Controls		
Gene	Type of rearrangement	POI score	Gene	Type of rearrangement	POI score
BMP15	**Complete gene deletion**	4	*LHCGR*	Position effect	4
CPEB1 (x2)	Intronic deletion, **complete gene deletion**	4	*FSHR (x2)*	Position effect	4
DIAPH2 (x2)	Intronic duplication, position effect	4	*PTN*	**Partial gene deletion**	3
*TP63 (x2)**	**Intragenic duplication**	3	*ADAMTS9*	Partial gene duplication	3
VLDLR	**Partial gene deletion**	3	*PDE9A*	Partial gene duplication	3
WHAMM	**Complete gene deletion**	3	*PRKCG*	Complete gene duplication	3
BNC1	**Complete gene deletion**	3	*DAB1*	Intronic deletion	3
PCSK5	Partial gene duplication	3	*GREM2*	Position effect	3
MAP2K2	Complete gene duplication	3	*FMN2*	Position effect	3
*PRKAA1 (x2)**	Intronic deletion	3	*IL6ST*	Position effect	3
BMP4	Position effect	3	*ADAMTS19*	Position effect	3
PARD3	Position effect	3	*AKAP2*	Position effect	3

***** gene found in familial cases;
Rearrangements with the most probably deleterious effect are shown in bold and underlined.
(x2): found in two CNVs

3.4 Molecular characterization of selected CNVs and study of their putative involvement in POI pathogenesis

3.4.1 POI 6: intronic duplication at Xq21.33 (*DIAPH2*)

The 53 kb duplication inherited from the healthy father, ranges from oligonucleotide A_16_P03733864 to A_16_P03733968 involving intron 26 of the gene and the entire long non-coding RNA named *DIAPH2*-antisense RNA 1 (*RBSG1*, chrX:96,692,826-96,819,534, hg19). As antisense RNAs play a regulatory role against the complementary genes, RT-qPCR on *DIAPH2* was performed to evaluate possible alterations at the transcript level. Two TaqMan assay probes were used: one specific for exons 15-16 junction, and the other specific for exons 26-27 (including the IVS26 duplication). Both assays detected no mRNA level alterations (**Fig.14**).

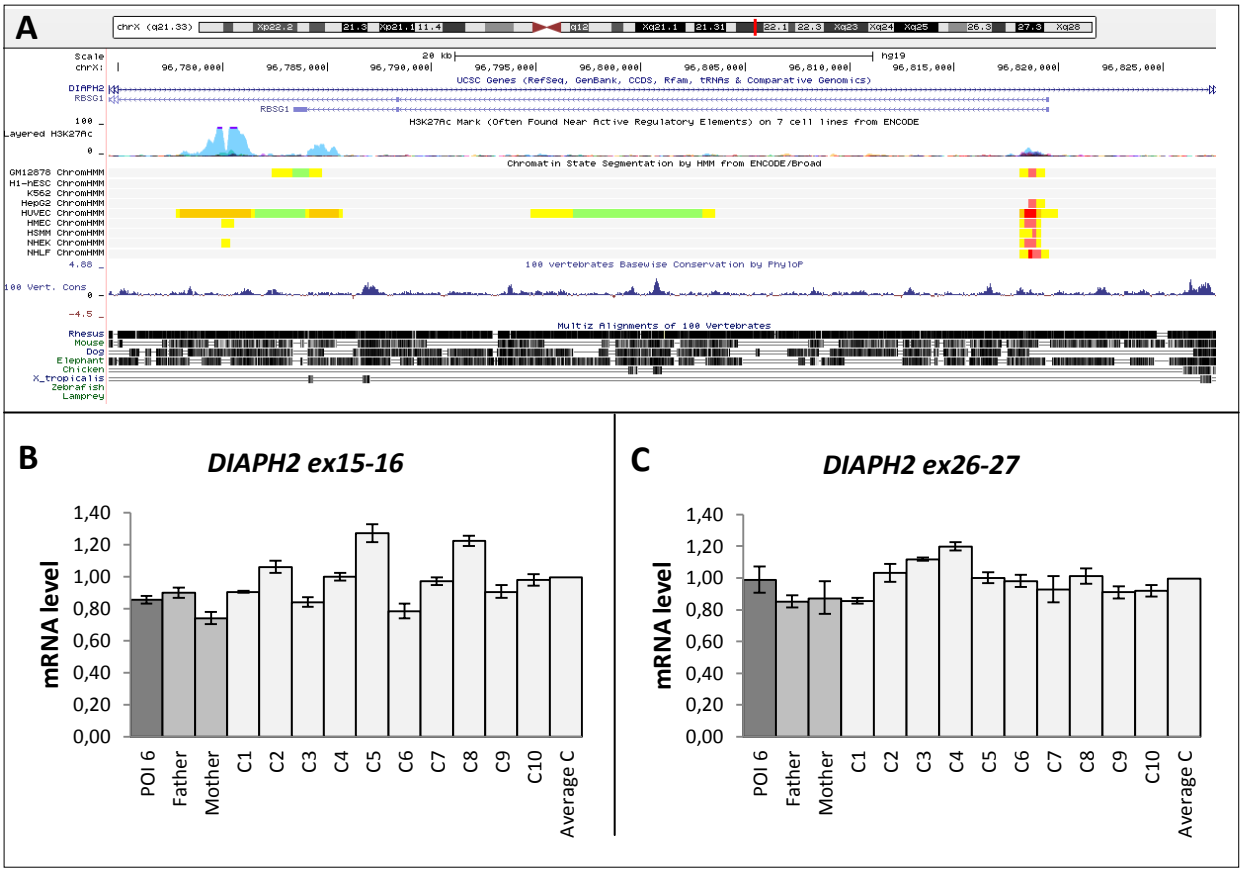

Fig.14. A) *UCSC genome browser view of the region involved in DIAPH2 IVS26 gain (chrX:96,774,546-96,827,670, hg19). UCSC genes are shown in blue. Regulatory elements prediction based on ENCODE chromatin state segmentation in nine human cell lines is shown: red, promoter; yellow, weak enhancer; orange, strong enhancer; light green, weak transcribed; light grey, heterochromatin. Conservation in other species is shown.* ***B)*** *RT-qPCR results using TaqMan assay for exons 15-16; TBP normalization is shown.* ***C)*** *RT-qPCR results using TaqMan assay for exons 26-27; GUSB normalization is shown. C1-C10, controls (light grey bars). POI patient is depicted in dark grey.*

3.4.2 POI 8 and 10: intragenic duplication at 3q28 (*TP63*)

The 161 kb paternal intragenic duplication ranges from oligonucleotide A_16_P16543639 to A_16_P16544117 and involves exons 2 to 9 of the complete*TP63* transcript composed of 14 exons, whereas it involves 5'UTR to exon 7 of the shortest transcripts (**Fig.15**). The outer duplication of about 170.6 kb starts after A_16_P00894458 and ends before A_16_P36592945 (chr3:189,422,152-189,592,727, hg19). Based on these results it is unknown whether exon 10 (or 8 for the shortest isoforms) is involved in the gain.

Fig.15. UCSC genome browser view of the region involved in the intragenic TP63 gain (chr3:189,427,139-189,588,477, hg19). *UCSC genes are shown in blue. Regulatory elements prediction based on ENCODE chromatin state segmentation in nine human cell lines is shown: red, promoter; yellow, weak enhancer; orange, strong enhancer; dark green, transcriptional elongation; light green, weak transcribed; light blue, insulator; dark grey, polycomb-repressed; light grey, heterochromatin. Conservation in other species is shown.*

To understand how the duplication is relocated, LR-PCR was performed on patients DNA at the breakpoint junctions of the duplicated portions. In detail, a direct tandem relocation was considered as shown in **Fig.16**.

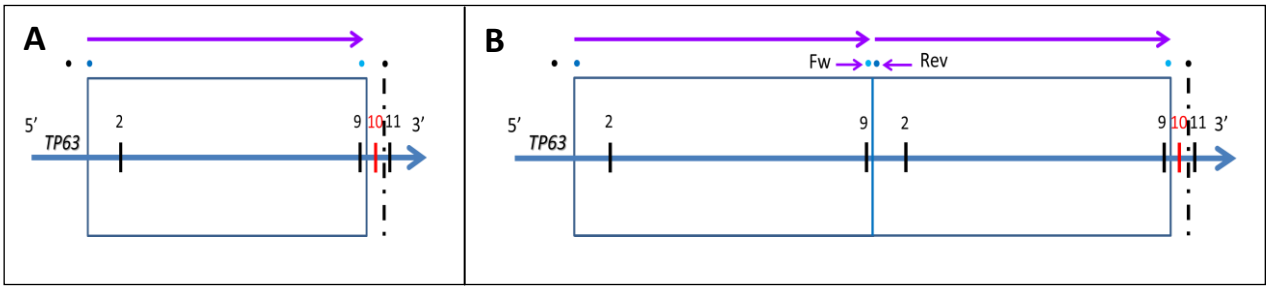

Fig.16. Schematic representation of the duplication repositioning hypothesis involving TP63. A) *The genomic duplication is shown as a box with the exons included in black (2 and 9 are shown). Dark and light blue points indicate the first and the last duplicated array spots respectively; black points indicate the first and the last array spots not duplicated. Inclusion of exon 10 (in red) in the duplication is unknown. Exon 11 (in black) is not included.* ***B)*** *A direct tandem orientation is depicted as indicated by the long purple arrows. Primers for LR-PCR are shown as short purple arrows.*

Moreover, considering that the distance between: i) the last probe not duplicated and the first duplicated (black and dark blue in **Fig.16**) is 4.9 kb; ii) the last probe duplicated and the first not duplicated (light blue and black in **Fig.16**) is 4.2 kb, an LR-PCR amplification of maximum 9.1 kb is expected. As a result, a product of about 7 kb was obtained and the subsequent Sanger sequencing revealed the expected sequence: the duplication relocated in a direct tandem orientation and exon 10 was found included in the gain (**Fig.17**). The shortest transcripts (composed by 8 exons) resulted fully duplicated. As shown in **Fig.17** the duplication of 167.8 kb starts within intron 1 (31191 bp before exon 2) and ends within intron 11 (1360 bp after exon 10). An overlapping region of about 11 bp between the two introns was found. Considering the overlapping sequence as part of intron 11 the gain was refined at chr3:189,424,338-189,592,155 (hg19).

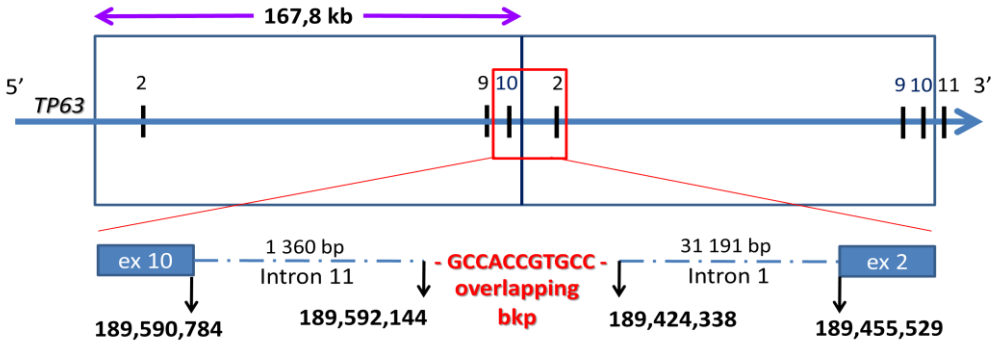

Fig.17. Schematic view of the duplication repositioning. *The red box highlights the enlargement of the breakpoint junction shown below.*

To clarify whether the intragenic duplication results in an aberrant mRNA, a PCR reaction specific for the junction between exon 10 and 2 (**Fig.18**) was performed. The PCR product was then sequenced and revealed the expected sequence confirming that the duplication produces an aberrant transcript (**Fig.18**).

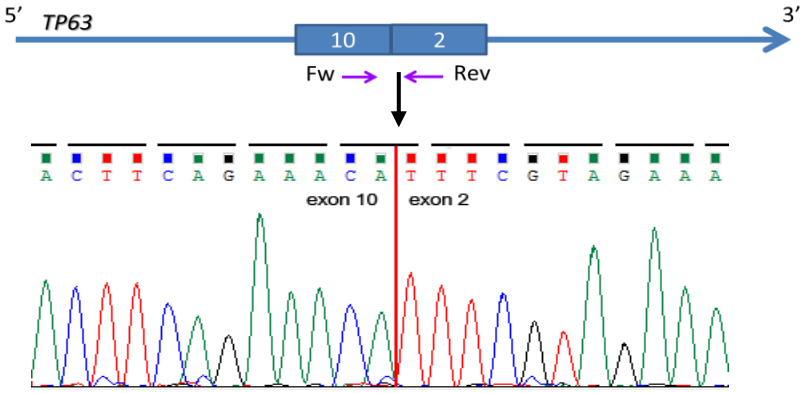

Fig.18. TP63 cDNA representation of the duplication junction. *Above, exons 10 and 2 are shown as blue boxes. PCR primers are depicted as short purple arrows. Below, the electropherogram of the breakpoint junction between the two exons is shown. The black lines represent the reading frame.*

RT-qPCR analysis on *TP63* was carried out using two TaqMan assays as shown in **Fig.19A**. The first one is located inside the duplication (ex 7-8, **Fig.19A**, in red) whereas the second is located outside (ex 11-12, **Fig.19A**, in blue). Likely due to the extreme variability among controls mRNA levels, neither the first assay nor the second detected any up or down regulation of the POI 10 transcript levels when compared to controls (**Fig.19B-C**). Conversely, normalizing the Ct values of the first "duplicated" assay with the Ct values of the second "not duplicated" assay, it was possible to evidence an overexpression statistically significant in the patients compared to controls which is indicative of the amount of the aberrant transcript (**Fig.19D**). Thus, the RT-qPCR results indicate that POI 10 presents a *TP63* mRNA level comparable to the average variability of controls mRNA, but a small amount of this transcript is represented by the duplicated aberrant mRNA (about 30% of the total mRNA level).

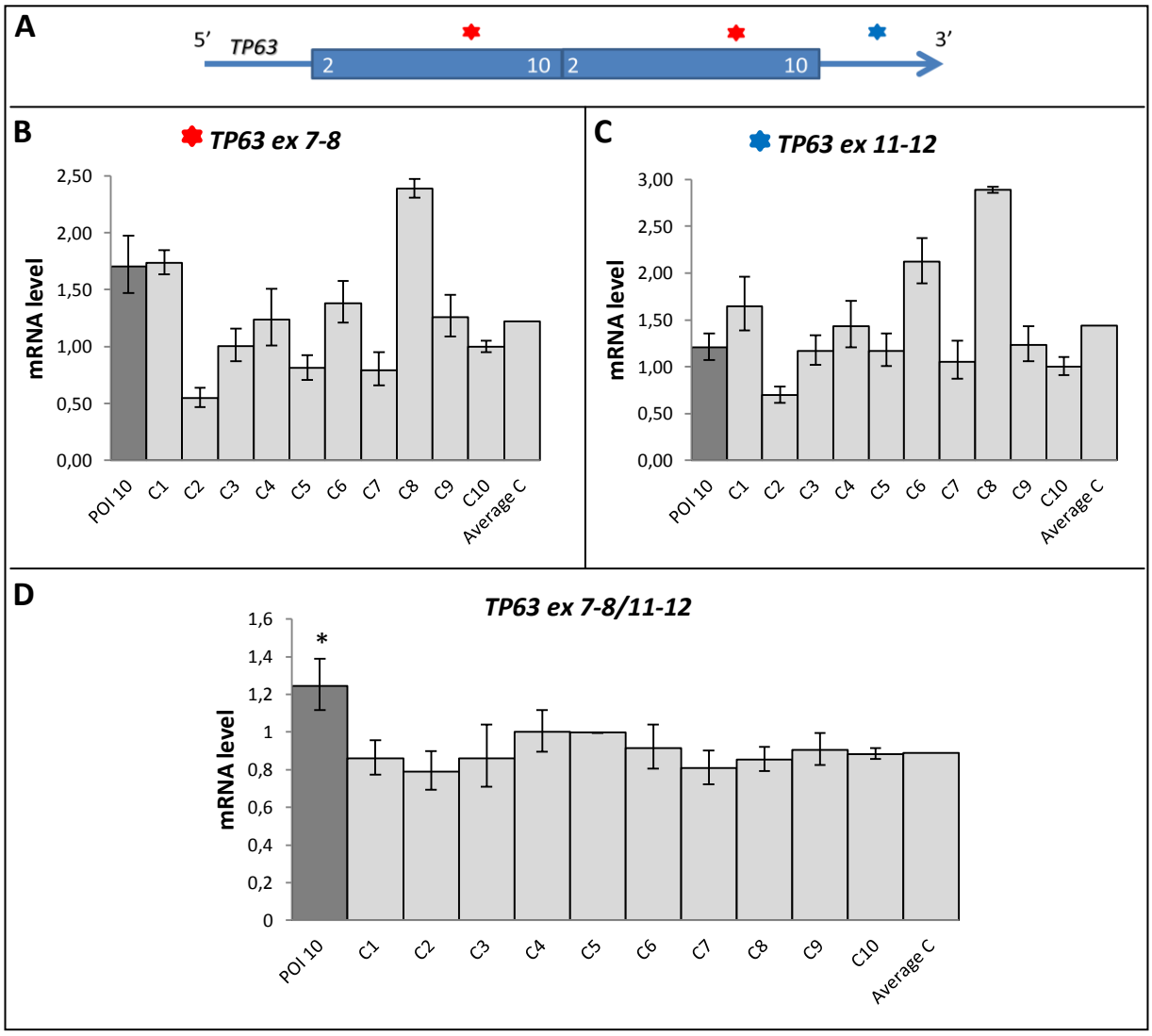

Fig.19. RT-qPCR results. A) *Schematic representation of the TP63 aberrant transcript; TaqMan assays 7-8 and 11-12 are shown as red and blue stars respectively.* ***B-C)*** *RT-qPCR results indicative of the "duplicated" (B) and "not duplicated" assay (C) normalized against RPLP0.* ***D)*** *RT-qPCR result indicative of the aberrant mRNA amount. C1-C10, controls (light grey bars); POI patient is depicted in dark grey.* $* p<0.05$ *vs avg C.*

The previous cDNA sequence analysis of the exon junction between exon 10 and 2, revealed that no changes in the reading frame occurred in the *TP63* aberrant mRNA. Particularly, the new overlap splice site formed by CA nucleotides of exon 10 and T nucleotide of exon 2, encodes a hystidine (H) (**Fig.20**).

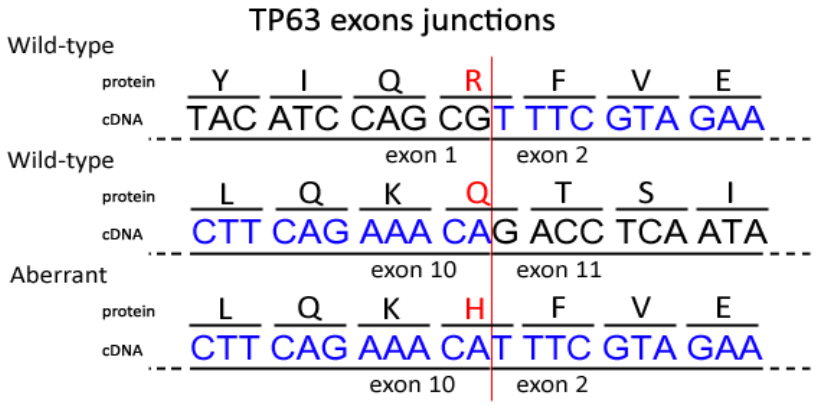

Fig.20. Tp63 exon junction representation. *Above, junction between wild-type exons 1-2; in the middle, junction between wild-type exons 10-11; below junction between aberrant exons 10-2. Protein and nucleotide sequences are shown. Exons 2 and 10 involved in the duplication are depicted in blue; the residues of overlap splice site are represented in red. The red vertical line indicates the junction between exons.*

The wild-type Tp63 protein sequence is constituted by 680 aa (**Fig.21A**). Considering that the intragenic duplication does not alter the reading frame, the hypothetical aberrant Tp63 protein might be formed by 1109 aa (**Fig.21B**). In detail, the duplication ranges from residue 22 to 449, involving a total of 429 aa, 428 of which completely duplicated in tandem (underlined in **Fig.21A-B**) and 1 aa derived from the new overlap splice site between exon 10 and 2 (H shown in red big and bolded in **Fig.21B**). The 1109 aa prediction was also confirmed by the "Expasy translate tool" (http://web.expasy.org/translate/).

Furthermore, as shown in **Fig.21C**, the nine duplicated exons encode the principal domain of the protein, a DNA binding domain belonging to the p53 superfamily.

A

TP63 wild-type Protein Sequence

MNFETSRCATLQYCPDPYIQR**F**VETPAHFSWKESYYRSTMSQSTQTNEFLSPEVFQHIWDFLEQPICSVQPIDLNFVDEPSEDGAT

NKIEISMDCIRMQDSDLSDPMWPQYTNLGLLNSMDQQIQNGSSSTSPYNTDHAQNSVTAPSPYAQPSSTFDALSPSPAIPSNTD

YPGPHSFDVSFQQSSTAKSATWTYSTELKKLYCQIAKTCPIQIKVMTPPPQGAVIRAMPVYKKAEHVTEVVKRCPNHELSREFNEG

QIAPPSHLIRVEGNSHAQYVEDPITGRQSVLVPYEPPQVGTEFTTVLYNFMCNSSCVGGMNRRPILIIVTLETRDGQVLGRRCFEAR

ICACPGRDRKADEDSIRKQQVSDSTKNGDGTKRPFRQNTHGIQMTSIKKRRSPDDELLYLPVRGRETYEMLLKIKESLELMQYLPQ

HTIETYRQQQQQQHQHLLQ**K**QTSIQSPSSYGNSSPPLNKMNSMNKLPSVSQLINPQQRNALTPTTIPDGMGANIPMMGTHM

PMAGDMNGLSPTQALPPPLSMPSTSHCTPPPPYPTDCSIVSFLARLGCSSCLDYFTTQGLTTIYQIEHYSMDDLASLKIPEQFRHAI

WKGILDHRQLHEFSSPSHLLRTPSSASTVSVGSSETRGERVIDAVRFTLRQTISFPPRDEWNDFNFDMDARRNKQQRIKEEGE

B

TP63 aberrant Protein Sequence

MNFETSRCATLQYCPDPYIQR**F**VETPAHFSWKESYYRSTMSQSTQTNEFLSPEVFQHIWDFLEQPICSVQPIDLNFVDEPSEDGAT

NKIEISMDCIRMQDSDLSDPMWPQYTNLGLLNSMDQQIQNGSSSTSPYNTDHAQNSVTAPSPYAQPSSTFDALSPSPAIPSNTD

YPGPHSFDVSFQQSSTAKSATWTYSTELKKLYCQIAKTCPIQIKVMTPPPQGAVIRAMPVYKKAEHVTEVVKRCPNHELSREFNEG

QIAPPSHLIRVEGNSHAQYVEDPITGRQSVLVPYEPPQVGTEFTTVLYNFMCNSSCVGGMNRRPILIIVTLETRDGQVLGRRCFEAR

ICACPGRDRKADEDSIRKQQVSDSTKNGDGTKRPFRQNTHGIQMTSIKKRRSPDDELLYLPVRGRETYEMLLKIKESLELMQYLPQ

HTIETYRQQQQQQHQHLLQ**KH****F**VETPAHFSWKESYYRSTMSQSTQTNEFLSPEVFQHIWDFLEQPICSVQPIDLNFVDEPSED

GATNKIEISMDCIRMQDSDLSDPMWPQYTNLGLLNSMDQQIQNGSSSTSPYNTDHAQNSVTAPSPYAQPSSTFDALSPSPAIPS

NTDYPGPHSFDVSFQQSSTAKSATWTYSTELKKLYCQIAKTCPIQIKVMTPPPQGAVIRAMPVYKKAEHVTEVVKRCPNHELSREF

NEGQIAPPSHLIRVEGNSHAQYVEDPITGRQSVLVPYEPPQVGTEFTTVLYNFMCNSSCVGGMNRRPILIIVTLETRDGQVLGRRC

FEARICACPGRDRKADEDSIRKQQVSDSTKNGDGTKRPFRQNTHGIQMTSIKKRRSPDDELLYLPVRGRETYEMLLKIKESLELMQ

YLPQHTIETYRQQQQQQHQHLLQ**K**QTSIQSPSSYGNSSPPLNKMNSMNKLPSVSQLINPQQRNALTPTTIPDGMGANIPMMG

THMPMAGDMNGLSPTQALPPPLSMPSTSHCTPPPPYPTDCSIVSFLARLGCSSCLDYFTTQGLTTIYQIEHYSMDDLASLKIPEQFR

HAIWKGILDHRQLHEFSSPSHLLRTPSSASTVSVGSSETRGERVIDAVRFTLRQTISFPPRDEWNDFNFDMDARRNKQQRIKEEGE

C

TP63 Protein Domain

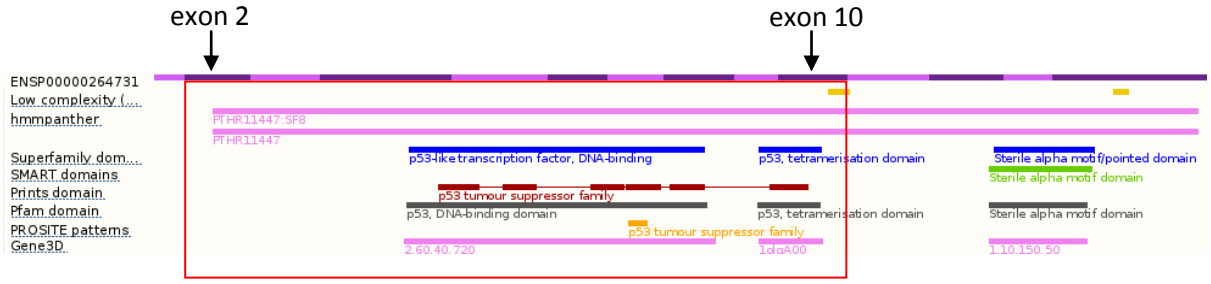

Fig.21. Tp63 protein representations. A-B) *Protein sequence of the wild-type (A) and aberrant (B) Tp63; the underlined sequence represents the duplicated portion; alternating exons are shown in black and blue; the residues of overlap splice site are represented in red; aminoacid F big and bold indicates the beginning of exon 2; aminoacid K big and bold indicates the end of exon 10; aminoacid H big and bold indicates the new residue of overlap splice site between exon 10 and exon 2.* ***C)*** *Schematic view of Tp63 protein domain from ensembl database (www.ensembl.org/); black arrows indicate exons 2 and 10; a red line highlights the duplicated domain.*

3.4.3 POI 9 and 11: intronic deletion at 5p13.1 (*PRKAA1*)

The deletion found in POI sisters and inherited from the father involves *PRKAA1* intron 1. Based on array CGH results the inner deletion of about 6.5 kb ranges from oligonucleotide A_16_P37159287 to A_16_P17099229 (chr5:40,779,232-40,785,703, hg19; **Fig.22A**), whereas the outer deletion of about 20 kb starts after probe A_16_P17099202 and ends before A_14_P135087 (chr5:40,773,231-40,793,258, hg19). Considering that part of the exonic sequence of *PRKAA1* (exons 2, 3, and 4) was not covered by array CGH probes, qPCR on patients DNA using six primer pairs (above in the **Fig.22A**), was carried out to establish the extent of the deletion. The analysis excluded an involvement of the conding region, confirming an intronic deletion (**Fig.22B**).

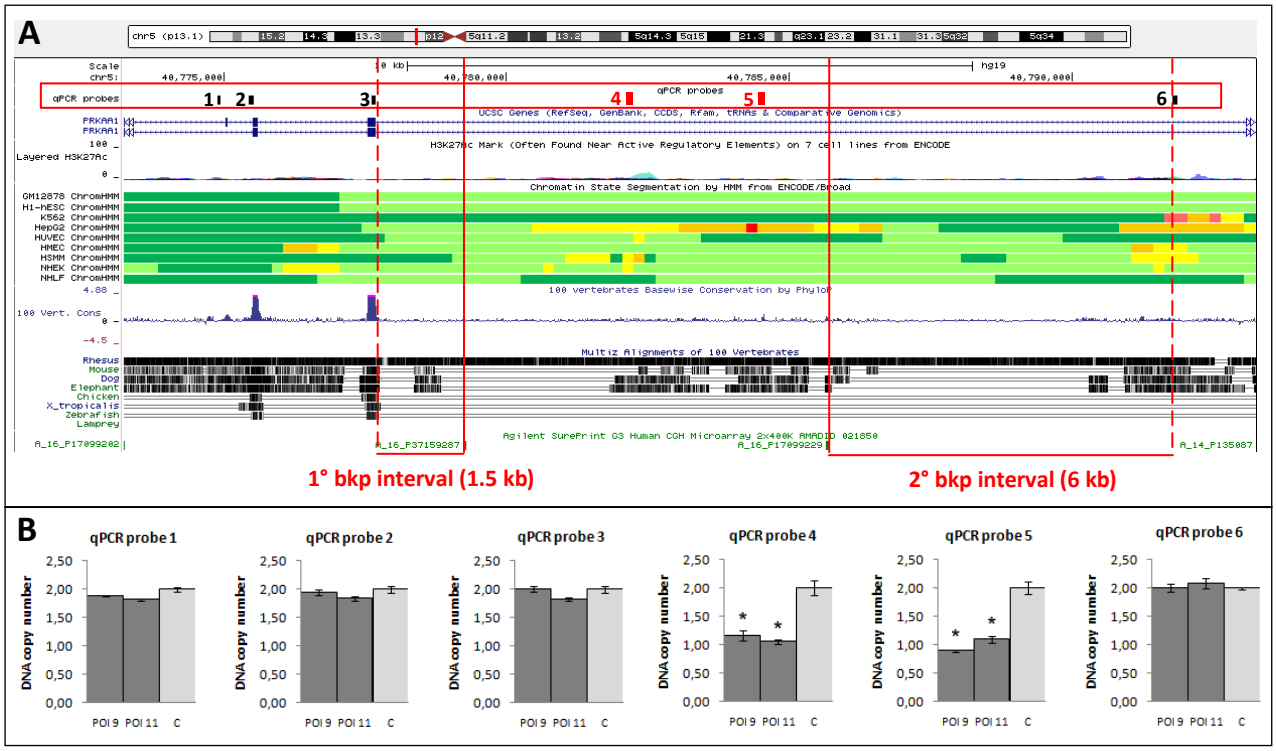

Fig.22. A) *UCSC genome browser view of the region involved in the loss CNV (chr5:40,773,231-40,793,258, hg19). UCSC genes are shown in blue. Regulatory elements prediction based on ENCODE chromatin state segmentation in nine human cell lines is shown: red, promoter; yellow, weak enhancer; orange, strong enhancer; dark green, transcriptional elongation; light green, weak transcribed. Conservation in other species as well as array CGH 400K probes are shown.* ***B)*** *qPCR results relative to each of the 6 probes used (*p<0.01). PCNT normalization is shown. C, control (light grey bars). POI patients are depicted in dark grey.*

To test whether the intronic deletion alter mRNA level, RT-qPCR analysis was performed on both sisters. Two TaqMan assays (ex1-2, ex9-10) were used and Ct values were normalized against three different housekeeping genes. Unfortunately, the results obtained failed to provide reproducible and interpretable data on blood tissue analyzed.

3.4.4 POI 44: partial gene deletion at 9p24.2 (*VLDLR*)

Array CGH analysis revealed a 14.6 kb deletion raging from probe A_16_P38628086 to A_16_P18531563 (chr9:2,654,203-2,668,811, hg19) and involving *VLDLR* 3'UTR (exon 19) and some predicted regulatory elements as enhancer and insulator (**Fig.23**).

Fig.23. UCSC Genome Browser view of the region involved in the CNV. *UCSC genes are shown in blue. Regulatory elements prediction based on ENCODE chromatin state segmentation in nine human cell lines is shown: light blue, insulator; yellow, weak enhancer; orange, strong enhancer; dark green, transcriptional elongation; light green, weak transcribed; light grey, heterochromatin. Conservation in other species as well as array CGH 400K probes are shown.*

Because the array probes which identified the rearrangement were 4, a validation was needed. Hence, exons 16, 17, and 18 resulted not covered by the array oligonucleotides. Thus, LR-PCR was carried out on patient DNA using primers specific for the deleted allele, taking into account the outer deletion raging from the non deleted probes A_16_P18531505 and A_16_P18531574 (black spots in **Fig.24A**). The amplicon was then sequenced and revealed that the entire coding sequence of *VLDLR* gene was preserved. Therefore, just 1 out of 4 array probes was not validated. The subsequent complete reading of the sequence amplified, revealed that four breakpoints occurred and generated two deletions of about 3 kb and 9.4 kb respectively, separated by 140 bp not involved in the loss (**Fig.24A**). Thus, the effective deletion includes only the predicted regulatory region downstream of *VLDLR*. The detailed sequence of the breakpoints junction is shown (**Fig.24A**).

Then, to figure out whether the deletion might have an effect on *VLDLR* gene expression by a "position effect", RT-qPCR was performed on patient cDNA: the analysis revealed a statistically significant transcript downregulation compared to controls (**Fig.24B**). Moreover, the identification on patients DNA of an heterozygous SNP (rs6148) in the *VLDLR* coding region (exon 14), allowed to point out a monoallelic mRNA expression of the gene, suggesting that the

deletion was likely responsible for the lack of expression of the deleted allele in the blood tissue analyzed (Fig.24B).

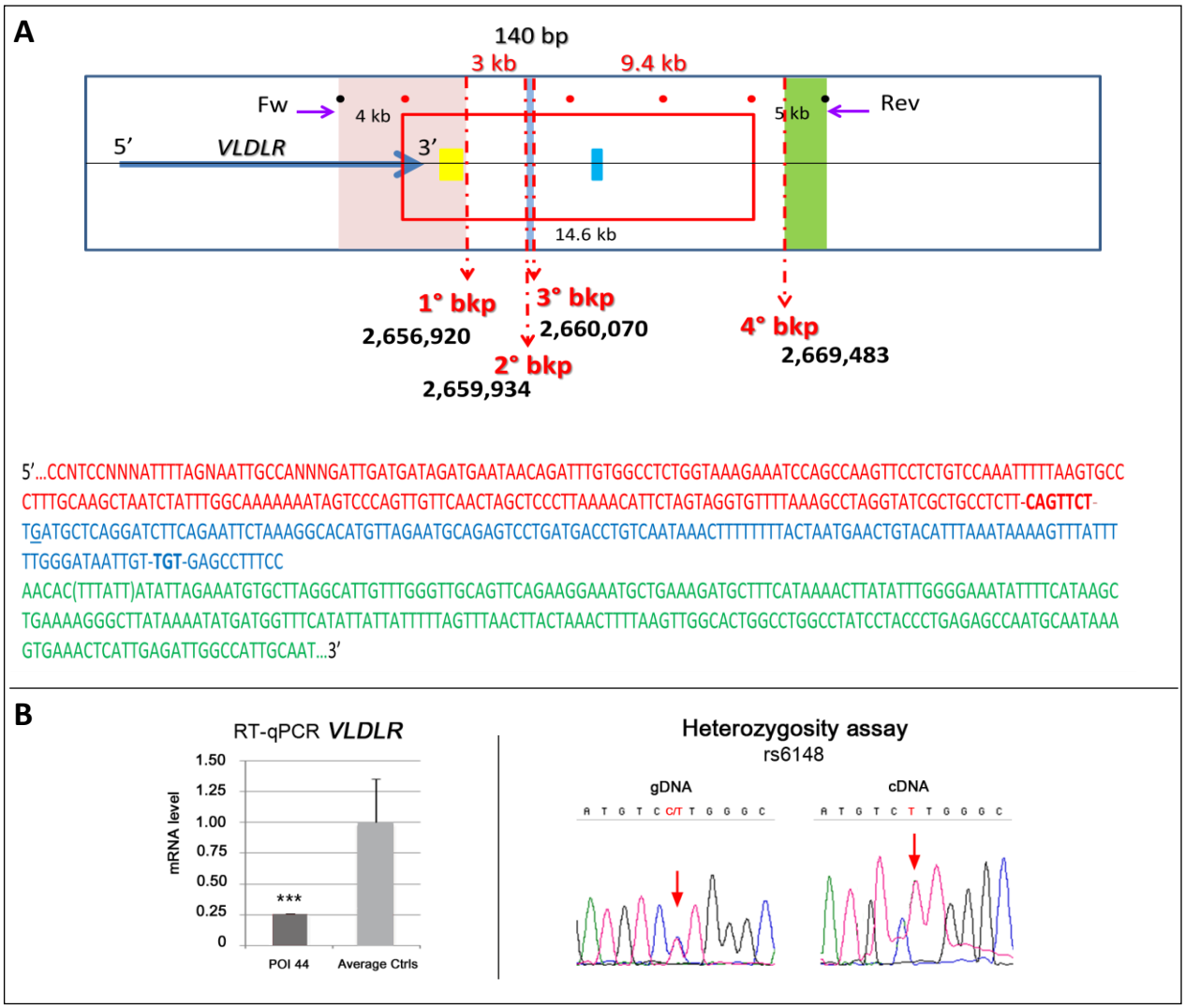

Fig.24. A) *Schematic view of the array result. The 14.6 kb deletion detected by array CGH which maches with the four deleted spots (in red) is represented by a red box. Inside the red box: 3' UTR of VLDLR is included, the enhancer element predicted is shown in yellow, the insulator in light blue. The array probes not deleted are shown as dark spots. Red arrows indicate the deletions breakpoints and the respective nucleotide position is written. Forward and Reverse primers used for LR-PCR are depicted as purple arrows. The regions not deleted after molecular characterization are highlighted in red, blue and green. Nucleotide sequence of the breakpoints junction is show: the sequence portions representative of the not deleted region (highlighted in A), are colored in red, blue, and green as well. The bases added and not present in the wild-type sequence are shown In bold and delimited by dashes. The bases included in parenthesis are not present in the breakpoints sequence contrary to the wild type sequence.* ***B)*** *RT-qPCR results on POI patient (dark grey) relative to VLDLR mRNA level. Average Controls levels are shown in light grey. *** p<0.001. Electopherogram showing the heterozygosity assay for SNP rs6148 identified in POI 44 (gDNA and cDNA sequences are reported).*

Therefore, to understand which is the role of the deleted genomic region, several luciferase assays were carried out using plasmid pGL3 (SV40 promoter) which was transfected in both granulosa (COV434) and HeLa cells. The cloning of different portions of the deleted region

(**Fig.25B**) revealed a strong enhancer role of the region including the predicted insulator and the following ~3000 bp (orange bars in **Fig.25C**); whereas the cloning of the only predicted insulator sequence, showed no signals (**Fig.25C**). This evidence was observed mainly in the COV analyzed (**Fig.25C**). Furthermore, the validation of the result was supported by cloning the *VLDLR* promoter region downstream the deleted region (insulator + 3000 bp) and upstream the reporter gene (plasmid pGL3b was used): a dramatic overexpression was observed (**Fig.26B**).

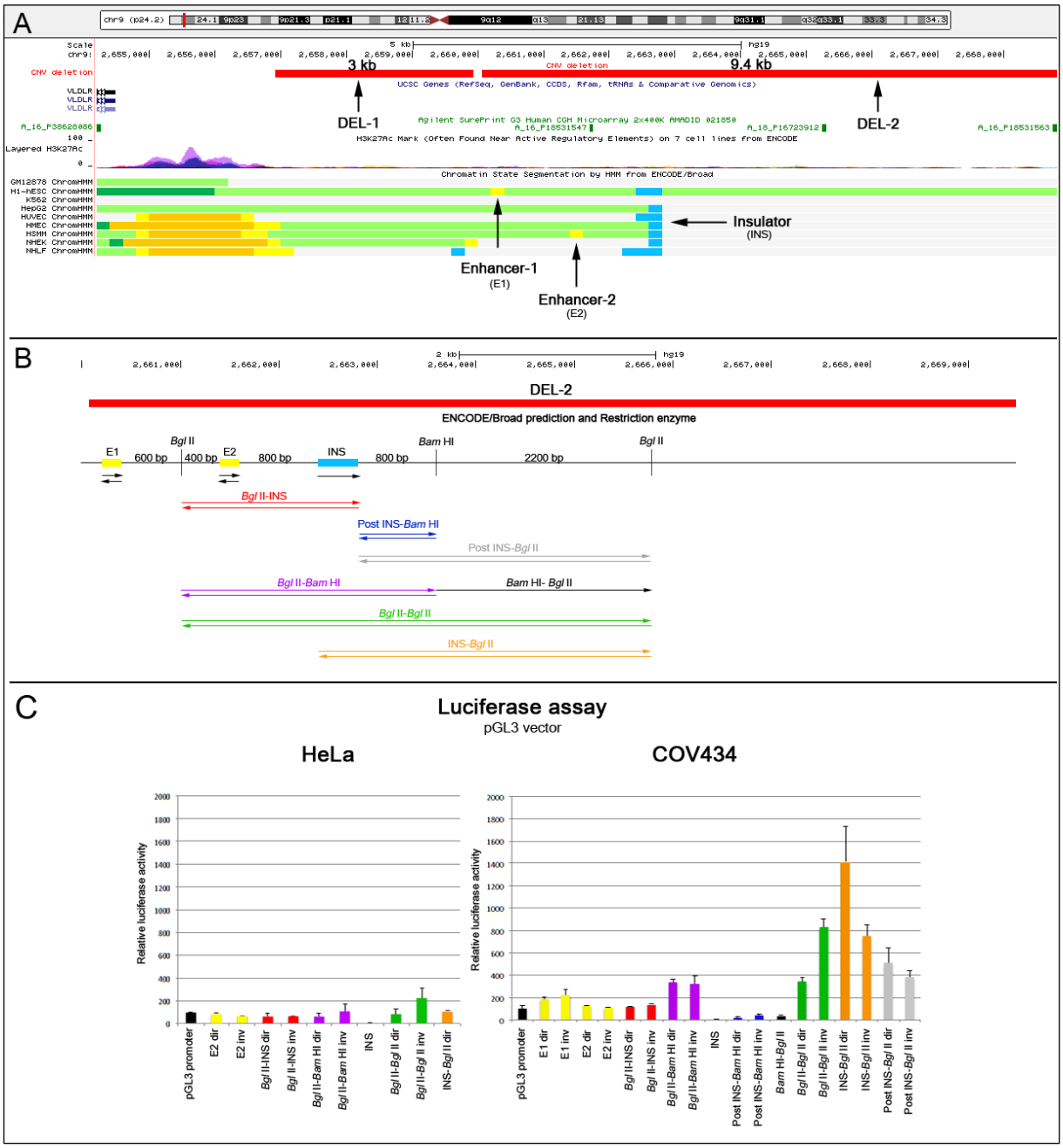

Fig.25. A) *UCSC Genome browser view of the deleted regions (DEL-1, DEL-2) involving the two predicted enhancer (in yellow) and insulator (in light blue).* ***B)*** *Schematic view of the luciferase constructs (arrows in different colors).* ***C)*** *Luciferase assay results using pGL3 vector (SV40 promoter) performed in both HeLa and COV434 cell lines.*

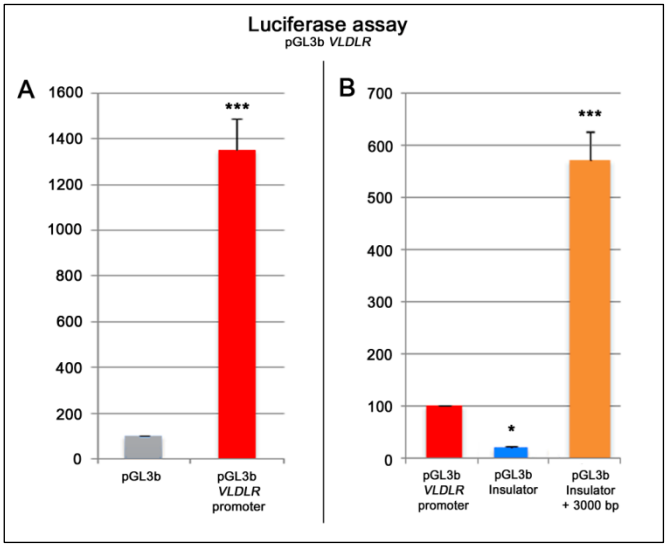

Fig.26. Results of Luciferase assay performed by transfecting plasmid pGL3b with VLDLR promoter region into COV434 cells. A) *Evaluation of the functionality of the region predicted as VLDLR promoter (in red). pGL3b control luminescence in grey.* ***B)*** *In red, luminescent signal of the VLDLR promoter region; in blue, luminescent signal of the insulator region cloned directly in pGL3b-VLDLR; in orange, luminescent signal of the insulator + 3000 bp construct cloned in pGL3b-VLDLR. *p<0.05; ***p<0.001.*

3.4.5 POI 46: intronic homozygous deletion at 15q25.2 (*CPEB1*)

The 8.7 kb homozygous deletion found in POI 46 and involving *CPEB1*, ranges from the array probe A_16_P40479359 to A_16_P03072633 (chr15:83,305,533-83,314,278, hg19; **Fig.27A**). Because the outer deletion extends from A_16_P40479345 to A_16_P03072640 (chr15:83,301,729-83,318,591, hg19), it was not possible to assess the involvement of the *CPEB1* promoter. Moreover, the region deleted in POI 46 was covered only by three oligonucleotide probes, thus making worthwhile a further validation.

The 17 qPCR assays used on POI DNA, allowed both to validate the CNV and to further characterize the homozygous deletion which does not involve the 5'UTR region but only part of intron 1 (**Fig.27A-B**). A subsequent Long Range PCR refined the loss at nucleotide level, which extends from 83,303,439 to 83,314,290, for a total of 10.8 kb deleted (**Fig.27A**).

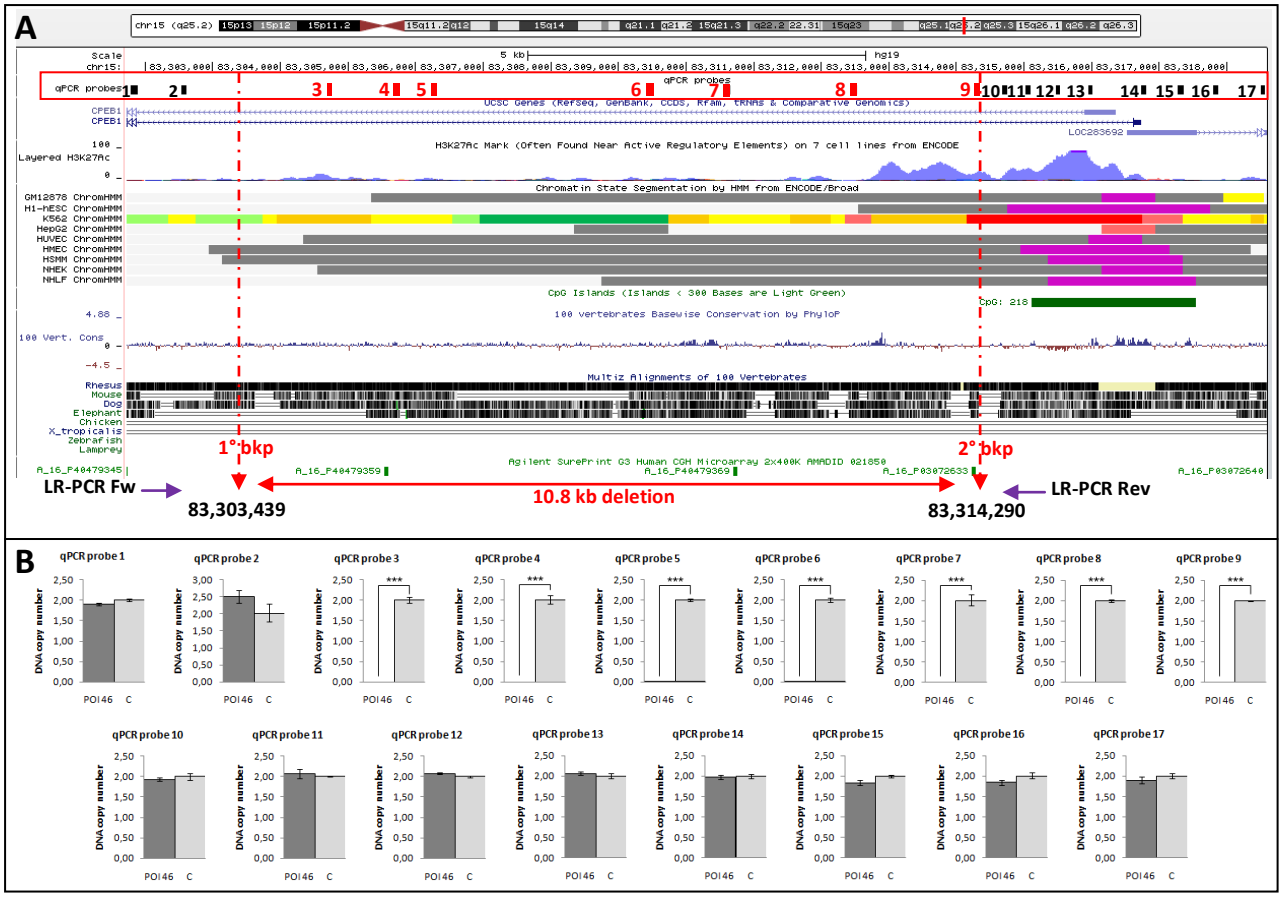

Fig.27. A) *UCSC Genome Browser view of the region involved in the CNV (chr15:83301729-83318591, hg19). Genes are shown in blue. Regulatory elements prediction based on ENCODE chromatin state segmentation in nine human cell lines is shown: purple, inactive promoter; red, active promoter; yellow, weak enhancer; orange, strong enhancer; dark green, transcriptional elongation; light green, weak transcribed; dark grey, polycomb-repressed; light grey, heterochromatin. CpG island, conservation in other species as well as array CGH 400K probes are shown. 17 qPCR probes are colored in black (not deleted) and red (deleted). Loss breakpoints are delimited by red arrows. LR-PCR primers are shown as purple arrows.* ***B)*** *qPCR results relative to each of the 17 probes used (***p<0.0001). PCNT normalization is shown. C, control (light grey bars). POI patient is depicted in dark grey.*

3.4.6 Common CNV: duplication and deletion at 10q26.3 (*SYCE1*)

The common CNV found both in POI patients and controls, involves *SYCE1* and *CYP2E1* (**Fig.28**). Considering that a block of segmental duplication (SD) is near the 5'UTR of *SYCE1*, the interpretation of the array data is not very informative for this region. Particularly, no 400K array probes cover the SD, whereas a 244K probe maps inside it but it is not informative. Thus it remains unclear whether *SYCE1* 5'UTR is involved or not in the CNV. Several qPCR experiments were carried out to assess the presence of putative differences in the type of rearrangement between patients and controls. They aimed at clarifying *SYCE1* 5'UTR complete or partial involvement in the CNV and consequently, its putative pathogenetic or susceptibility role in POI etiology. qPCR analyses revealed that *SYCE1* results completely included in the CNVs (duplicated or deleted) and that there are no differences between POI women and controls (Fig.28B).

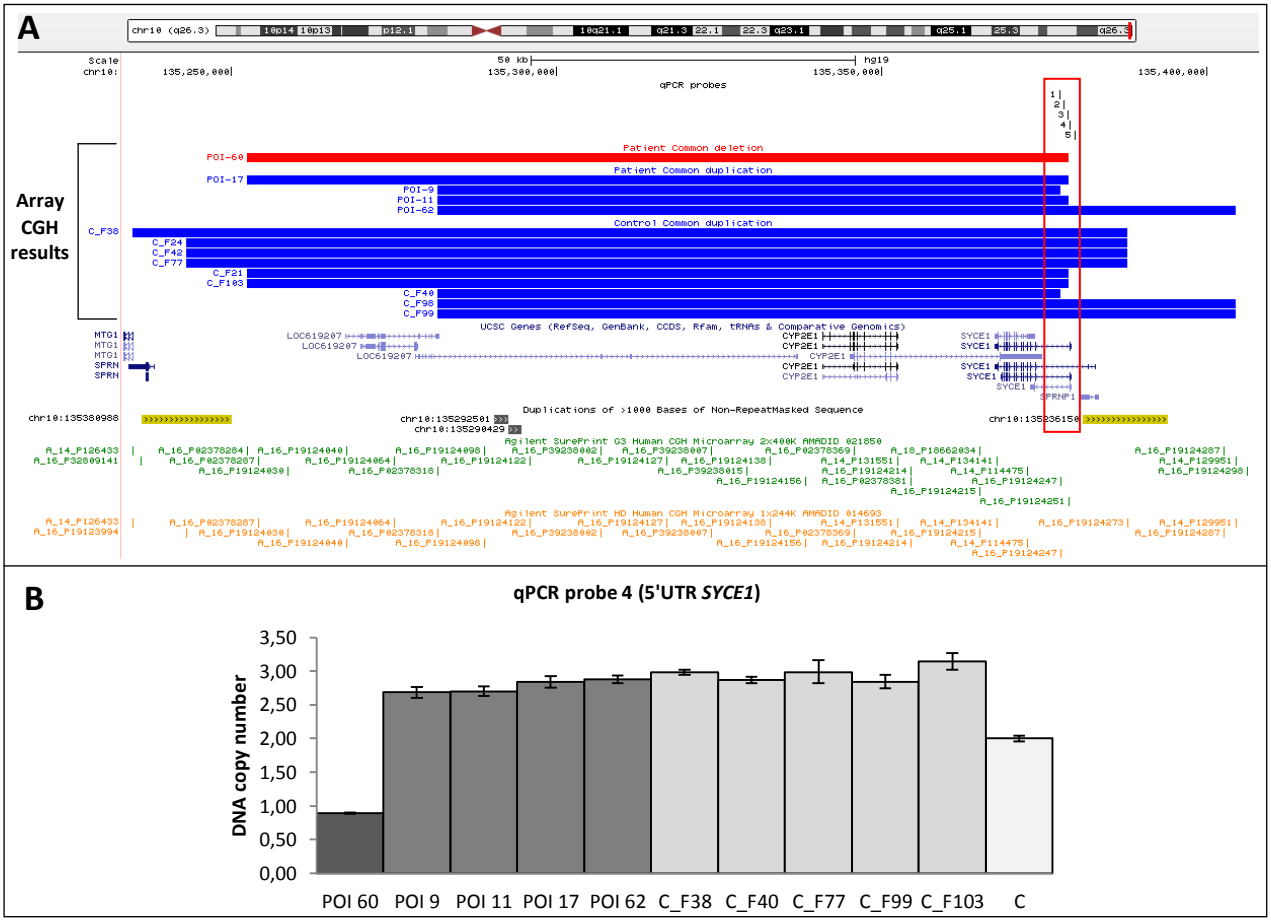

Fig.28. A) *UCSC Genome Browser view of the region involved in the common CNV (chr10:135,233,327-135,408,523, hg19). CNV deletion is shown as a red horizontal bar. CNVs duplications are shown as blue horizontal bars. UCSC genes are represented in blue. The red box highlights the qPCR probes used and the region to be investigated. Segmental duplication are colored in yellow and grey. Array CGH 244K and 400K probes are shown.* ***B)*** *RT-qPCR result of the assay number 4 which covers the 5'UTR of SYCE1. 5 out of 9 "duplicated" controls are shown. Dark grey, patients; light grey, controls; C, control without CNV in the region analyzed.*

3.5 Whole Exome Sequencing preliminary analysis

17 out of 67 primary amenorrhea patients (POI 48, 49, 50, 51, 52, 53, 54, 55, 57, 59, 60, 61, 62, 63, 65, 66, 67) were processed by exome targeted Next Generation Sequencing to discover rare mutations possibly implicated in the onset of their POI phenotype. In detail, the preliminary results of the exome sequencing data analysis consist in Single Nucleotide Variants (SNVs) filtering based on:

- **41** ovarian genes probably perturbed by rare and common CNVs identified after analyzing array CGH data (*AMFR*, *BMP15*, *BMP4*, *BNC1*, *BTBD1*, *CAPZA1*, *CPEB1*, *CYP2E1*, *DDX1*, *DIAPH2*, *DUSP22*, *EEF2*, *ELAVL2*, *FZD6*, *GRIA1*, *HBE1*, *LRIG2*, *MAP2K2*, *MOV10*, *MSL3*, *MYCN*, *NAIP*, *NRP1*, *PARD3*, *PCSK5*, *PRKAA1*, *RHOC*, *RREB1*, *RYR3*, *SH3GL3*, *SIRT6*, *SLC5A5*, *SLMAP*, *SMYD3*, *STS*, *SYCE1*, *TP63*, *TSHZ2*, *VLDLR*, *WHAMM*, *XPO1*);
- **135** genes derived from the collection of genes known to be involved in i) non-syndromic POI, or ii) syndromic form of POI, or iii) ovarian growth and development [see introduction] (*ADAMTS19*, *AFF2*, *AIRE*, *ALOX12*, *AMH*, *AMHR2*, *AR*, *ATM*, *BAX*, *BCL2L1*, *BDNF*, *BICD*, *BMP15*, *BMP2*, *BMP7*, *BMP8B*, *BMPR1A*, *BMPR1B*, *BMPR2*, *CDKN1A*, *CDKN1B*, *CGGBP1*, *CNOT6*, *CXCL12*, *CXCR4*, *CXCR7*, *CYP11A1*, *CYP17A1*, *CYP19A1*, *CYP1B1*, *DAZAP1*, *DAZL*, *DICER1*, *DMC1*, *DNAJC8*, *DNMT1*, *EIF2B*, *EIF2B1*, *EIF2B2*, *EIF2B3*, *EIF2B4*, *EIF2B5*, *EIF2C2*, *eIF4ENIF1*, *EPB41L5*, *EPS8*, *ESR1*, *FANCA*, *FANCC*, *FANCG*, *FGF8*, *FIGLA*, *FMR1*, *FOXE1*, *FOXL2*, *FOXO1*, *FOXO3*, *FOXO4*, *FSHB*, *FSHR*, *FST*, *GALT*, *GDF9*, *GDNF*, *GHR*, *GJA1*, *GJA4*, *GNAS*, *HDC*, *HFM1*, *INHA*, *INHBA*, *INHBB*, *KIT*, *KITLG*, *LHCGR*, *LHR*, *LHX8*, *LHX9*, *LSH*, *MCM8*, *MCM9*, *MGAT1*, *MOS*, *MSH4*, *MSH5*, *MSY2*, *NANOS3*, *NGF*, *NOBOX*, *NOG*, *NR5A1*, *NT4*, *NTRK1*, *NTRK2*, *NXF5*, *OCT4*, *PDPK1*, *PGRMC1*, *PIN1*, *PMM2*, *POF1B*, *POG*, *POLG*, *PTEN*, *PTHB1*, *PTPRO*, *RAD51C*, *RALB*, *REC8*, *RERG*, *RFPL4*, *RPA2*, *RPL10*, *RSPO1*, *SALL4*, *SERPINE1*, *SMAD1*, *SMAD3*, *SMAD5*, *SMPDL3B*, *SOHLH1*, *SOHLH2*, *SPO11*, *STAG3*, *STAR*, *STRAP*, *TAF4B*, *TIAR*, *TMEM35*, *TSC2*, *WNT4*, *XPNPEP2*, *ZAR1*, *ZFX*, *ZNF654*).
- **15** genes known to be interactors of the ***VLDLR*** (*APOE*, *APOER2*, *DAB1*, *FYN*, *RELN*), ***TP63*** (*CHUK*, *ITCH*, *MDM2*, *TP53*, *TP63*), and ***CPEB1*** (*APLP1*, *AURKA*, *BICC1*, *PLK1*, *PUM2*) genes. These data were obtained from STRING analysis.

The purpose of this approach was to find:

- mutations in the array genes in order to validate their role in the disorder;
- mutations in the known POI causative genes;

- mutations in new POI candidate ovarian genes.

Only exonic variants were considered and based on the filtering steps shown in **Fig.29** a total of 50 variants were selected.

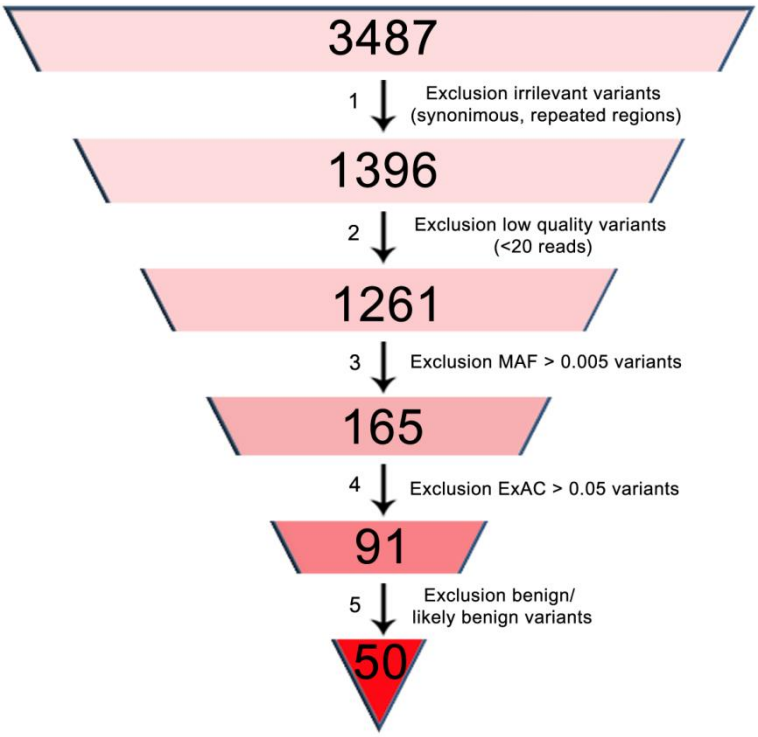

Fig.29. *Flowchart of the filtering steps (1-5) carried out in the selection of rare exonic SNVs in the 17 patients analyzed. The numbers in the boxes refer to the variants identified in the 191 genes.*

A detailed list of the 50 exonic SNVs is shown in **Tab.18**. It includes:

- 46 nonsynonimous SNVs;
- 1 frameshift insertion leading to a premature stop codon;
- 1 frameshift deletion leading to a premature stop codon;
- 1 nonframeshift deletion;
- 1 nonsense mutation.

19 out of 50 variants have never been reported so far in the databases investigated. Moreover, considering that 10 variants were reported as "low quality" sequencing data, Sanger validation was performed. No one of the 10 SNVs was confirmed.

Up to now, other 6 variants were processed by Sanger sequencing (in progress analysis) and for all, validation was obtained (**Tab.18**).

Tab.18. Pathogenic or likely pathogenic SNVs found in the 17 POI patients analyzed by WES

ID	Chr	Start[#]	End[#]	Ref	Alt	Gene	ExonicFunction	AAChange	1000G	ExAC	dbSNP*	HGMD	P1	P2	P3	V
48	chr6	121768925	121768925	C	-	*GJA1*	frameshift deletion	NM_000165:exon2:c.932delC:p.A311fs	NA	0,0072	rs778110855	/	/	/	/	
	chr12	124111657	124111657	T	C	*EIF2B1*	nonsynonymous SNV	NM_001414:exon5:c.A416G:p.E139G	NA	NA	/	/	T	P	D	
49	chr2	120776755	120776755	A	G	*EPB41L5*	nonsynonymous SNV	NM_001184937:exon2:c.A95G:p.H32R	NA	2,47E-05	rs138231769	/	D	D	D	
	chr17	7192970	7192970	G	A	*YBX2*	nonsynonymous SNV	NM_015982:exon7:c.C923T:p.P308L	NA	8,34E-06	rs756982334	/	T	B	D	
	chr19	4174745	4174745	C	A	*SIRT6*[LQ]	nonsynonymous SNV	NM_001193285:exon7:c.G856T:p.A286S	0,001	0,0049	rs183444295	/	T	B	N	-
50	chr2	219254672	219254672	T	C	*SLC11A1*	nonsynonymous SNV	NM_000578:exon9:c.T875C:p.L292P	0,0006	0,0009	rs145536242	/	D	D	D	
	chr15	83481979	83481979	C	T	*WHAMM*	nonsynonymous SNV	NM_001080435:exon2:c.C734T:p.P245L	NA	NA	/	/	D	D	D	+
	chr17	56774110	56774110	A	G	*RAD51C*[LQ]	nonsynonymous SNV	NM_058216:exon3:c.A461G:p.E154G	NA	0,0404	rs758847241	/	D	P	D	-
	chrX	84569452	84569452	G	A	*POF1B*	nonsynonymous SNV	NM_024921:exon9:c.C943T:p.R315C	0,0003	0,0008	rs149085556	/	D	D	D	+
51	chr1	113242873	113242873	A	G	*MOV10*	nonsynonymous SNV	NM_001130079:exon20:c.A2831G:p.Y944C	NA	NA	/	/	T	B	D	
	chr15	89868870	89868870	G	A	*POLG*	nonsynonymous SNV	NM_001126131:exon10:c.C1760T:p.P587L	0,0008	0,0017	rs113994096	CM031330	D	D	A	
	chr15	89873415	89873415	G	A	*POLG*	nonsynonymous SNV	NM_001126131:exon3:c.C752T:p.T251I	0,0008	0,0017	rs113994094	CM021660	T	B	A	
	chr16	89842176	89842176	C	G	*FANCA*	nonsynonymous SNV	NM_000135:exon21:c.G1874C:p.C625S	0,0012	0,0028	rs139235751	CM112454	D	D	D	
	chr19	2251240	2251240	T	C	*AMH*	nonsynonymous SNV	NM_000479:exon5:c.T967C:p.F323L	0,0002	0	rs577002391	/	T	B	N	
52	chr5	42700021	42700021	C	T	*GHR*	nonsynonymous SNV	NM_001242460:exon4:c.C469T:p.R157C	0,0014	0,0041	rs121909362	CM930313	D	B	A	
53	chr2	237490111	237490111	T	G	*ACKR3*	nonsynonymous SNV	NM_020311:exon2:c.T1003G:p.S335A	NA	NA	/	/	/	B	D	
	chr6	108985176	108985176	-	G	*FOXO3*	frameshift insertion	NM_001455:exon2:c.1141dupG:p.D380fs	NA	NA	rs34133353	/	/	/	/	
	chr15	89865023	89865023	C	T	*POLG*	nonsynonymous SNV	NM_001126131:exon16:c.G2542A:p.G848S	0,0002	0,0002	rs113994098	CM021662	D	D	A	
54	chr1	113232516	113232516	C	A	*MOV10*	nonsynonymous SNV	NM_001130079:exon5:c.C632A:p.P211Q	NA	NA	/	/	T	D	D	
	chr4	48496243	48496243	C	G	*ZAR1*	nonsynonymous SNV	NM_175619:exon4:c.C1257G:p.S419R	0,0002	0,0004	rs150750718	/	D	D	D	
	chr6	72111150	72111150	C	A	*RREB1*[LQ]	stopgain	NM_001003698:exon7:c.C539A:p.S180X	NA	NA	/	/	/	/	/	-
	chr7	103197600	103197600	G	T	*RELN*	nonsynonymous SNV	NM_005045:exon38:c.C5621A:p.P1874Q	NA	NA	/	/	T	D	D	
	chr9	78953201	78953201	G	T	*PCSK5*[LQ]	nonsynonymous SNV	NM_001190482:exon34:c.G4723T:p.G1575W	NA	NA	/	/	D	/	N	-
	chr12	15652521	15652521	C	G	*PTPRO*	nonsynonymous SNV	NM_002848:exon4:c.C654G:p.H218Q	NA	NA	/	/	T	D	D	
55	chr1	91846517	91846517	G	T	*HFM1*[LQ]	nonsynonymous SNV	NM_001017975:exon7:c.C825A:p.F275L	NA	NA	/	/	D	D	D	-
	chr5	128863507	128863507	C	A	*ADAMTS19*[LQ]	nonsynonymous SNV	NM_133638:exon5:c.C1135A:p.L379M	NA	NA	/	/	/	D	D	-
	chr6	31138093	31138093	C	A	*POU5F1*[LQ]	nonsynonymous SNV	NM_002701:exon1:c.G305T:p.G102V	NA	NA	/	/	T	P	N	-

Tab.18. Continued

	chr7	103474110	103474110	G	T	*RELN*[LQ]	nonsynonymous SNV	NM_005045:exon3:c.C347A:p.S116T	NA	NA	/	/	D	P	D	-
	chr15	84255764	84255764	C	A	*SH3GL3*[LQ]	nonsynonymous SNV	NM_001301108:exon7:c.C447A:p.F149L	NA	NA	/	/	T	B	D	-
	chr16	2121909	2121909	C	T	*TSC2*	nonsynonymous SNV	NM_000548:exon19:c.C2071T:p.R691C	NA	6,94E-05	rs760489473	/	D	D	D	
	chrX	128884450	128884450	C	T	*XPNPEP2*	nonsynonymous SNV	NM_003399:exon8:c.C644T:p.T215I	0,0013	0,0044	rs138365897	/	D	B	D	
57	chr13	41134286	41134286	T	C	*FOXO1*	nonsynonymous SNV	NM_002015:exon2:c.A1342G:p.S448G	NA	NA	/	/	T	B	D	
	chr15	33999240	33999240	G	A	*RYR3*	nonsynonymous SNV	NM_001036:exon43:c.G6604A:p.A2202T	NA	NA	/	/	T	D	D	+
	chr19	3984203	3984203	C	T	*EEF2*	nonsynonymous SNV	NM_001961:exon2:c.G149A:p.R50Q	NA	NA	/	/	T	B	D	
59	chr1	156849863	156849863	G	A	*NTRK1*	nonsynonymous SNV	NM_001012331:exon15:c.G2101A:p.E701K	NA	8,28E-06	rs747855434	/	D	D	D	
	chr2	38298394	38298394	C	T	*CYP1B1*	nonsynonymous SNV	NM_000104:exon3:c.G1103A:p.R368H	0,0042	0,0062	rs79204362	CM000137	D	D	A	
60	chr9	97879642	97879642	A	C	*FANCC*	nonsynonymous SNV	NM_000136:exon11:c.T1027G:p.Y343D	NA	NA	/	/	T	D	N	
	chrX	147026489	147026489	C	T	*FMR1*[LQ]	nonsynonymous SNV	NM_001185081:exon13:c.C1238T:p.A413V	0,0005	0,0012	rs143889976	/	T	/	D	-
61	chr15	83932648	83932648	G	A	*BNC1*	nonsynonymous SNV	NM_001301206:exon4:c.C1334T:p.P445L	NA	8,24E-06	rs778369152	/	D	P	D	
62	chr19	36361822	36361822	C	G	*APLP1*	nonsynonymous SNV	NM_001024807:exon3:c.C316G:p.R106G	0,0002	0,0002	rs144561889	/	D	D	D	
63	chr15	51520044	51520044	T	C	*CYP19A1*	nonsynonymous SNV	NM_000103:exon4:c.A383G:p.H128R	NA	5,77E-05	rs375975652	/	D	D	D	+
65	chr9	34648433	34648433	C	T	*GALT*	nonsynonymous SNV	NM_001258332:exon5:c.C340T:p.R114C	NA	2,47E-05	rs111033750	/	D	D	D	+
	chr16	8941654	8941654	G	A	*PMM2*	nonsynonymous SNV	NM_000303:exon8:c.G713A:p.R238H	NA	0,0006	rs151319324	/	T	D	N	
	chr17	56811542	56811542	A	G	*RAD51C*	nonsynonymous SNV	NM_058216:exon9:c.A1090G:p.S364G	NA	NA	rs587782565	/	D	B	N	
	chr19	45412314	45412314	T	A	*APOE*	nonsynonymous SNV	NM_000041:exon4:c.T761A:p.V254E	NA	0,0014	rs199768005	/	D	P	N	
66	chr18	23872246	23872246	G	A	*TAF4B*	nonsynonymous SNV	NM_001293725:exon8:c.G1642A:p.V548M	NA	NA	/	/	T	D	D	
	chr20	50407448	50407448	C	T	*SALL4*	nonsynonymous SNV	NM_020436:exon2:c.G1574A:p.G525E	NA	8,25E-06	rs746415699	/	T	P	D	
	chrX	24197772	24197774	AGA	-	*ZFX*	nonframeshift deletion	NM_001178084:exon3:c.531_533del:p.177_178del	NA	2,28E-05	rs762873800	/	/	/	/	
67	chr9	2643709	2643709	G	A	*VLDLR*	nonsynonymous SNV	NM_001018056:exon6:c.G902A:p.R301Q	0,0004	0,0027	rs139671268	/	T	D	D	+
	chr15	89869870	89869870	C	T	*POLG*	nonsynonymous SNV	NM_001126131:exon9:c.G1685A:p.R562Q	NA	NA	rs781168350	CM033443	D	D	D	

[#]Physical position of the identified SNVs based on the UCSC Genome Browser, hg19, released February 2009
*dbSNP release 144
LQ: Low Quality site
P1: SIFT prediction (T, tolerated; D, damaging)
P2: Polyphen2 prediction (B, benign; P, possibly damaging; D, probably damaging)
P3: MutationTaster prediction (N, polymorphism; D, disease causing; A, disease causing automatic)
V: Sanger validation (in progress analysis); +, validated; -, not validated
NA, not available
In grey: SNVs detected in array-CGH genes

Among the 50 SNVs, 11 fall in the group of the candidate ovarian genes detected by array-CGH: 4 resulted false-positive (low quality SNVs), 3 were positive (*WHAMM* in POI 50, *RYR3* in POI 57, *VLDLR* in POI 67), and the remaining 4 SNVs have validation in progress.

To evaluate the presence of positive *VLDLR* SNVs in the other 50 patients of the present cohort, exon 6 sanger sequencing was performed. rs139671268 was further found in POI 46 for a total of two positive patients. Frequencies of the identified SNVs are reported below in **Tab.19**.

Tab.19. Frequencies relative to rs139671268

Total Patients	Positive patients	Allelic frequency	Allelic frequency 1000 Genomes	Allelic frequency ExAC	Genotype frequency	Genotype frequency 1000 Genomes	Genotype frequency ExAC
67	2	0,015 (2/134)	0,001 (1/1006)	0,0034 (229/66730)	0,03 (2/67)	0,002 (1/503)	0.007 (229/33365)

Interestingly, the analysis performed on POI 67's parents showed the presence of *VLDLR* SNVs also in her mother who suffered of premature menopause.

4. DISCUSSION

4.1 Sample cohort selection and investigation

Primary Ovarian Insufficiency is an extremely heterogeneous disorder with a great impact on women's life, not only from the perspective of reproduction/fertility but also from a general point of view, as the young women have to deal with problems arising from low estrogen levels. The etiology of the disease is complex and encompasses infectious, metabolic, and autoimmune causes, iatrogenic treatments, and environmental factors. Also the genetic causes are thought to be prevalent in the pathogenesis of idiopathic POI patients, as supported by the occurrence of families with several affected women, corresponding to an incidence from 4 to 31%, depending on the studied population [10,60]. Thus far, genetic screening of *FMR1* premutations is critical to evaluate non-syndromic forms of POI. Also the assessment of patient's karyotype is crucial in order to identify X chromosome abnormalities and therefore exclude the presence of syndromic conditions of POI (e.g., Turner Syndrome).

Several causative genes for POI onset are already known (*BMP15*, *GDF9*, *HFM1*, *DIAPH2* etc.). However, the existence of idiopathic cases suggests that other genes should be involved. Thus, it appears more and more obvious that POI is a very heterogeneous disease and that the molecular pathways underlying women's reproductive life are extremely complex and not yet clarified.

Based on these evidences, the aim of the present thesis work was to analyze by high resolution array-CGH (244K and 400K) a cohort of stringently selected non-syndromic patients affected by the most severe phenotype, namely 67 women with PA. The analysis might allow to identify CNVs possibly implicated in POI pathogenesis, involving genes with a role in the early phases of ovarian follicles growth and development. These CNVs, if disregulated, are supposed to have a more deleterious effect. Up to date, there are five publications in which array-CGH technique (whole genome or custom for X chromosome) was used to identify rare CNVs possibly implicated in POI etiology [159,160,163,164,165]: however the sample cohorts screened have always been phenotypically "mixed", including both PA and SA affected women. Only the study by Liao *et al.* [153] selected women with a POI extreme phenotype (PA), although it applied genome-wide association as technological approach.

The cohort recruited for the current study consists of 53 sporadic (79.1%) and 14 familial (20.9%) cases including six sisters pairs, and two patients with relatives affected by SA (not included in the analyzed cohort). The familial cases will support the importance of the genetic component and therefore may help to identify more easily the molecular defect. At the same time the analysis of a large group of sporadic cases could help to discover defects in the same genes, thus supporting their role in the onset of the disease.

In the present work the identification of rare CNVs, never reported or reported at a very low frequency, might allow to identify high-penetrance variants with a more severe impact. However, it cannot be excluded that "common" CNVs may have modulated the phenotype of the reported POI patients, increasing the complexity of inter-individual genetic variability. It is for this reason that "common" CNVs previously associated with POI [152,161] were also valued.

Another point deserving attention is the control group to which all the CNVs data were compared. The DGV tool, which contains CNVs reported in healthy controls of different populations whose gender and age are unknown, does not match with the ideal set of population control for POI disease. For this reason, as done in the study of Norling *et al.* [159], an *ad hoc* control cohort was considered. This cohort includes 140 healthy women with normal reproductive life and physiological menopause, who have been previously analyzed with the same method of patients.

First, rare CNVs according to the DGV were selected and those involving "ovarian" genes were considered. Secondly the selected CNVs were investigated in the control group to determine whether there was a real enrichment in ovarian genes in the POI cohort [see chapter 2].

The evidence that 39 patients resulted negative for rare ovarian CNVs (58.2%), suggested to combine different genomic approaches in order to increase the detection rate of the disorder. In the present work, 17 out of 67 collected patients, were then submitted to WES analysis. Particularly, the preliminary WES screening attempted to highlight i) the presence of rare variants involving the array-CGH candidate genes in order to support their role in POI etiology; ii) the presence of likely pathogenic variants in genes already associated to POI or in genes which play a role in ovarian maturation and differentiation.

To our knowledge this is the first study in which a combined approach of high-resolution array CGH and WES is used to analyze a cohort of women affected by the most severe POI phenotype.

4.2 Array-CGH: rare CNVs detection and *ad hoc* control group analysis

28 out of 67 women resulted positive to array-CGH analysis because having at least one rare CNV involving genes with a role in oocyte maturation and differentiation, for a total number of 32 CNVs. Thus the detection rate of the method resulted 48.3% which is comparable to that reported in two whole genome array-CGH publications by Norling *et al.* (42.3%: 11 rare ovarian CNVs in 11/26 patients [159]), and Ledig *et al.* (41.9%: 44 rare CNVs in 31/74 patients [160]).

Based on array-CGH results, 4 patients (POI 13, 39, 46, 54) have more than one rare "ovarian" CNV. Moreover, 14 women positive for rare "ovarian" CNVs (POI 6, 9, 11, 13, 33, 39, 42, 44, 48, 50, 52, 53, 60, 62), have at least one common CNVs with a putative susceptibility role (**Tab.10**). These evidences support the existence of a genetic model characterized by oligogenic heterozygosity i.e., the simultaneous presence in a single patient of multiple heterozygous quantitative variants/rare mutations, both *de novo* and/or inherited, affecting multiple genes. In this model, when more than one rare CNVs is found in a single patient, it is difficult to attribute a "major" causative role to one of them.

To better understand their contribution in POI onset, the same array-CGH data analysis was performed on *ad hoc* controls population composed by 140 healthy women. The majority of rare "ovarian" CNVs detected in patients were not found in the control group thus supporting their association with POI. Only 4/32 rare CNVs involving *STS*, *MSL3*, *XPO1*, *MOV10* genes were found also in controls (**Tab.15**). Particularly, the CNVs including *STS* have identical extension both in patients and controls, whereas the CNVs involving the other three genes have no recurrent breakpoints. Hence, their causative role remain unclear, but considering POI onset as an oligogenic heterozygosity model, their involvement cannot be ruled out.

Moreover, to evaluate the consistency of our methodological approach, rare "ovarian" CNVs were searched in the control population too. In 44 women, 49 rare "ovarian" CNVs were found for a detection rate of 31.4% which is lower if compared to the patients' one (48.3%).

Furthermore, to evaluate the presence of a significant enrichment in "ovarian" CNVs in the POI group, several statistical analyses were performed comparing patients' to controls' data (number of positive and negative people to array-CGH; number of exonic, non-exonic CNVs; number of genes with higher and lower association score), revealing no differences. This finding, which was never assessed in the previous studies, emphasizes that POI is not a disease statistically enriched in "ovarian" CNVs, as it is reported for other disorders (e.g.: autism, schizophrenia, etc.)[156,157,158]. Notewortly, regarding the type of rearrangement (certain or hypothetically

deleterious) involving the ovarian genes, a difference between patients and controls can be observed: if the genes with the highest POI association score are considered (score 3,4), the CNVs assumed to be more harmful involve 5 genes in patients' group (*BMP15*, *CPEB1*, *TP63*, *WHAMM*, *BNC1*) but only 1 gene in controls' group (*PTN*, growth factor expressed in bovine and mouse follicular fluid during oocyte development [228,229]) (**Tab.17**). In detail, the most deleterious rearrangements considered, were:

- i) complete gene deletion and partial gene deletion that might lead to haploinsufficiency (*BMP15*, *CPEB1*, *WHAMM*, *BNC1*, *PTN*); ii) intragenic duplication/deletion that might lead to a new gene formation with duplicated or deleted exons, possibly resulting in aberrant mRNA (if translated in aberrant protein) or leading to haploinsufficiency (*TP63*).

Moreover, some of the controls' genes with high score (3,4; **Tab.17**) and an already reported POI association, are not directly involved by the CNVs, but might be disrupted by a position effect. This is a condition that does not always occur, it relies on the regulatory elements eventually contained in the CNVs, and it is difficult to predict if they have a real effect on the adjacent genes [230,231]. For instance it referes to *LHCGR* and *FSHR* (gonadotropin receptors) as well as *GREM2* (BMP antagonist with a role inregulating ovarian primordial to primary follicle transition [232]), *FMN2* (formin-2 with a role in regulating actin assembly during meiosis [233]), *IL6ST* (membrane receptor involved in ovarian steroid metabolism [234]), *ADAMTS19* (metalloproteinase whose SNPs were associated to POI [148,150]), and *AKAP2* (protein kinase involved in cAMP signaling implicated in meiotic arrest e meiotic resumption [235]).

Besides position effect, the other less harmful CNVs considered were:

- i) complete duplication leading to an increase of the corresponding protein, as well as mild or null effect; ii) partial gene duplication for which it is difficult to predict the real effect of the rearrangement on the involved genes, as it depends on the orientation in which the duplicated portion is relocated: in case of a tandem direct orientation, mild or null perturbation of the gene expression might be expected; iii) intronic duplication/deletion which generally does not result in a deleterious effect unless splicing regulatory elements are embedded.

Regarding the "common" CNVs identified, no difference about the frequency of these variants in patients compared to controls (38.8% vs 32.1%) was found confirming the data reported in Castronovo *et al.* [67]. However, as mentioned before, it cannot be excluded a susceptibility role in combinations with other common/rare CNVs, then supporting the oligogenic heterozygosity model.

4.3 Gene content analysis and CNVs characterization: genotype-phenotype correlations

The finding of 37 genes likely to be involved in POI etiology, highlights and confirms the high genetic heterogeneity associated with the disorder. Indeed, the same affected loci were detected only in few unrelated patients (*DIAPH2* in patients 6 and 55, *CPEB1* in patients 46 and 54, and *TSHZ2* in patients 52 and 62) and the presence of 4 patients having more than one rare CNVs, further corroborates the genetic complexity of POI.

A classification based on their known/possible role in the pathogenesis was made, giving a score from 1 (low association) to 4 (high) (**Tab.12**). Most of the selected genes have never been reported in association with POI (score 1, 2, 3), and thus they are suggested as new candidate loci. Among them, the genes with score 3 have strong molecular evidences supporting their involvement in the disorder (e.g.: *TP63*, *VLDLR*, *WHAMM*, *BNC1*, *PRKAA1*, *BMP4*). Whereas few genes with an already known association with POI were found (score 4, e.g.: *DIAPH2*, *BMP15*, *CPEB1*).

The identified genes were then divided in 4 classes according to their function (**Fig.9**) (genes with the highest POI score are marked in bold):

- genes involved in ovarian metabolism and homeostasis (19%: *RYR3*, ***PCSK5***, *SLC5A5*, ***VLDLR***, *AMFR*, *STS*, *BTBD1*): e.g. ***PCSK5***, involved in the processing of mouse ovarian inhibin subunits during folliculogenesis [199,200], and ***VLDLR,*** which plays an important role in cholesterol and lipoprotein endocytosis in ovarian follicle [202,203,204,205,236];
- genes with a role in the intracellular trafficking, and in the regulation of cytoskeleton dynamics for oocyte asymmetric division (21.6%: ***DIAPH2***, *RREB1*, *RHOC*, *CAPZA1*, *XPO1*, ***PARD3***, ***WHAMM***, *SH3GL3*): e.g. ***DIAPH2***, causative of POF2A (OMIM #300511), encodes a protein associated with cytoskeleton and involved in metaphase chromosome alignment during oogenesis [129,175,176]; ***PARD3,*** known to play a role in establishing asymmetric division in mouse and Xenopus oocytes [206,207]; ***WHAMM,*** required for meiotic spindle migration and asymmetric cytokinesis in mouse oocytes [217];
- genes involved in transcription and translation regulation, chromatin remodelling, and maintenance of oocyte genomic integrity (35.1%: ***TP63***, *SMYD3*, *MOV10*, *LRIG2*, *DDX1*, *NRP1*, ***CPEB1***, *TSHZ2*, *SLMAP*, ***BNC1***, *MSL3*, *EEF2*, *SIRT6*): e.g. ***TP63,*** plays a role in the maintainance of oogonia genomic integrity and may be implicated in meiosis and cell

cycle control of germ cells in the mouse ovary [179,180,181,182,237]; ***CPEB1***, plays a role in regulating mRNA translation of proteins involved in synaptonemal complex formation during oocyte maturation [209,210,211]; ***BNC1***, may play a role in the differentiation of mouse oocytes increasing transcription of the ribosomal RNA genes during mouse oogenesis [219,220];

- genes playing a role in ovarian differentiation, follicular development, and meiotic resumption (24.3%: *ELAVL2*, ***PRKAA1***, *HBE1*, ***BMP15***, ***BMP4***, *MYCN*, *GRIA1*, *FZD6*, ***MAP2K2***): e.g. ***PRKAA1,*** involved in murine and porcine oocyte meiotic resumption [184,185,186,187,188]; ***BMP15,*** plays a role in oocyte maturation and follicular development as homodimer or heterodimers with a related protein (Gdf9); moreover, several mutations were found in POI women [94,95,96,97,98]; ***BMP4,*** plays a role in follicular development and differentiation acting as a transition factor from primordial to primary mouse follicle [192,193,194,195,196]; ***MAP2K2,*** regulates spindle formation and meiotic resumption during mouse oocyte development [223,224,225].

Although these selected ovarian genes have an interesting function likely and/or hypothetically related to POI etiology, it is also important to validate and characterize the CNVs affecting them, as well as to study their possible perturbation at mRNA level, in order to perform a correct genotype-phenotype correlation. Hereafter some of the "ovarian" rare CNVs detected are discussed.

In **POI 6** an intronic duplication, inherited from the father, between exons 26-27 of *DIAPH2* at Xq21.33 was detected. Inside the duplicated region, an isoform of *DIAPH2*-antisense1 (*RBSG1*) was found, resulting fully duplicated.

Considering that *DIAPH2* is an already reported POI gene with a peculiar action in female gonads [74,78,79,129,238], it is most likely that the CNVs contributed to the patients phenotype. Although intronic duplications are considered less harmful rearrangements than exonic duplications, it is possible that in POI 6 the antisense mRNA duplication resulted in a deleterious effect on the target gene.

Based on the function of mRNA-antisense in modulating the transcription of the complementary gene, Real Time RT-PCR analysis was performed on *DIAPH2* mRNA, but no differences between POI 6 and controls were detected in the peripheral blood tissue analyzed. *DIAPH2* maps on X chromosome and the controls used in the analysis were both female and male. RT-qPCR showed the same mRNA levels between female and male controls indicating an X inactivation (preferential or casual) of the gene in blood. Moreover, the father who has the

same X-linked CNV, did not show any transcript downregulation in the tissue analyzed, highlighting that in blood the CNV might not result in transcript perturbation. However, based on these evidences and taking into account that the gene may escape the X inactivation [239], it cannot be excluded that mRNA alterations occurred in other tissues, mainly in the early stages of development in the tissues involved in ovarian growth.

A paternal intragenic duplication at 3q28 involving *TP63* gene was found in sisters **POI 8 and 10 (Fig.30)**. The characterization of the CNV on POI 10 gDNA was performed by Long-Range PCR and Sanger sequencing, revealing the involvement from exon 2 to exon 10 of the main isoform. This molecular deepening was crucial to understand that the resultant allele has a direct tandem duplication from exon 2 to 10; particularly, an overlapping region of 11 bp (present both in intron 11 and 1) corresponds to the breakpoint junction of the rearrangement (**Fig.17**).

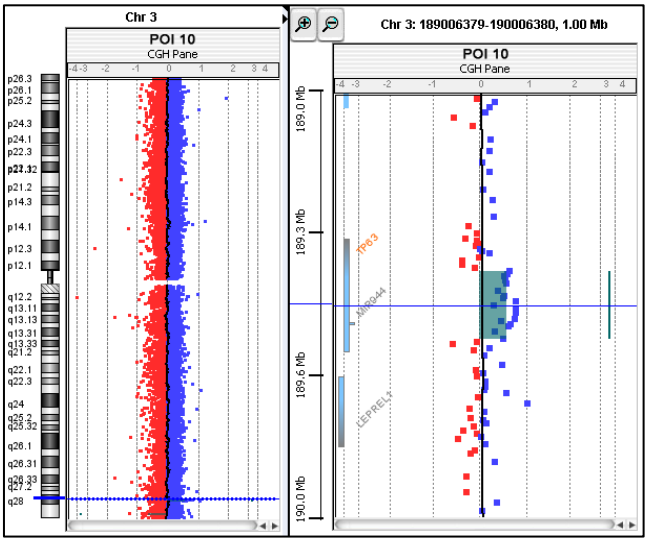

Fig.30. *array-CGH result showing TP63 intragenic duplication in POI 10.*

The extension of the analysis on POI 10 cDNA by RT-qPCR and direct sequencing between exons junction 10-2, allowed to confirm that the aberrant gene results in a stable aberrant mRNA as well. RT-qPCR using two different TaqMan probes (one outside and the other inside the duplicated region), also revealed that the aberrant mRNA consists about 30% of the total transcript. Regarding the reading frame, no changes were identified, thus, if NMD (non-sense mediated decay) process does not occur, probably an aberrant protein with the main functional domain completely duplicated (p-53 DNA binding domain) is translated (**Fig.31B**).

TP63 could be considered a new candidate POI gene: it encodes a protein of tumor suppressor p53 family and it is a main regulator to protect the genomic integrity of female germ cells during

meiotic arrest. Moreover, null mice are infertile, because their oocytes do not undergo apoptosis after DNA damage [181].

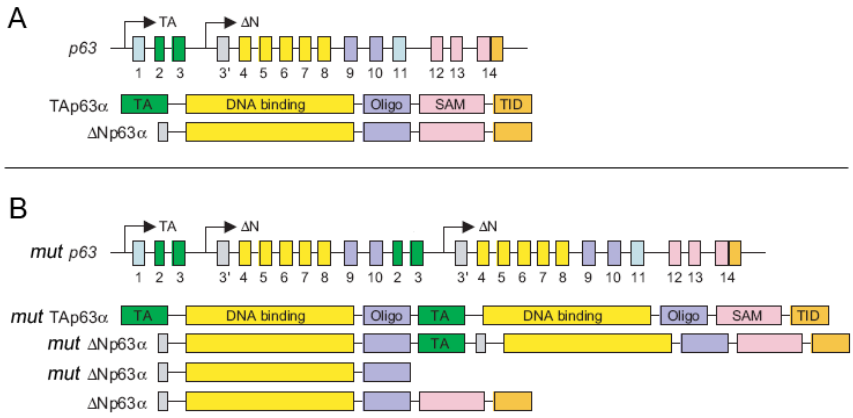

Fig.31. A) *TP63 mRNA and protein are shown. Main isoform has in the promoter region a transactivating domain (TA), whereas the shortest isoform has a ΔN promoter region, lacking the TA domain. Both isoforms have: DNA binding domain, oligomerization domain (Oligo), a sterile alpha motif (SAM) and a transactivation inhibitory domain (TID).* ***B)*** *Mutated TP63 mRNA and hypothetical proteins are shown. The image was modified from [182].*

Indeed, after DNA damage, Tp63 is activated by kinases phosforylation and dimers opening lead to tetramerization (**Fig.32**). Thus, Tp63 triggers apoptosis, by activation of pro-apoptotic factors Puma and Noxa, whose function is to inhibit the survival family Bcl [237,240].

It is worthwhile to infer that whether the aberrant *TP63* transcript in POI 8 and 10 results in no translation, the adequate protein level mandatory for oocyte maturation might be missed from the earliest stages of fetal development, thus resulting in a deleterious effect for the correct oogenesis. Contrary, the presence of an aberrant protein, might have altered the proper protein folding, preventing the tetramer formation and resulting in few functional wild-type proteins able to protect oocyte genome [180].

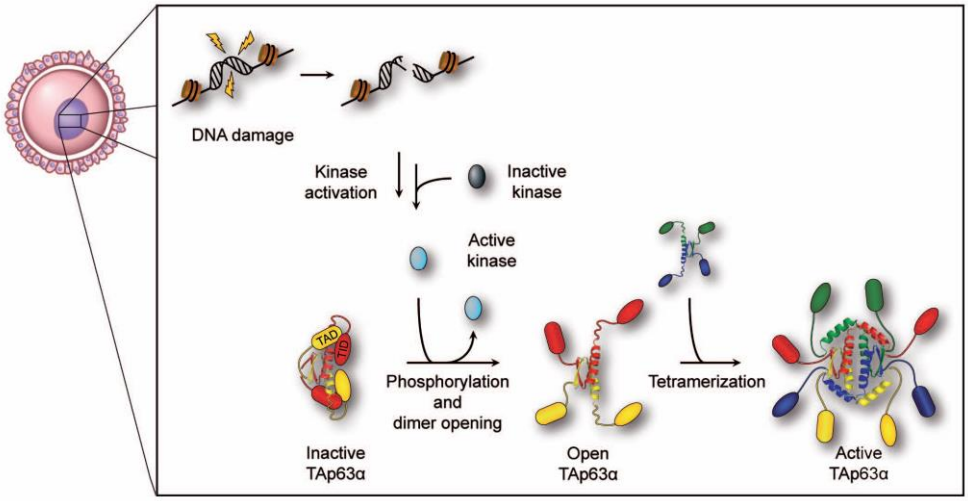

Fig.32*. Tp63 activation (TAp63α) after DNA damage during oocyte development [180,237].*

Notably the patients' father, having the same CNV, died in young age from cancer. Since Tp63 is a tumor suppressor, the deleterious effect of the CNV is further supported. Thus, it could be strongly suggested that the alteration in Tp63 in POI 8 and 10 leads to their fertility problems: nevertheless it is appropriate to address the patients also to a proper cancer screening surveillance.

In sisters **POI 9 and 11**, an intronic deletion in *PRKAA1* inherited from the father was detected (**Fig.33**). CNV validation and characterization was crucial also in this case because only two array spots identified the loss and the adjacent exons 2, 3, and 4 were not covered by the array probes. qPCR on gDNA excluded the perturbation of *PRKAA1* coding regions confirming an intronic deletion and the subsequent RT-qPCR did not allow to evaluate a transcript alteration due to the extreme variability of mRNA levels in the blood tissue analyzed. Nevertheless, it cannot be excluded that mRNA alterations occurred in other tissues during ovarian development. Indeed, murine, porcine and hen animal models showed the involvement of *PRKAA1* in oocyte meiotic resumption [184,185,186,188,241]. Particularly, the process is regulated by the kinase in a dose-dependent manner.
The CNV detected is reported at a very low frequency (5/2562) in DGV: thus, its causative role has yet to be elucidated.

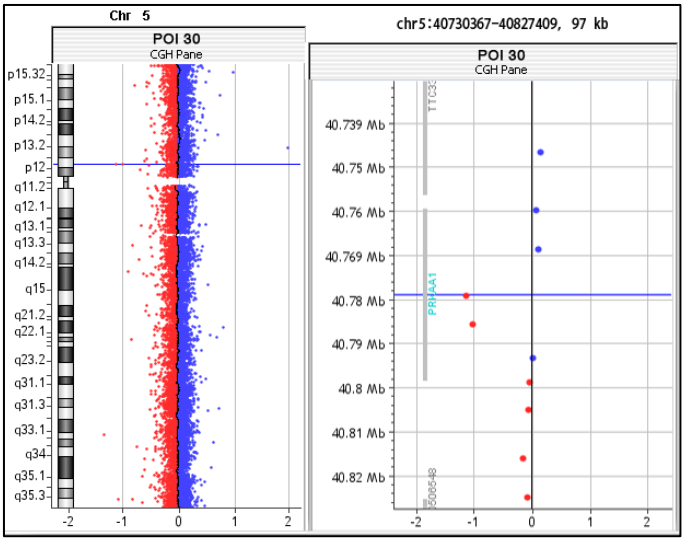

Fig.33. *array-CGH result showing the two spots deletion involving PRKAA1 (POI 30 is depicted)*

Another example of the importance of array CGH data validation and characterization, is represented by a 4 spots deletion involving *VLDLR* 3'UTR (exon 19) and the proximal 14 kb identified in **POI 44 (Fig.34)**. Array-CGH analysis was not performed on patient's parents because they were not available.

As in the previous characterization, exons 16, 17, 18 were not covered by array probes. Long-Range PCR and Sanger sequencing allowed not only to exclude the involvement of the three exons, but also to evaluate that the array probe covering *VLDLR* 3'UTR was a fake. In addition, sequencing of the deletion breakpoints revealed how the formation of genomic rearrangements maay be complex. In detail, the deletion downstream *VLDLR* includes the proximal 3 kb, and after 140 bp conserved, the following 9.4 kb which correspond to the 3 deleted array spots (**Fig.24A**).

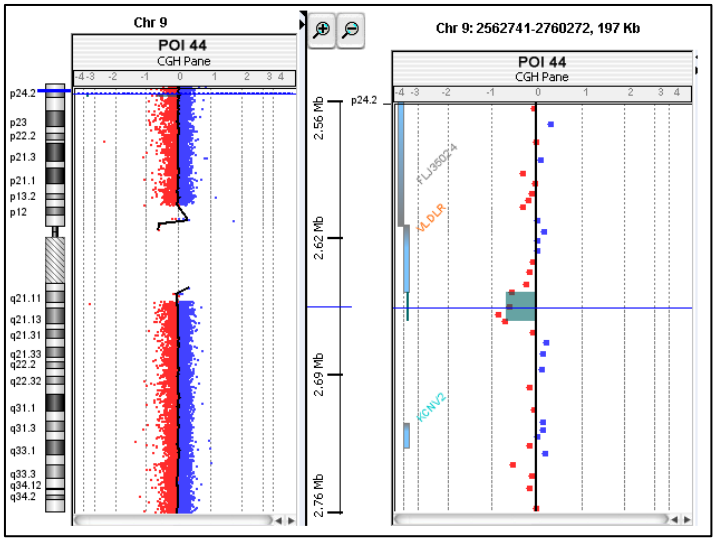

Fig.34. *array-CGH result showing the VLDLR partial deletion in POI 44.*

Thus, the gene is not included in the CNV but might be perturbed by a position effect mediated by the downstream deletion. In support to this hypothesis, the analysis performed at mRNA level showed a drastic transcript downregulation and the heterozygosity assay confirmed a monoallelic expression of *VLDLR* in the patient. Furthermore, functional studies carried out using luciferase assay in a granulosa cell line, highlighted an enhancer role of the region including the ENCODE predicted insulator and the 3000 bp downstream deleted in the patient (**Fig.35**). These findings support that the CNV deregulate *VLDLR* gene by position effect in the blood tissue analyzed and that the deleted region has a regulatory role in oocyte. Particularly, luciferase assay using pGL3b vector with *VLDLR* promoter, showed that the deleted region plays a crucial role in *VLDLR* regulation (**Fig.26**).

Remarkably, the prediction performed in ENCODE project considers 9 different cell lines (GM12878: B-lymphocyte, lymphoblastoid cell line; H1-hESC: embryonic stem cells; HepG2: hepatocellular carcinoma; HMEC: mammary epithelial cells; HSMM: skeletal muscle myoblasts; HUVEC: umbilical vein endothelial cells; K562: leukemia; NHEK: epidermal keratinocytes;

NHLF: lung fibroblasts[242]), among which no ovarian cell line is present. Thus, it is likely that the ENCODE prediction of "heterochromatin" region (**Fig.35**, red box) does not occur in ovary.

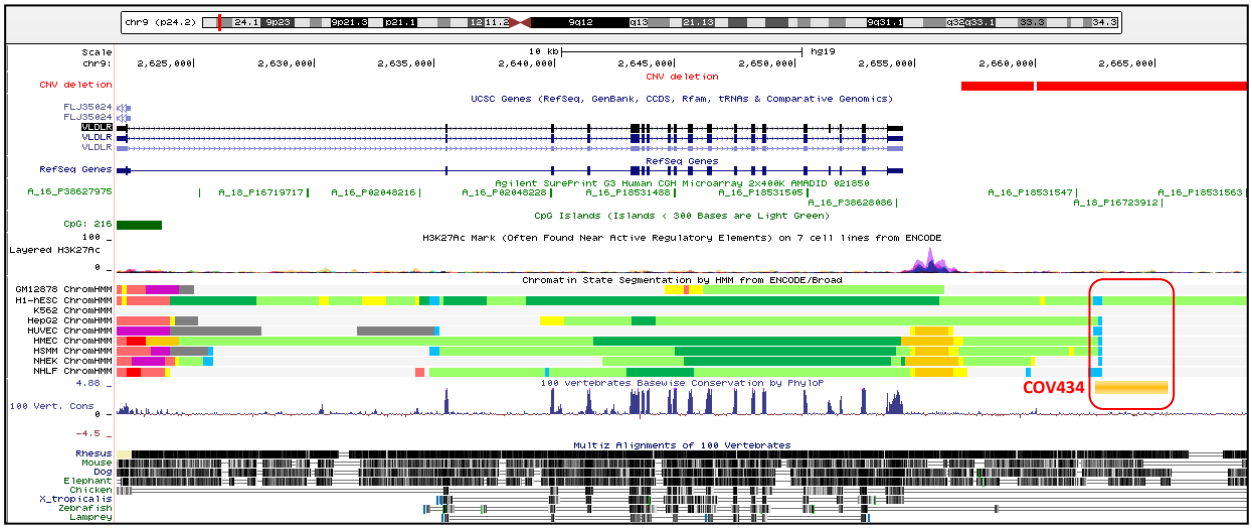

Fig.35. UCSC genome browser view (hg19). *VLDLR gene is shown in blue; downstream the deleted regions are shown in red. The region with an enhancer role in COV434 is shown in orange. Regulatory elements prediction based on ENCODE chromatin state segmentation in nine human cell lines is shown: red, promoter; pourple, poised promoter; yellow, weak enhancer; orange, strong enhancer; dark green, transcriptional elongation; light green, weak transcribed; light blue, insulator; dark grey, polycomb-repressed; light grey, heterochromatin. Conservation in other species is shown.*

Mutations in the gene have never been reported in human POI, but homozygous mutations are associated to Cerebellar hypoplasia and mental retardation with or without quadrupedal locomotion 1 (OMIM #224050). However, *VLDLR* in human is very important for steroidogenesis; in particular, the protein regulates lipoprotein endocytosis in order to ensure the correct follicular growth and its mRNA is mainly found in granulosa cells of preovulatory follicles [243,244]. This finding is supported by several publications in which the role of lipoprotein receptor is discussed in animal models:

- hens lacking a functional receptor due to a point mutation are sterile, the oocytes fail to growth, become necrotic and do not ovulate. Moreover, mutant chickens present also severe hyperlipidemia and show atherosclerotic lesions [203,204];
- in bovine, Argov *et al.* [205,236] reported for the first time the expression of *VLDLR* throughout different stages of follicular development. In particular *VLDLR* was expressed in the granulosa cells in early and preovulatory follicle, and in theca cells in preovulatory follicle;
- it is also associated with egg performance in duck, in which two *VLDLR* transcripts were identified: *VLDLR*-a mainly expressed in muscle tissue, and *VLDLR*-b in reproductive

organs, with a huge amount in theca and granulosa cells from early stages of follicle development [202,245].

Based on gene function, it might be hypothesized that in POI 44, *VLDLR* haploinsufficiency is responsible for a modified cholesterol endocytosis in oocytes leading to an improper estrogen production, and resulting in hypogonadism and then POI.

Concerning the functional pathways in which it is involved, Vldlr protein associates with another receptor, namely Apoer2 (cholesterol transport protein apolipoprotein E). Besides cholesterol endocytosis, the receptors are involved in the signaling pathway mediated by Reelin (*RELN*), which is important for neuronal migration during brain development, and learning and memory during adult life [246,247]. Particularly, Reelin interacts with Vldlr and Apoer2 extracellular component, causing their clusterization and the activation of the signaling pathway. Dab1, one of the intracellular interactors, is then phosphorylated and activates a pathway involved in cellular motility, adhesion and shape (**Fig.36**) [248,249,250,251]. Besides its role in central nervous system (CNS), Reelin resulted also expressed in mouse ovary [252] and in a publication by Fayad *et al.* its mRNA was found in bovine oocytes, mainly in theca cells of dominant follicles [253].

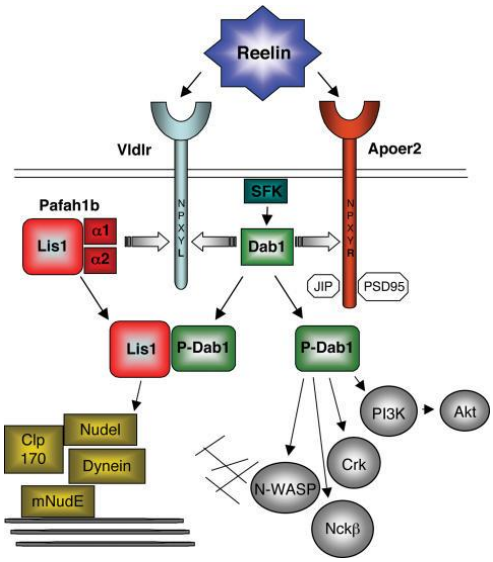

Fig.36. Reelin signaling pathway. *Reelin binds Vldlr and Apoer2 receptors triggering src-family kinase (SFK) activation and Dab1 phosphorylation. Dab1 binds to the NPxY motif of both receptors. Upon Reelin stimulation, phosphoDab1 (P-Dab1) interacts with Lis1 and with other signal transduction molecules (grey circles). Signaling molecules downstream of Lis1 and Dab1 affect cytoskeleton dynamics by acting on microtubules (thick lines) or actin filaments (thin lines), thereby controlling cellular migration and formation[246].*

This finding sheds light on the brain-ovarian functions of several genes. An example is the neurotrophin family, which includes: nerve growth factor (*NGF*), neurotrophin 3 (*NT3*),

neurotrophin 4/5 (*NT4/5*), brain-derived neurotrophic factor (*BDNF*), glial derived neurotrophic factor (*GDNF*) and vasointestinal peptide (*VIP*). Besides their predominant role in brain, these factors act also during ovarian physiological and pathological processes, as recently reported by Streiter *et al.* [254]. They promote follicular cell proliferation, ovulation, inhibit follicular apoptosis, promote oocyte maturation, and stimulate steroidogenesis.
Consistently, among the ovarian genes identified by array-CGH (**Tab.13**), three have also a brain function: *PARD3* (POI 45) is required for establishment of neuronal polarity and normal axon formation in cultured hippocampal neurons but it is also involved in establishing asymmetric division in mouse and xenopus oocytes [206,207]; *NRP1* (POI 45) plays a role in the formation of certain neuronal circuits but it is also expressed both in granulosa and theca cells in bovine ovary [208]; *GRIA1* (POI 48) encodes a glutamate receptor which is the predominant excitatory neurotransmitter receptors in the mammalian brain and is activated in a variety of normal neurophysiological processes, but it is also suggested as a critical mediator of ovulation in cow [212,213].

In **POI 46**, a homozygous intronic deletion at 15q25.2 in *CPEB1* (**Fig.37**) was characterized in order to evaluate the involvement of gene promoter region in the CNV. The analyses performed on gDNA confirmed a homozygous deletion in intron 1 of about 10.8 kb. Unfortunately, gDNA from patient's parents, as well as mRNA from POI 46 were not available for further analyses.
The involved region is predicted to be transcriptionally repressed, and the promoter region of the isoform included in the CNV results a poised promoter in the cellular lines investigated by the ENCODE project [242]. Based on the functional studies performed on *VLDRL* one may speculate that it is an ovarian specific isoform and that the deletion might involve some regulatory elements (not predicted by ENCODE project) with a pivotal function in *CPEB1* regulation. Thus, its haploinsufficiency might disrupt the proper ooctyes maturation and development.
Furthermore, the gene was found fully deleted in **POI 54**, who has a CNV of about 1.6 Mb which involves other three ovarian genes (**Tab.11**). To support its role in POI onset, one very similar variation was also reported by McGuire *et al.* [161] in a patient with PA.
CPEB1 is preferentially expressed in oocyte and is involved in synaptonemal complex formation during oocyte maturation as proved by several animal models (mouse, xenopus, porcine [209,210,211]). Female null mice lack follicles due to oocytes premature loss, and female embryos

show oocytes stalled at pachitene [209]. *CPEB1* is also important in regulating cellular cycle progression during mitotic embryonic division [255].

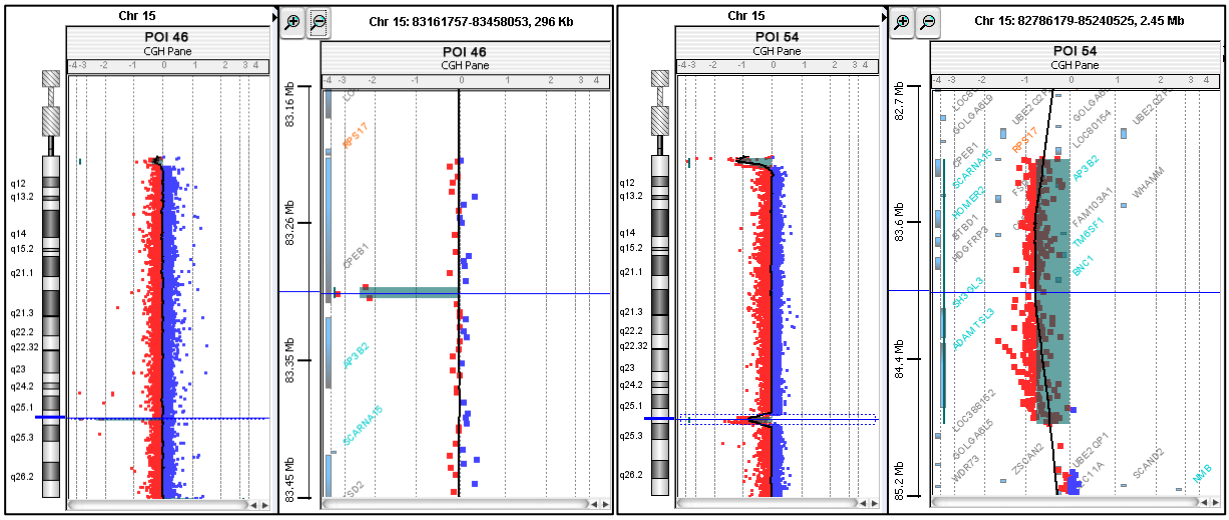

Fig.37. *array-CGH results showing the CNVs involving CPEB1 in POI 46 and 54, respectively.*

Particularly, during zygotene to pachytene transition in meiosis, Cpeb1 is phosphorylated by Aurora A kinase (**Fig.38**) and activated with the subsequent transcription of proteins involved in synaptonemal complex formation (Scp1, Scp3). In mouse KO, these proteins are not translated and the complex is not formed with the following stall of follicular development. When wild-type Cpeb1 progresses to diplotene, is dephosphorylated and thus inactivated. During ovulation, oocyte resumes meiosis and Cpeb1 is activated again promoting Mos and CyclinB1 transcription (**Fig.38**) [256].

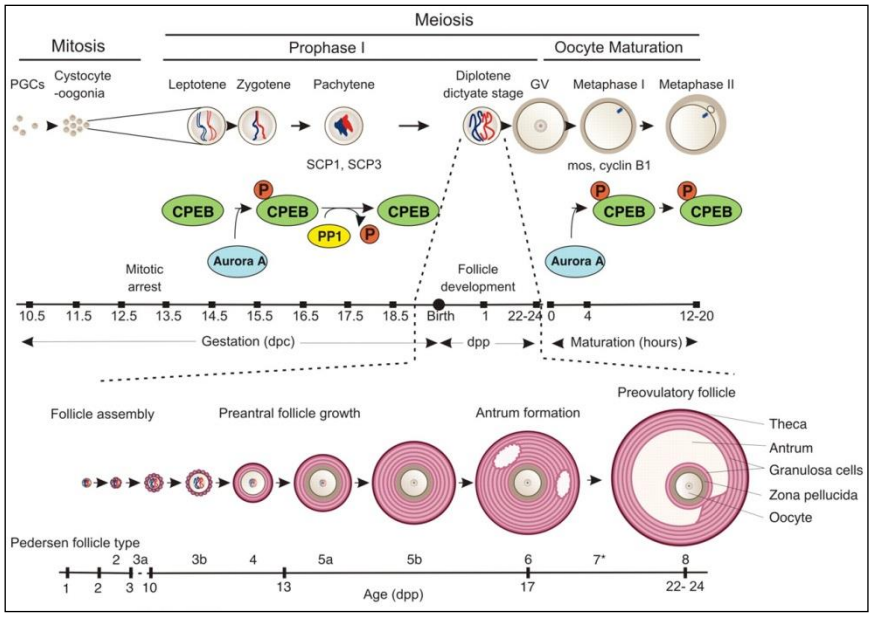

Fig.38. *Cpeb1 activity during oogenesis and oocyte maturation in mice [256].*

Therefore its role in POI etiology is clear: the haploinsufficiency in human might accelerate the loss of primordial germ cells and thus it may contribute to the extreme phenotype onset shown by POI 46 and 54.

As aforementioned, the frequency of "common" CNVs does not change between patients and controls. Among these CNVs, that with the most relevant evidence about its putative role in POI onset is the duplication/deletion involving *SYCE1* gene. The duplication which appears most frequently according to DGV, was found in 4 patients and detected nine times in controls. By contrast, the only deletion found in POI 60 was never reported in the *ad hoc* control group, but it was previously detected in a patient with premature SA (21 years) discussed by McGuire *et al.* [161] (**Fig.39**).

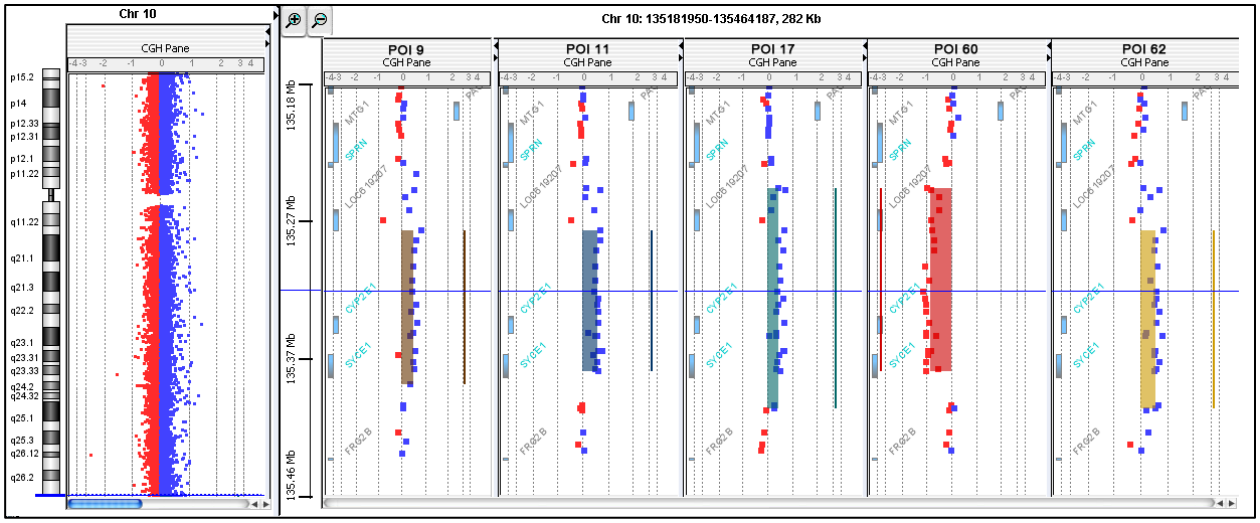

Fig.39. *array-CGH results showing the CNVs involving SYCE1 gene found in POI cohort.*

Syce1 is part of the synaptonemal complex, plays a role in ensuring complex stability together with other proteins (Cesc1, Sycp1) and in recombination during meiosis (**Fig.40**). In Syce1 nullmouse, DNA damage repair pathway is changed and after DSB, oocytes undergo apoptosis leading to infertility [147]. In humans, haploinsufficiency might result in a similar effect, thus in POI 60, the *SYCE1* deletion could have a major contribution in POI etiology, leading to premature oocytes loss. This hypothesis is supported by a recent publication in which *SYCE1* haploinsufficiency by a homozygous mutation was associated to POI [146]. Hence, it will be appropriate to exclude any loss-of-function alterations in the wild-type allele of POI 60.

Regarding duplications, it was not possible to understand both in patients and in controls whether *SYCE1* was completely involved in the CNVs due to the presence of segmental duplication near the promoter region. Even if no differences about the frequency were shown, it seemed crucial to assess whether the size of the rearrangement was the same between patients

and controls thus leading to a different effect. qPCR revealed no differences and *SYCE1* was fully duplicated both in POI and controls. However, one cannot ruled out that these CNVs, in combination with others, might have a susceptibility role in POI onset.

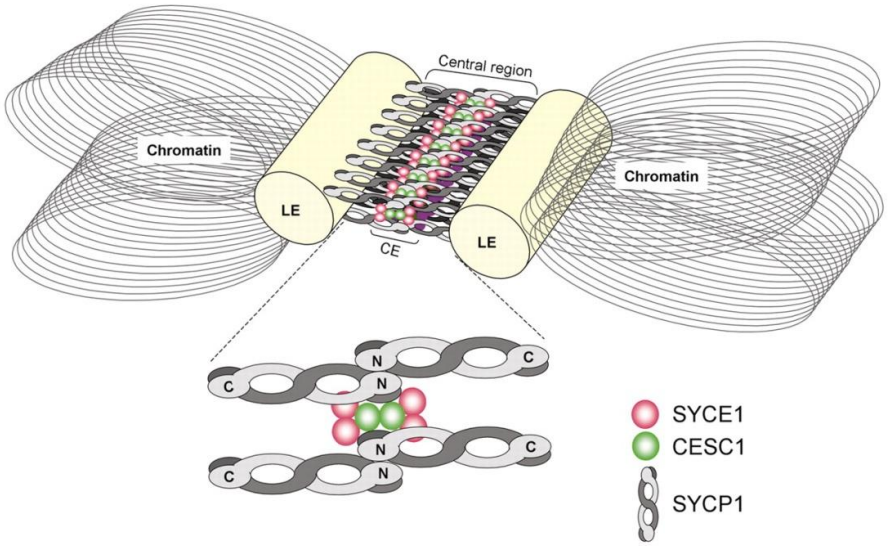

***Fig.40. Schematic representation of the synaptonemal complex**. Dimers or multimers of SYCE1 and CESC1 associate with the N-terminal region of SYCP1 forming the central region. C, carboxyl-terminus; CE, central element; LE, lateral element; N, amino-terminus [144].*

4.3.1 Bioinformatic analysis

To detect a potential link between array candidate genes and the ones known to be involved in POI onset, bioinformatic analysis was performed. Since few genes were detected (37) it was not possible to predict a statistically significance enrichment in term of Gene Ontology (GO), biological processes, as well as KEGG pathways. For this reason STRING software was used to verify any protein connections. Some interesting data were obtained, thus supporting the possible causative role of the following genes in POI etiology (**Fig.41**):

- Sts, the steroidogenic factor, as well as Amfr and Cyp2e1, were annotated in the same pathway "Steroid Hormone biosynthesis" of Cyp17a1 and Cyp19a1 [KEGG];
- Smyd3, (regulator of transcription factors) shows a moderate binding score (0.576) with Esr1, (estrogen receptor) [257]. Particularly, it is required for transcription of genes regulated by Esr1 in estrogen signaling pathway;
- Prkaa1 is involved in post-translation modification of FOXO family members in the "FoxO signaling pathway" [KEGG];

- Sycel shows a high binding score (0.900) with Stag3, in synaptonemal complex formation [Reactome].

Due to the identification of several ovarian genes which could be selected as new candidates for POI onset, array CGH is therefore a very useful approach to be adopted, both in familial and idiopathic cases.

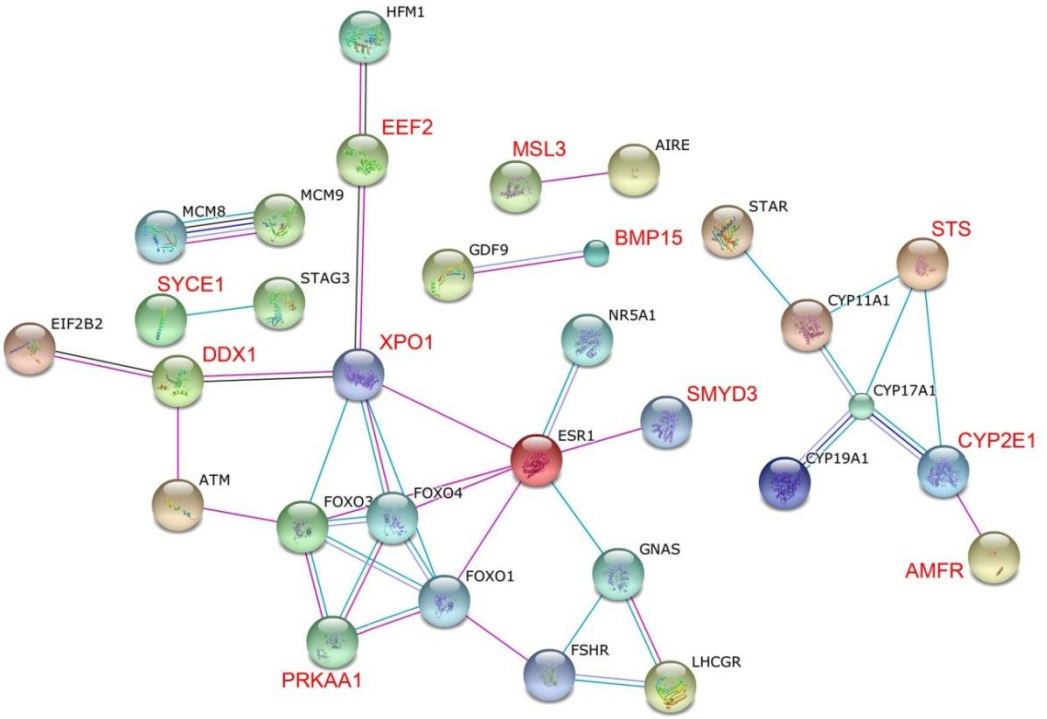

Fig.41. Output representation of STRING analysis. *The proteins encoded by array-CGH detected genes are shown in red. The proteins encoded by already reported POI genes are shown in black. Purple line: experimental/biochemical data; light purple line: homology; blue line: co-occurence across genome; light blue line: association in curated databases; black line: coexpression.*

4.4 Preliminary WES analysis

Based on the present literature, few WES analyses have been carried out, mainly on POI non-syndromic familial cases with women affected by both PA and SA [116,135,136,146,167]. Only in the study by Fonseca *et al.* [166], twelve sporadic POI patients affected by SA were analyzed.
Here, it is reported the first WES investigation on non-syndromic unrelated patients affected by the extreme POI phenotype (PA).
Particularly, WES analysis was performed in 17 out of 67 patients, thanks to a collaboration with Genetics of Common Disorders Unit at San Raffaele Hospital. The preliminary approach in the identification of rare SNVs consisted in the screening of a total of 191 selected genes derived from array-CGH data, and literature regarding POI and ovary (http://omim.org/,[155]). In addition, the interactors of the three array genes *VLDLR*, *TP63*, *CPEB1* (http://string-db.org/), were considered in the research, based on their function, their high association to POI and the molecular evidences obtained in this thesis work.
The variants obtained were filtered according to the steps shown in **Fig.6 and Fig.29**. In detail, considering POI a heterogeneous complex disorder and based on the purpose of this study, the most rare SNVs with MAF < 0.5% were considered. Then based on the evidence that ExAC browser collects about 60,700 exome sequencing data from several pathological individuals, variants with a higher frequency (> 5%) were excluded because considered benign according to a recent publication [258].
The variant annotation software used (wANNOVAR) collects all the information from several databases but does not consider the HGMD mutation database. Moreover for some submitted SNVs, dbSNP ID is not reported. Thus, the variants obtained were analyzed individually consulting several bioinformatic databases, to clarify their benign or pathological effect. HGMD, dbSNP, ensembl, ClinVar as well as aminoacid substitution prediction softwares (SIFT [170], Polyphen [171], MutationTaster [172]) were interrogated and a total of 50 "uncertain significance", "likely pathological", and "pathological" variants [258] were obtained. 10 out of 50 SNVs were signed as low quality data, indeed they were not confirmed by Sanger validation.
Among the array genes, 7 "likely pathological" variants were detected and validation for 3 of them was obtained supporting their putative role in the disorder. They are discussed below.

- In POI 50, positive to array-CGH analysis (*STS* fully duplicated), a *WHAMM* missense variant c.C734T was detected. It resulted never reported in the databases investigated and predicted "damaging" for the three main softwares consulted. The protein encoded acts as a nucleation-promoting factor (NPF) that stimulates actin polymerization both

at the Golgi apparatus and along tubular membranes. Only one publication highlights its role in ovary: Huang *et al.* [217], examined its localization and function in mouse oocytes during meiosis. Particularly, *WHAMM* silencing by interfering siRNA lead to failure of spindle migration and alterations in the correct asymmetric division during oocyte maturation.

- In POI 57, negative to array CGH analysis, the yet unreported *RYR3* missense variant c.G6604A was found. The gene encodes a receptor which contributes to cellular calcium ion homeostasis and plays a role in cellular calcium signaling. It is expressed on the oocyte surface of mature bovine and anphibia [173,174]. The SNV with a predicted "damaging" effect could have changed the ovarian calcium levels which are essential for the activation and the development of oocytes.
- In POI 67, negative to array CGH analysis, a *VLDLR* missense variant c.G902A was detected. It is reported as rs139671268, with a very low frequency (MAF: 0.001).
  After Sanger sequencing of *VLDLR* exon 6 in the rest of POI cohort, the SNV was also found in POI 46, (carrier of the homozygous intronic deletion in *CPEB1*). rs139671268 allele frequency in the present cohort resulted higher (0.015) compared to the control population (0.001) **(Tab.19)** thus supporting an association with POI onset. In addition, the mother of POI 67 who underwent premature menopause, presents the same SNV.
  It is further interesting to note that variants involving genes with a role in the same *VLDLR* pathways were discovered: i.e. *RELN*, c.C5621A in POI54 (carrier of two rare CNVs involving *SLMAP* and *CPEB1*, *WHAMM* with other genes, respectively); *APOE*, rs199768005 in POI 65 (negative to array-CGH).

Notably, for *RYR3* and *VLDLR* missense variants, "damaging" prediction was obtained in Polyphen and MutationTaster, but it was "tolerated" in SIFT. The discordance between these predictions is quite common and depends on type of source and algorithm used by the different tools. Indeed, SIFT is considered a prediction software based on "Evolutionary conservation", whereas Polyphen and Mutation Taster are also based on " protein structure and function" [258]. In the present thesis work, Polyphen was mainly considered in the interpretation of data because its prediction is correlated to human phenotypes [171].

Noteworthy the present WES preliminary analysis allowed to confirm the importance of some array CGH candidate genes in POI onset: e.g.*VLDLR* gene, for which more molecular and literature evidences are present. Moreover, the exome sequencing findings support the heterogeneity of POI: every patient has at least one rare SNV in the ovarian genes selected and

some of POI women are also carriers of rare "ovarian" CNVs. The combination of these molecular evidences, with major or minor contribution, might define the underlying mechanism of their POI phenotype.

4.5 Conclusions and future perspectives

The data obtained in the present work, confirm the extreme heterogeneity of Primary Ovarian Insufficiency, in which the combination of more rare/common variants seems to concur in the phenotypic onset.

The analysis of POI patients affected by the most severe phenotype enabled to detect new candidate ovarian genes never reported before. Indeed, the combined genome wide approach of array-CGH and WES techniques, resulted an efficient tool to identify rare variants (CNVs and SNVs) involving both genes already reported in POI, and new candidate genes with a role in oocyte maturation and differentiation.

Hence, the present tool is strongly recommended to be adopted in the continuous growing understanding of the genetic basis of POI. The characterization and the discovery of POI molecular causes in each patient, will help not only to shed light on the pathways underlying women reproduction but also to ensure a proper diagnosis and then an appropriate therapeutic intervention. Moreover, the new candidate genes could be specifically analyzed in the future for both diagnostic and prognostic purposes.

Finally, it will be appropriate to:

- deepen the most interesting results emerged from array-CGH analysis in order to support their putative pathogenic effect (e.g. Tp63 Western Blot to detected the presence of an aberrant protein; functional studies on the main candidate genes);
- complete WES analysis on the 17 patients and process the other 50 POI;
- validate WES data by Sanger sequencing;
- confirm the actual rarity of the identified variants screening the *ad hoc* control group with the same WES approach.

REFERENCES

1. McGee EA, Hsueh AJ (2000) Initial and cyclic recruitment of ovarian follicles. Endocr Rev 21: 200-214.

2. Pepling ME (2006) From primordial germ cell to primordial follicle: mammalian female germ cell development. Genesis 44: 622-632.

3. Saladin KS (2012) Anatomia Umana.

4. Stanfield CL, Germann WJ (2009) Fisiologia.

5. Lobo RA (2003) Early ovarian ageing: a hypothesis. What is early ovarian ageing? Hum Reprod 18: 1762-1764.

6. Baker TG, Franchi LL (1967) The fine structure of oogonia and oocytes in human ovaries. J Cell Sci 2: 213-224.

7. Sathananthan AH, Selvaraj K, Trounson A (2000) Fine structure of human oogonia in the foetal ovary. Mol Cell Endocrinol 161: 3-8.

8. McLaughlin EA, McIver SC (2009) Awakening the oocyte: controlling primordial follicle development. Reproduction 2009 Jan;137(1):1-11. doi: 10.1530/REP-1508-0118.

9. Matsuda-Minehata F, Inoue N, Goto Y, Manabe N (2006) The regulation of ovarian granulosa cell death by pro- and anti-apoptotic molecules. J Reprod Dev: 2006 Dec;2052(2006):2695-2705.

10. Persani L, Rossetti R, Cacciatore C (2010) Genes involved in human premature ovarian failure. J Mol Endocrinol 2010 Nov;45(5):257-79. doi: 10.1677/JME-1610-0070.

11. Gougeon A (1996) Regulation of ovarian follicular development in primates: facts and hypotheses. Endocr Rev 17: 121-155.

12. Hirshfield AN (1991) Development of follicles in the mammalian ovary. Int Rev Cytol 124: 43-101.

13. Richards JS, Fitzpatrick SL, Clemens JW, Morris JK, Alliston T, et al. (1995) Ovarian cell differentiation: a cascade of multiple hormones, cellular signals, and regulated genes. Recent Prog Horm Res 50: 223-254.

14. Reynaud K, Driancourt MA (2000) Oocyte attrition. Mol Cell Endocrinol 163: 101-108.

15. Choi Y, Rajkovic A (2006) Genetics of early mammalian folliculogenesis. Cell Mol Life Sci 63: 579-590.

16. Block E (1953) A quantitative morphological investigation of the follicular system in newborn female infants. Acta Anat 17: 201-206.

17. Forabosco A, Sforza C, De Pol A, Vizzotto L, Marzona L, et al. (1991) Morphometric study of the human neonatal ovary. Anat Rec 231: 201-208.

18. Tilly JL, Kowalski KI, Johnson AL, Hsueh AJ (1991) Involvement of apoptosis in ovarian follicular atresia and postovulatory regression. Endocrinology 129: 2799-2801.

19. Kaipia A, Hsueh AJ (1997) Regulation of ovarian follicle atresia. Annu Rev Physiol 59: 349-363.

20. Erickson GF, Shimasaki S (2000) The role of the oocyte in folliculogenesis. Trends Endocrinol Metab 11: 193-198.

21. Erickson GF, Danforth DR (1995) Ovarian control of follicle development. Am J Obstet Gynecol 172: 736-747.

22. Araujo VR, Gastal MO, Figueiredo JR, Gastal EL (2014) In vitro culture of bovine preantral follicles: a review. Reprod Biol Endocrinol: 2014 Aug 2013;2012:2078.

23. Guigon CJ, Magre S (2006) Contribution of germ cells to the differentiation and maturation of the ovary: insights from models of germ cell depletion. Biol Reprod: 2006 Mar;2074(2003):2450-2008.

24. Coucouvanis EC, Sherwood SW, Carswell-Crumpton C, Spack EG, Jones PP (1993) Evidence that the mechanism of prenatal germ cell death in the mouse is apoptosis. Exp Cell Res 209: 238-247.

25. Pesce M, De Felici M (1994) Apoptosis in mouse primordial germ cells: a study by transmission and scanning electron microscope. Anat Embryol 189: 435-440.

26. De Pol A, Vaccina F, Forabosco A, Cavazzuti E, Marzona L (1997) Apoptosis of germ cells during human prenatal oogenesis. Hum Reprod 12: 2235-2241.

27. Picton H, Briggs D, Gosden R (1998) The molecular basis of oocyte growth and development. Mol Cell Endocrinol 145: 27-37.

28. Wallace WH, Kelsey TW (2010) Human ovarian reserve from conception to the menopause. PLoS One: 2010 Jan 2027;2015(2011):e8772.

29. Skinner MK (2005) Regulation of primordial follicle assembly and development. Hum Reprod Update: 2005 Sep-Oct;2011(2005):2461-2071.

30. Hutt KJ, McLaughlin EA, Holland MK (2006) KIT/KIT ligand in mammalian oogenesis and folliculogenesis: roles in rabbit and murine ovarian follicle activation and oocyte growth. Biol Reprod: 2006 Sep;2075(2003):2421-2033.

31. Bachvarova R (1985) Gene expression during oogenesis and oocyte development in mammals. Dev Biol 1: 453-524.

32. Wassarman PM, Liu C, Litscher ES (1996) Constructing the mammalian egg zona pellucida: some new pieces of an old puzzle. J Cell Sci 109: 2001-2004.

33. Simon AM, Goodenough DA, Li E, Paul DL (1997) Female infertility in mice lacking connexin 37. Nature 385: 525-529.

34. Beyer EC (1993) Gap junctions. Int Rev Cytol: 1-37.

35. Kumar NM, Gilula NB (1996) The gap junction communication channel. Cell 84: 381-388.

36. Fortune JE, Cushman RA, Wahl CM, Kito S (2000) The primordial to primary follicle transition. Mol Cell Endocrinol 163: 53-60.

37. Erickson GF, Magoffin DA, Dyer CA, Hofeditz C (1985) The ovarian androgen producing cells: a review of structure/function relationships. Endocr Rev 6: 371-399.

38. Yamoto M, Shima K, Nakano R (1992) Gonadotropin receptors in human ovarian follicles and corpora lutea throughout the menstrual cycle. Horm Res 1: 5-11.

39. Li R, Phillips DM, Mather JP (1995) Activin promotes ovarian follicle development in vitro. Endocrinology 136: 849-856.

40. Yoshida H, Takakura N, Kataoka H, Kunisada T, Okamura H, et al. (1997) Stepwise requirement of c-kit tyrosine kinase in mouse ovarian follicle development. Dev Biol 184: 122-137.

41. Trounson A, Anderiesz C, Jones GM, Kausche A, Lolatgis N, et al. (1998) Oocyte maturation. Hum Reprod 3: 52-62.

42. Zeleznik AJ (2004) The physiology of follicle selection. Reprod Biol Endocrinol 2: 31.

43. Hutt KJ, Albertini DF (2007) An oocentric view of folliculogenesis and embryogenesis. Reprod Biomed Online 14: 758-764.

44. Goswami D, Conway GS (2007) Premature ovarian failure. Horm Res: 2007;2068(2004):2196-2202.

45. Morabia A, Costanza MC (1998) International variability in ages at menarche, first livebirth, and menopause. World Health Organization Collaborative Study of Neoplasia and Steroid Contraceptives. Am J Epidemiol 148: 1195-1205.

46. Beck-Peccoz P, Persani L (2006) Premature ovarian failure. Orphanet J Rare Dis 1: 9.

47. Cooper AR, Baker VL, Sterling EW, Ryan ME, Woodruff TK, et al. (2011) The time is now for a new approach to primary ovarian insufficiency. Fertil Steril 2011 May;95(6):1890-7. doi: 10.1016/j.fertnstert.2010.1001.1016.

48. Albright F, Smith PH, Fraser R (1942) A syndrome characterized by primary ovarian insufficiency and decerased stature. Am J Med Sci 204: 625-648.

49. Cox L, Liu JH (2014) Primary ovarian insufficiency: an update. Int J Womens Health 2014 Feb 20;6:235-43. doi: 10.2147/IJWH.S37636.

50. Goswami D, Conway GS (2005) Premature ovarian failure. Hum Reprod Update: 2005 Jul-Aug;2011(2004):2391-2410.

51. Nelson LM (2009) Clinical practice. Primary ovarian insufficiency. N Engl J Med: 2009 Feb 2005;2360(2006):2606-2014.

52. Welt CK (2008) Primary ovarian insufficiency: a more accurate term for premature ovarian failure. Clin Endocrinol (Oxf): 2008 Apr;2068(2004):2499-2509.

53. Fortuno C, Labarta E (2014) Genetics of primary ovarian insufficiency: a review. J Assist Reprod Genet 2014 Dec;31(12):1573-85. doi: 10.1007/s10815-10014-10342-10819.

54. Kovanci E, Schutt AK (2015) Premature ovarian failure: clinical presentation and treatment. Obstet Gynecol Clin North Am: 2015 Mar;2042(2011):2153-2061.

55. Cordts EB, Christofolini DM, Dos Santos AA, Bianco B, Barbosa CP (2011) Genetic aspects of premature ovarian failure: a literature review. Arch Gynecol Obstet 2011 Mar;283(3):635-43. doi: 10.1007/s00404-00010-01815-00404.

56. Jin M, Yu Y, Huang H (2012) An update on primary ovarian insufficiency. Sci China Life Sci 2012 Aug;55(8):677-86. doi: 10.1007/s11427-11012-14355-11422.

57. Visser JA, Schipper I, Laven JS, Themmen AP (2012) Anti-Mullerian hormone: an ovarian reserve marker in primary ovarian insufficiency. Nat Rev Endocrinol: 2012 Jan 2010;2018(2016):2331-2041.

58. Younis JS (2012) Ovarian aging and implications for fertility female health. Minerva Endocrinol 37: 41-57.

59. Persani L, Rossetti R, Cacciatore C, Bonomi M (2009) Primary Ovarian Insufficiency: X chromosome defects and autoimmunity. J Autoimmun 2009 Aug;33(1):35-41. doi: 10.1016/j.jaut.2009.1003.1004.

60. van Kasteren YM, Hundscheid RD, Smits AP, Cremers FP, van Zonneveld P, et al. (1999) Familial idiopathic premature ovarian failure: an overrated and underestimated genetic disease? Hum Reprod 14: 2455-2459.

61. Bachelot A, Rouxel A, Massin N, Dulon J, Courtillot C, et al. (2009) Phenotyping and genetic studies of 357 consecutive patients presenting with premature ovarian failure. Eur J Endocrinol 2009 Jul;161(1):179-87. doi: 10.1530/EJE-1509-0231.

62. Toniolo D (2006) X-linked premature ovarian failure: a complex disease. Curr Opin Genet Dev: 2006 Jun;2016(2003):2293-2300.

63. Vegetti W, Grazia Tibiletti M, Testa G, de Lauretis Y, Alagna F, et al. (1998) Inheritance in idiopathic premature ovarian failure: analysis of 71 cases. Hum Reprod 13: 1796-1800.

64. Davis CJ, Davison RM, Payne NN, Rodeck CH, Conway GS (2000) Female sex preponderance for idiopathic familial premature ovarian failure suggests an X chromosome defect: opinion. Hum Reprod 15: 2418-2422.

65. Sullivan SD, Castrillon DH (2011) Insights into primary ovarian insufficiency through genetically engineered mouse models. Semin Reprod Med 2011 Jul;29(4):283-98. doi: 10.1055/s-0031-1280914.

66. Sybert VP, McCauley E (2004) Turner's syndrome. N Engl J Med 351: 1227-1238.

67. Castronovo C, Rossetti R, Rusconi D, Recalcati MP, Cacciatore C, et al. (2014) Gene dosage as a relevant mechanism contributing to the determination of ovarian function in Turner syndrome. Hum Reprod 2014 Feb;29(2):368-79. doi: 10.1093/humrep/det1436.

68. Villanueva AL, Rebar RW (1983) Triple-X syndrome and premature ovarian failure. Obstet Gynecol 62: 70s-73s.

69. Goswami R, Goswami D, Kabra M, Gupta N, Dubey S, et al. (2003) Prevalence of the triple X syndrome in phenotypically normal women with premature ovarian failure and its association with autoimmune thyroid disorders. Fertil Steril 80: 1052-1054.

70. Otter M, Schrander-Stumpel CT, Curfs LM (2010) Triple X syndrome: a review of the literature. Eur J Hum Genet 2010 Mar;18(3):265-71. doi: 10.1038/ejhg.2009.1109.

71. Collen RJ, Falk RE, Lippe BM, Kaplan SA (1980) A 48,XXXX female with absence of ovaries. Am J Med Genet 6: 275-278.

72. Zinn AR (2001) The X chromosome and the ovary. J Soc Gynecol Investig 8: S34-36.

73. Moyses-Oliveira M, Guilherme Rdos S, Dantas AG, Ueta R, Perez AB, et al. (2015) Genetic mechanisms leading to primary amenorrhea in balanced X-autosome translocations. Fertil Steril 2015 May;103(5):1289-96.e2. doi: 10.1016/j.fertnstert.2015.1001.1030.

74. Genesio R, Mormile A, Licenziati MR, De Brasi D, Leone G, et al. (2015) Short stature and primary ovarian insufficiency possibly due to chromosomal position effect in a balanced X;1 translocation. Mol Cytogenet 2015 Jul 15;8:50. doi: 10.1186/s13039-13015-10154-13033.

75. Kawano Y, Narahara H, Matsui N, Miyakawa I (1998) Premature ovarian failure associated with a Robertsonian translocation. Acta Obstet Gynecol Scand 77: 467-469.

76. Hens L, Devroey P, Van Waesberghe L, Bonduelle M, Van Steirteghem AC, et al. (1989) Chromosome studies and fertility treatment in women with ovarian failure. Clin Genet 36: 81-91.

77. Therman E, Laxova R, Susman B (1990) The critical region on the human Xq. Hum Genet 85: 455-461.

78. Rizzolio F, Bione S, Sala C, Goegan M, Gentile M, et al. (2006) Chromosomal rearrangements in Xq and premature ovarian failure: mapping of 25 new cases and review of the literature. Hum Reprod: 2006 Jun;2021(2006):1477-2083.

79. Marozzi A, Manfredini E, Tibiletti MG, Furlan D, Villa N, et al. (2000) Molecular definition of Xq common-deleted region in patients affected by premature ovarian failure. Hum Genet 107: 304-311.

80. Dixit H, Rao L, Padmalatha V, Raseswari T, Kapu AK, et al. (2010) Genes governing premature ovarian failure. Reprod Biomed Online 2010 Jun;20(6):724-40. doi: 10.1016/j.rbmo.2010.1002.1018.

81. Fierabracci A, Bizzarri C, Palma A, Milillo A, Bellacchio E, et al. (2012) A novel heterozygous mutation of the AIRE gene in a patient with autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy syndrome (APECED). Gene 2012 Dec 10;511(1):113-7. doi: 10.1016/j.gene.2012.1009.1029.

82. Schiffmann R, Tedeschi G, Kinkel RP, Trapp BD, Frank JA, et al. (1997) Leukodystrophy in patients with ovarian dysgenesis. Ann Neurol 41: 654-661.

83. Harris SE, Chand AL, Winship IM, Gersak K, Aittomaki K, et al. (2002) Identification of novel mutations in FOXL2 associated with premature ovarian failure. Mol Hum Reprod 8: 729-733.

84. Laissue P, Lakhal B, Benayoun BA, Dipietromaria A, Braham R, et al. (2009) Functional evidence implicating FOXL2 in non-syndromic premature ovarian failure and in the regulation of the transcription factor OSR2. J Med Genet 2009 Jul;46(7):455-7. doi: 10.1136/jmg.2008.065086.

85. Schmidt D, Ovitt CE, Anlag K, Fehsenfeld S, Gredsted L, et al. (2004) The murine winged-helix transcription factor Foxl2 is required for granulosa cell differentiation and ovary maintenance. Development: 2004 Feb;2131(2004):2933-2042.

86. Laissue P (2015) Aetiological coding sequence variants in non-syndromic premature ovarian failure: From genetic linkage analysis to next generation sequencing. Mol Cell Endocrinol 2015 Aug 15;411:243-57. doi: 10.1016/j.mce.2015.1005.1005.

87. Wittenberger MD, Hagerman RJ, Sherman SL, McConkie-Rosell A, Welt CK, et al. (2007) The FMR1 premutation and reproduction. Fertil Steril: 2007 Mar;2087(2003):2456-2065.

88. Sullivan AK, Marcus M, Epstein MP, Allen EG, Anido AE, et al. (2005) Association of FMR1 repeat size with ovarian dysfunction. Hum Reprod: 2005 Feb;2020(2002):2402-2012.

89. Murray A, Webb J, Dennis N, Conway G, Morton N (1999) Microdeletions in FMR2 may be a significant cause of premature ovarian failure. J Med Genet 36: 767-770.

90. Dube JL, Wang P, Elvin J, Lyons KM, Celeste AJ, et al. (1998) The bone morphogenetic protein 15 gene is X-linked and expressed in oocytes. Mol Endocrinol 12: 1809-1817.

91. Moore RK, Shimasaki S (2005) Molecular biology and physiological role of the oocyte factor, BMP-15. Mol Cell Endocrinol 234: 67-73.

92. Hashimoto O, Moore RK, Shimasaki S (2005) Posttranslational processing of mouse and human BMP-15: potential implication in the determination of ovulation quota. Proc Natl Acad Sci U S A: 2005 Apr 2012;2102(2015):5426-2031.

93. Galloway SM, McNatty KP, Cambridge LM, Laitinen MP, Juengel JL, et al. (2000) Mutations in an oocyte-derived growth factor gene (BMP15) cause increased ovulation rate and infertility in a dosage-sensitive manner. Nat Genet 25: 279-283.

94. Di Pasquale E, Beck-Peccoz P, Persani L (2004) Hypergonadotropic ovarian failure associated with an inherited mutation of human bone morphogenetic protein-15 (BMP15) gene. Am J Hum Genet: 2004 Jul;2075(2001):2106-2011.

95. Rossetti R, Di Pasquale E, Marozzi A, Bione S, Toniolo D, et al. (2009) BMP15 mutations associated with primary ovarian insufficiency cause a defective production of bioactive protein. Hum Mutat: 2009 May;2030(2005):2804-2010.

96. Di Pasquale E, Rossetti R, Marozzi A, Bodega B, Borgato S, et al. (2006) Identification of new variants of human BMP15 gene in a large cohort of women with premature ovarian failure. J Clin Endocrinol Metab: 2006 May;2091(2005):1976-2009.

97. Dixit H, Rao LK, Padmalatha VV, Kanakavalli M, Deenadayal M, et al. (2006) Missense mutations in the BMP15 gene are associated with ovarian failure. Hum Genet: 2006 May;2119(2004):2408-2015.

98. Laissue P, Christin-Maitre S, Touraine P, Kuttenn F, Ritvos O, et al. (2006) Mutations and sequence variants in GDF9 and BMP15 in patients with premature ovarian failure. Eur J Endocrinol 154: 739-744.

99. Settas N, Anapliotou M, Kanavakis E, Fryssira H, Sofocleous C, et al. (2015) A novel FOXL2 gene mutation and BMP15 variants in a woman with primary ovarian insufficiency and blepharophimosis-ptosis-epicanthus inversus syndrome. Menopause 18: 18.

100. Tiotiu D, Alvaro Mercadal B, Imbert R, Verbist J, Demeestere I, et al. (2010) Variants of the BMP15 gene in a cohort of patients with premature ovarian failure. Hum Reprod 2010 Jun;25(6):1581-7. doi: 10.1093/humrep/deq1073.

101. Aaltonen J, Laitinen MP, Vuojolainen K, Jaatinen R, Horelli-Kuitunen N, et al. (1999) Human growth differentiation factor 9 (GDF-9) and its novel homolog GDF-9B are expressed in oocytes during early folliculogenesis. J Clin Endocrinol Metab 84: 2744-2750.

102. Dong J, Albertini DF, Nishimori K, Kumar TR, Lu N, et al. (1996) Growth differentiation factor-9 is required during early ovarian folliculogenesis. Nature 383: 531-535.

103. Takebayashi K, Takakura K, Wang H, Kimura F, Kasahara K, et al. (2000) Mutation analysis of the growth differentiation factor-9 and -9B genes in patients with premature ovarian failure and polycystic ovary syndrome. Fertil Steril 74: 976-979.

104. Dixit H, Rao LK, Padmalatha V, Kanakavalli M, Deenadayal M, et al. (2005) Mutational screening of the coding region of growth differentiation factor 9 gene in Indian women with ovarian failure. Menopause: 2005 Nov-Dec;2012(2006):2749-2054.

105. Kovanci E, Rohozinski J, Simpson JL, Heard MJ, Bishop CE, et al. (2007) Growth differentiating factor-9 mutations may be associated with premature ovarian failure. Fertil Steril: 2007 Jan;2087(2001):2143-2006.

106. Wang TT, Ke ZH, Song Y, Chen LT, Chen XJ, et al. (2013) Identification of a mutation in GDF9 as a novel cause of diminished ovarian reserve in young women. Hum Reprod 2013 Sep;28(9):2473-81. doi: 10.1093/humrep/det1291.

107. Groome NP, Illingworth PJ, O'Brien M, Cooke I, Ganesan TS, et al. (1994) Detection of dimeric inhibin throughout the human menstrual cycle by two-site enzyme immunoassay. Clin Endocrinol 40: 717-723.

108. Shelling AN, Burton KA, Chand AL, van Ee CC, France JT, et al. (2000) Inhibin: a candidate gene for premature ovarian failure. Hum Reprod 15: 2644-2649.

109. Marozzi A, Porta C, Vegetti W, Crosignani PG, Tibiletti MG, et al. (2002) Mutation analysis of the inhibin alpha gene in a cohort of Italian women affected by ovarian failure. Hum Reprod 17: 1741-1745.

110. Harris SE, Chand AL, Winship IM, Gersak K, Nishi Y, et al. (2005) INHA promoter polymorphisms are associated with premature ovarian failure. Mol Hum Reprod: 2005 Nov;2011(2011):2779-2084.

111. Corre T, Schuettler J, Bione S, Marozzi A, Persani L, et al. (2009) A large-scale association study to assess the impact of known variants of the human INHA gene on premature ovarian failure. Hum Reprod 2009 Aug;24(8):2023-8. doi: 10.1093/humrep/dep1090.

112. Woad KJ, Pearson SM, Harris SE, Gersak K, Shelling AN (2009) Investigating the association between inhibin alpha gene promoter polymorphisms and premature ovarian failure. Fertil Steril 2009 Jan;91(1):62-6. doi: 10.1016/j.fertnstert.2007.1011.1012.

113. Abel MH, Wootton AN, Wilkins V, Huhtaniemi I, Knight PG, et al. (2000) The effect of a null mutation in the follicle-stimulating hormone receptor gene on mouse reproduction. Endocrinology 141: 1795-1803.

114. Zhang FP, Poutanen M, Wilbertz J, Huhtaniemi I (2001) Normal prenatal but arrested postnatal sexual development of luteinizing hormone receptor knockout (LuRKO) mice. Mol Endocrinol 15: 172-183.

115. Aittomaki K, Lucena JL, Pakarinen P, Sistonen P, Tapanainen J, et al. (1995) Mutation in the follicle-stimulating hormone receptor gene causes hereditary hypergonadotropic ovarian failure. Cell 82: 959-968.

116. Katari S, Wood-Trageser MA, Jiang H, Kalynchuk E, Muzumdar R, et al. (2015) Novel Inactivating Mutation of the FSH Receptor in Two Siblings of Indian Origin With Premature Ovarian Failure. J Clin Endocrinol Metab 2015 Jun;100(6):2154-7. doi: 10.1210/jc.2015-1401.

117. Latronico AC, Anasti J, Arnhold IJ, Rapaport R, Mendonca BB, et al. (1996) Brief report: testicular and ovarian resistance to luteinizing hormone caused by inactivating mutations of the luteinizing hormone-receptor gene. N Engl J Med 334: 507-512.

118. Jeyasuria P, Ikeda Y, Jamin SP, Zhao L, De Rooij DG, et al. (2004) Cell-specific knockout of steroidogenic factor 1 reveals its essential roles in gonadal function. Mol Endocrinol: 2004 Jul;2018(2007):1610-2009.

119. Lourenco D, Brauner R, Lin L, De Perdigo A, Weryha G, et al. (2009) Mutations in NR5A1 associated with ovarian insufficiency. N Engl J Med 2009 Mar 19;360(12):1200-10. doi: 10.1056/NEJMoa0806228.

120. Castrillon DH, Miao L, Kollipara R, Horner JW, DePinho RA (2003) Suppression of ovarian follicle activation in mice by the transcription factor Foxo3a. Science 301: 215-218.

121. Liu L, Rajareddy S, Reddy P, Du C, Jagarlamudi K, et al. (2007) Infertility caused by retardation of follicular development in mice with oocyte-specific expression of Foxo3a. Development 134: 199-209.

122. Huntriss J, Hinkins M, Picton HM (2006) cDNA cloning and expression of the human NOBOX gene in oocytes and ovarian follicles. Mol Hum Reprod: 2006 May;2012(2005):2283-2009.

123. Rajkovic A, Pangas SA, Ballow D, Suzumori N, Matzuk MM (2004) NOBOX deficiency disrupts early folliculogenesis and oocyte-specific gene expression. Science 305: 1157-1159.

124. Qin Y, Choi Y, Zhao H, Simpson JL, Chen ZJ, et al. (2007) NOBOX homeobox mutation causes premature ovarian failure. Am J Hum Genet: 2007 Sep;2081(2003):2576-2081.

125. Bouilly J, Bachelot A, Broutin I, Touraine P, Binart N (2011) Novel NOBOX loss-of-function mutations account for 6.2% of cases in a large primary ovarian insufficiency cohort. Hum Mutat 2011 Oct;32(10):1108-13. doi: 10.1002/humu.21543.

126. Bouilly J, Roucher-Boulez F, Gompel A, Bry-Gauillard H, Azibi K, et al. (2015) New NOBOX mutations identified in a large cohort of women with primary ovarian insufficiency decrease KIT-L expression. J Clin Endocrinol Metab 2015 Mar;100(3):994-1001. doi: 10.1210/jc.2014-2761.

127. Hu W, Gauthier L, Baibakov B, Jimenez-Movilla M, Dean J (2010) FIGLA, a basic helix-loop-helix transcription factor, balances sexually dimorphic gene expression in postnatal oocytes. Mol Cell Biol 2010 Jul;30(14):3661-71. doi: 10.1128/MCB.00201-00210.

128. Zhao H, Chen ZJ, Qin Y, Shi Y, Wang S, et al. (2008) Transcription factor FIGLA is mutated in patients with premature ovarian failure. Am J Hum Genet 2008 Jun;82(6):1342-8. doi: 10.1016/j.ajhg.2008.1004.1018.

129. Bione S, Sala C, Manzini C, Arrigo G, Zuffardi O, et al. (1998) A human homologue of the Drosophila melanogaster diaphanous gene is disrupted in a patient with premature ovarian failure: evidence for conserved function in oogenesis and implications for human sterility. Am J Hum Genet 62: 533-541.

130. Rizzolio F, Sala C, Alboresi S, Bione S, Gilli S, et al. (2007) Epigenetic control of the critical region for premature ovarian failure on autosomal genes translocated to the X chromosome: a hypothesis. Hum Genet: 2007 May;2121(2003-2004):2441-2050.

131. Lacombe A, Lee H, Zahed L, Choucair M, Muller JM, et al. (2006) Disruption of POF1B binding to nonmuscle actin filaments is associated with premature ovarian failure. Am J Hum Genet: 2006 Jul;2079(2001):2113-2009.

132. Tanaka K, Miyamoto N, Shouguchi-Miyata J, Ikeda JE (2006) HFM1, the human homologue of yeast Mer3, encodes a putative DNA helicase expressed specifically in germ-line cells. DNA Seq 17: 242-246.

133. Guiraldelli MF, Eyster C, Wilkerson JL, Dresser ME, Pezza RJ (2013) Mouse HFM1/Mer3 is required for crossover formation and complete synapsis of homologous chromosomes during meiosis. PLoS Genet 2013 Mar;9(3):e1003383. doi: 10.1371/journal.pgen.1003383.

134. Wang J, Zhang W, Jiang H, Wu BL (2014) Mutations in HFM1 in recessive primary ovarian insufficiency. N Engl J Med: 2014 Mar 2016;2370(2010):2972-2014.

135. Caburet S, Arboleda VA, Llano E, Overbeek PA, Barbero JL, et al. (2014) Mutant cohesin in premature ovarian failure. N Engl J Med: 2014 Mar 2016;2370(2010):2943-2019.

136. Le Quesne Stabej P, Williams HJ, James C, Tekman M, Stanescu HC, et al. (2015) STAG3 truncating variant as the cause of primary ovarian insufficiency. Eur J Hum Genet: 2015 Jun 2010.

137. Suzuki H, Tsuda M, Kiso M, Saga Y (2008) Nanos3 maintains the germ cell lineage in the mouse by suppressing both Bax-dependent and -independent apoptotic pathways. Dev Biol 2008 Jun 1;318(1):133-42. doi: 10.1016/j.ydbio.2008.1003.1020.

138. Wu X, Wang B, Dong Z, Zhou S, Liu Z, et al. (2013) A NANOS3 mutation linked to protein degradation causes premature ovarian insufficiency. Cell Death Dis: 2013 Oct 2013;2014:e2825.

139. Santos MG, Machado AZ, Martins CN, Domenice S, Costa EM, et al. (2014) Homozygous inactivating mutation in NANOS3 in two sisters with primary ovarian insufficiency. Biomed Res Int 2014;2014:787465. doi: 10.1155/2014/787465.

140. AlAsiri S, Basit S, Wood-Trageser MA, Yatsenko SA, Jeffries EP, et al. (2015) Exome sequencing reveals MCM8 mutation underlies ovarian failure and chromosomal instability. J Clin Invest 2015 Jan;125(1):258-62. doi: 10.1172/JCI78473.

141. Wood-Trageser MA, Gurbuz F, Yatsenko SA, Jeffries EP, Kotan LD, et al. (2014) MCM9 mutations are associated with ovarian failure, short stature, and chromosomal instability. Am J Hum Genet: 2014 Dec 2014;2095(2016):2754-2062.

142. Lutzmann M, Grey C, Traver S, Ganier O, Maya-Mendoza A, et al. (2012) MCM8- and MCM9-deficient mice reveal gametogenesis defects and genome instability due to impaired homologous recombination. Mol Cell 2012 Aug 24;47(4):523-34. doi: 10.1016/j.molcel.2012.1005.1048.

143. Costa Y, Cooke HJ (2007) Dissecting the mammalian synaptonemal complex using targeted mutations. Chromosome Res 15: 579-589.

144. Costa Y, Speed R, Ollinger R, Alsheimer M, Semple CA, et al. (2005) Two novel proteins recruited by synaptonemal complex protein 1 (SYCP1) are at the centre of meiosis. J Cell Sci 118: 2755-2762.

145. Hamer G, Gell K, Kouznetsova A, Novak I, Benavente R, et al. (2006) Characterization of a novel meiosis-specific protein within the central element of the synaptonemal complex. J Cell Sci: 2006 Oct 2001;2119(Pt 2019):4025-2032.

146. de Vries L, Behar DM, Smirin-Yosef P, Lagovsky I, Tzur S, et al. (2014) Exome sequencing reveals SYCE1 mutation associated with autosomal recessive primary ovarian insufficiency. J Clin Endocrinol Metab 2014 Oct;99(10):E2129-32. doi: 10.1210/jc.2014-1268.

147. Bolcun-Filas E, Hall E, Speed R, Taggart M, Grey C, et al. (2009) Mutation of the mouse Syce1 gene disrupts synapsis and suggests a link between synaptonemal complex structural components and DNA repair. PLoS Genet 2009 Feb;5(2):e1000393. doi: 10.1371/journal.pgen.1000393.

148. Knauff EA, Franke L, van Es MA, van den Berg LH, van der Schouw YT, et al. (2009) Genome-wide association study in premature ovarian failure patients suggests ADAMTS19 as a possible candidate gene. Hum Reprod 2009 Sep;24(9):2372-8. doi: 10.1093/humrep/dep1197.

149. Kang H, Lee SK, Kim MH, Song J, Bae SJ, et al. (2008) Parathyroid hormone-responsive B1 gene is associated with premature ovarian failure. Hum Reprod 2008 Jun;23(6):1457-65. doi: 10.1093/humrep/den1086.

150. Pyun JA, Kim S, Cha DH, Kwack K (2013) Epistasis between IGF2R and ADAMTS19 polymorphisms associates with premature ovarian failure. Hum Reprod 2013 Nov;28(11):3146-54. doi: 10.1093/humrep/det1365.

151. Qin Y, Zhao H, Xu J, Shi Y, Li Z, et al. (2012) Association of 8q22.3 locus in Chinese Han with idiopathic premature ovarian failure (POF). Hum Mol Genet 2012 Jan 15;21(2):430-6. doi: 10.1093/hmg/ddr1462.

152. Aboura A, Dupas C, Tachdjian G, Portnoi MF, Bourcigaux N, et al. (2009) Array comparative genomic hybridization profiling analysis reveals deoxyribonucleic acid copy number variations associated with premature ovarian failure. J Clin Endocrinol Metab 2009 Nov;94(11):4540-6. doi: 10.1210/jc.2009-0186.

153. Liao C, Fu F, Yang X, Sun YM, Li DZ (2011) Analysis of Chinese women with primary ovarian insufficiency by high resolution array-comparative genomic hybridization. Chin Med J 124: 1739-1742.

154. Perry JR, Corre T, Esko T, Chasman DI, Fischer K, et al. (2013) A genome-wide association study of early menopause and the combined impact of identified variants. Hum Mol Genet 2013 Apr 1;22(7):1465-72. doi: 10.1093/hmg/dds1551.

155. Day FR, Ruth KS, Thompson DJ, Lunetta KL, Pervjakova N, et al. (2015) Large-scale genomic analyses link reproductive aging to hypothalamic signaling, breast cancer susceptibility and BRCA1-mediated DNA repair. Nat Genet: 2015 Sep 2028.

156. Kirov G (2015) CNVs in neuropsychiatric disorders. Hum Mol Genet 2015 Oct 15;24(R1):R45-9. doi: 10.1093/hmg/ddv1253.

157. Martin J, O'Donovan MC, Thapar A, Langley K, Williams N (2015) The relative contribution of common and rare genetic variants to ADHD. Transl Psychiatry: 2015 Feb 2010;2015:e2506.

158. Pinto D, Pagnamenta AT, Klei L, Anney R, Merico D, et al. (2010) Functional impact of global rare copy number variation in autism spectrum disorders. Nature 2010 Jul 15;466(7304):368-72. doi: 10.1038/nature09146.

159. Norling A, Hirschberg AL, Rodriguez-Wallberg KA, Iwarsson E, Wedell A, et al. (2014) Identification of a duplication within the GDF9 gene and novel candidate genes for primary ovarian insufficiency (POI) by a customized high-resolution array comparative genomic hybridization platform. Hum Reprod 2014 Aug;29(8):1818-27. doi: 10.1093/humrep/deu1149.

160. Ledig S, Ropke A, Wieacker P (2010) Copy number variants in premature ovarian failure and ovarian dysgenesis. Sex Dev 2010 Sep;4(4-5):225-32. doi: 10.1159/000314958.

161. McGuire MM, Bowden W, Engel NJ, Ahn HW, Kovanci E, et al. (2011) Genomic analysis using high-resolution single-nucleotide polymorphism arrays reveals novel microdeletions associated with premature ovarian failure. Fertil Steril 2011 Apr;95(5):1595-600. doi: 10.1016/j.fertnstert.2010.1012.1052.

162. Zhen XM, Sun YM, Qiao J, Li R, Wang LN, et al. (2013) [Genome-wide copy number scan in Chinese patients with premature ovarian failure]. Beijing Da Xue Xue Bao 45: 841-847.

163. Dudding TE, Lawrence O, Winship I, Froyen G, Vandewalle J, et al. (2010) Array comparative genomic hybridization for the detection of submicroscopic copy number variations of the X chromosome in women with premature ovarian failure. Hum Reprod 2010 Dec;25(12):3159-60; author reply 3160-1. doi: 10.1093/humrep/deq1284.

164. Quilter CR, Karcanias AC, Bagga MR, Duncan S, Murray A, et al. (2010) Analysis of X chromosome genomic DNA sequence copy number variation associated with premature ovarian failure (POF). Hum Reprod 2010 Aug;25(8):2139-50. doi: 10.1093/humrep/deq1158.

165. Knauff EA, Blauw HM, Pearson PL, Kok K, Wijmenga C, et al. (2011) Copy number variants on the X chromosome in women with primary ovarian insufficiency. Fertil Steril 2011 Apr;95(5):1584-8.e1. doi: 10.1016/j.fertnstert.2011.1001.1018.

166. Fonseca DJ, Patino LC, Suarez YC, de Jesus Rodriguez A, Mateus HE, et al. (2015) Next generation sequencing in women affected by nonsyndromic premature ovarian failure displays new potential causative genes and mutations. Fertil Steril 2015 Jul;104(1):154-162.e2. doi: 10.1016/j.fertnstert.2015.1004.1016.

167. Qin Y, Guo T, Li G, Tang TS, Zhao S, et al. (2015) CSB-PGBD3 Mutations Cause Premature Ovarian Failure. PLoS Genet 2015 Jul 28;11(7):e1005419. doi: 10.1371/journal.pgen.1005419.

168. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402-408.

169. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, et al. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010 Sep;20(9):1297-303. doi: 10.1101/gr.107524.107110.

170. Kumar P, Henikoff S, Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 2009;4(7):1073-81. doi: 10.1038/nprot.2009.1086.

171. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, et al. (2010) A method and server for predicting damaging missense mutations. Nat Methods: 2010 Apr;2017(2014):2248-2019.

172. Schwarz JM, Cooper DN, Schuelke M, Seelow D (2014) MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods: 2014 Apr;2011(2014):2361-2012.

173. Wang L, White KL, Reed WA, Campbell KD (2005) Dynamic changes to the inositol 1,4,5-trisphosphate and ryanodine receptors during maturation of bovine oocytes. Cloning Stem Cells 7: 306-320.

174. Toranzo GS, Buhler MC, Buhler MI (2014) Participation of IP3R, RyR and L-type Ca2+ channel in the nuclear maturation of Rhinella arenarum oocytes. Zygote 2014 May;22(2):110-23. doi: 10.1017/S0967199412000287.

175. Simpson JL, Rajkovic A (1999) Ovarian differentiation and gonadal failure. Am J Med Genet 89: 186-200.

176. Mandon-Pepin B, Oustry-Vaiman A, Vigier B, Piumi F, Cribiu E, et al. (2003) Expression profiles and chromosomal localization of genes controlling meiosis and follicular development in the sheep ovary. Biol Reprod 68: 985-995.

177. Gohin M, Fournier E, Dufort I, Sirard MA (2014) Discovery, identification and sequence analysis of RNAs selected for very short or long poly A tail in immature bovine oocytes. Mol Hum Reprod 2014 Feb;20(2):127-38. doi: 10.1093/molehr/gat1080.

178. Chalupnikova K, Solc P, Sulimenko V, Sedlacek R, Svoboda P (2014) An oocyte-specific ELAVL2 isoform is a translational repressor ablated from meiotically competent antral oocytes. Cell Cycle 2014;13(7):1187-200. doi: 10.4161/cc.28107.

179. Nakamuta N, Kobayashi S (2007) Expression of p63 in the mouse ovary. J Reprod Dev: 2007 Jun;2053(2003):2691-2007.

180. Deutsch GB, Zielonka EM, Coutandin D, Weber TA, Schafer B, et al. (2011) DNA damage in oocytes induces a switch of the quality control factor TAp63alpha from dimer to tetramer. Cell: 2011 Feb 2018;2144(2014):2566-2076.

181. Kim DA, Suh EK (2014) Defying DNA double-strand break-induced death during prophase I meiosis by temporal TAp63alpha phosphorylation regulation in developing mouse oocytes. Mol Cell Biol 2014 Apr;34(8):1460-73. doi: 10.1128/MCB.01223-01213.

182. Suh EK, Yang A, Kettenbach A, Bamberger C, Michaelis AH, et al. (2006) p63 protects the female germ line during meiotic arrest. Nature: 2006 Nov 2030;2444(7119):2624-2008.

183. Kayampilly PP, Menon KM (2009) Follicle-stimulating hormone inhibits adenosine 5'-monophosphate-activated protein kinase activation and promotes cell proliferation of primary granulosa cells in culture through an Akt-dependent pathway. Endocrinology 2009 Feb;150(2):929-35. doi: 10.1210/en.2008-1032.

184. Mayes MA, Laforest MF, Guillemette C, Gilchrist RB, Richard FJ (2007) Adenosine 5'-monophosphate kinase-activated protein kinase (PRKA) activators delay meiotic resumption in porcine oocytes. Biol Reprod: 2007 Apr;2076(2004):2589-2097.

185. Downs SM, Hudson ER, Hardie DG (2002) A potential role for AMP-activated protein kinase in meiotic induction in mouse oocytes. Dev Biol 245: 200-212.

186. Chen J, Hudson E, Chi MM, Chang AS, Moley KH, et al. (2006) AMPK regulation of mouse oocyte meiotic resumption in vitro. Dev Biol: 2006 Mar 2015;2291(2002):2227-2038.

187. Bertoldo MJ, Guibert E, Faure M, Rame C, Foretz M, et al. (2015) Specific deletion of AMP-activated protein kinase (alpha1AMPK) in murine oocytes alters junctional protein expression and mitochondrial physiology. PLoS One 2015 Mar 13;10(3):e0119680. doi: 10.1371/journal.pone.0119680.

188. Chen J, Downs SM (2008) AMP-activated protein kinase is involved in hormone-induced mouse oocyte meiotic maturation in vitro. Dev Biol: 2008 Jan 2001;2313(2001):2047-2057.

189. Melani M, Simpson KJ, Brugge JS, Montell D (2008) Regulation of cell adhesion and collective cell migration by hindsight and its human homolog RREB1. Curr Biol: 2008 Apr 2008;2018(2007):2532-2007.

190. Bai H, Li Y, Gao H, Dong Y, Han P, et al. (2015) Histone methyltransferase SMYD3 regulates the expression of transcriptional factors during bovine oocyte maturation and early embryonic development. Cytotechnology 7: 7.

191. Beverdam A, Koopman P (2006) Expression profiling of purified mouse gonadal somatic cells during the critical time window of sex determination reveals novel candidate genes for human sexual dysgenesis syndromes. Hum Mol Genet: 2006 Feb 2001;2015(2003):2417-2031.

192. Lawson KA, Dunn NR, Roelen BA, Zeinstra LM, Davis AM, et al. (1999) Bmp4 is required for the generation of primordial germ cells in the mouse embryo. Genes Dev 13: 424-436.

193. Fujiwara T, Dunn NR, Hogan BL (2001) Bone morphogenetic protein 4 in the extraembryonic mesoderm is required for allantois development and the localization and survival of primordial germ cells in the mouse. Proc Natl Acad Sci U S A: 2001 Nov 2020;2098(2024):13739-13744.

194. Park ES, Woods DC, Tilly JL (2013) Bone morphogenetic protein 4 promotes mammalian oogonial stem cell differentiation via Smad1/5/8 signaling. Fertil Steril 2013 Nov;100(5):1468-75. doi: 10.1016/j.fertnstert.2013.1007.1978.

195. Tanwar PS, O'Shea T, McFarlane JR (2008) In vivo evidence of role of bone morphogenetic protein-4 in the mouse ovary. Anim Reprod Sci: 2008 Jul;2106(2003-2004):2232-2040.

196. Bai T, Yang J, Chen B, Wang B, Ma X, et al. (2014) Genetic analysis of BMP4 gene in Chinese Han female population with premature ovarian insufficiency. Climacteric 2014 Jun;17(3):304-6. doi: 10.3109/13697137.13692013.13876619.

197. Nakano M, Kakiuchi Y, Shimada Y, Ohyama M, Ogiwara Y, et al. (2009) MOV10 as a novel telomerase-associated protein. Biochem Biophys Res Commun 2009 Oct 16;388(2):328-32. doi: 10.1016/j.bbrc.2009.1008.1002.

198. Pan H, O'Brien M J, Wigglesworth K, Eppig JJ, Schultz RM (2005) Transcript profiling during mouse oocyte development and the effect of gonadotropin priming and development in vitro. Dev Biol: 2005 Oct 2015;2286(2002):2493-2506.

199. Antenos M, Lei L, Xu M, Malipatil A, Kiesewetter S, et al. (2011) Role of PCSK5 expression in mouse ovarian follicle development: identification of the inhibin alpha- and beta-subunits as candidate substrates. PLoS One: 2011 Mar 2018;2016(2013):e17348.

200. Bae JA, Park HJ, Seo YM, Roh J, Hsueh AJ, et al. (2008) Hormonal regulation of proprotein convertase subtilisin/kexin type 5 expression during ovarian follicle development in the rat. Mol Cell Endocrinol 2008 Jul 16;289(1-2):29-37. doi: 10.1016/j.mce.2008.1004.1006.

201. Johnstone O, Deuring R, Bock R, Linder P, Fuller MT, et al. (2005) Belle is a Drosophila DEAD-box protein required for viability and in the germ line. Dev Biol 277: 92-101.

202. Wang C, Li SJ, Yu WH, Xin QW, Li C, et al. (2011) Cloning and expression profiling of the VLDLR gene associated with egg performance in duck (Anas platyrhynchos). Genet Sel Evol: 2011 Aug 2015;2043:2029.

203. Bujo H, Yamamoto T, Hayashi K, Hermann M, Nimpf J, et al. (1995) Mutant oocytic low density lipoprotein receptor gene family member causes atherosclerosis and female sterility. Proc Natl Acad Sci U S A 92: 9905-9909.

204. Bujo H, Hermann M, Kaderli MO, Jacobsen L, Sugawara S, et al. (1994) Chicken oocyte growth is mediated by an eight ligand binding repeat member of the LDL receptor family. Embo J 13: 5165-5175.

205. Argov N, Sklan D (2004) Expression of mRNA of lipoprotein receptor related protein 8, low density lipoprotein receptor, and very low density lipoprotein receptor in bovine ovarian cells during follicular development and corpus luteum formation and regression. Mol Reprod Dev 68: 169-175.

206. Duncan FE, Moss SB, Schultz RM, Williams CJ (2005) PAR-3 defines a central subdomain of the cortical actin cap in mouse eggs. Dev Biol 280: 38-47.

207. Nakaya M, Fukui A, Izumi Y, Akimoto K, Asashima M, et al. (2000) Meiotic maturation induces animal-vegetal asymmetric distribution of aPKC and ASIP/PAR-3 in Xenopus oocytes. Development 127: 5021-5031.

208. Shimizu T, Jayawardana BC, Nishimoto H, Kaneko E, Tetsuka M, et al. (2006) Hormonal regulation and differential expression of neuropilin (NRP)-1 and NRP-2 genes in bovine granulosa cells. Reproduction 131: 555-559.

209. Tay J, Richter JD (2001) Germ cell differentiation and synaptonemal complex formation are disrupted in CPEB knockout mice. Dev Cell 1: 201-213.

210. Nishimura Y, Kano K, Naito K (2010) Porcine CPEB1 is involved in Cyclin B translation and meiotic resumption in porcine oocytes. Anim Sci J: 2010 Aug 2011;2081(2014):2444-2052.

211. Kim JH, Richter JD (2008) Measuring CPEB-mediated cytoplasmic polyadenylation-deadenylation in Xenopus laevis oocytes and egg extracts. Methods Enzymol: 2008;2448:2119-2038.

212. Sugimoto M, Sasaki S, Watanabe T, Nishimura S, Ideta A, et al. (2010) Ionotropic glutamate receptor AMPA 1 is associated with ovulation rate. PLoS One: 2010 Nov 2013;2015(2011):e13817.

213. Cushman RA, Miles JR, Rempel LA, McDaneld TG, Kuehn LA, et al. (2013) Identification of an ionotropic glutamate receptor AMPA1/GRIA1 polymorphism in crossbred beef cows differing in fertility. J Anim Sci 2013 Jun;91(6):2640-6. doi: 10.2527/jas.2012-5950.

214. Otsuka Y, Yanaihara A, Iwasaki S, Hasegawa J, Yanaihara T, et al. (2005) Localization and gene expression of steroid sulfatase by RT-PCR in cumulus cells and relationship to serum FSH levels observed during in vitro fertilization. J Exp Clin Assist Reprod 2: 6.

215. Sugawara T, Fujimoto S (2004) The potential function of steroid sulphatase activity in steroid production and steroidogenic acute regulatory protein expression. Biochem J 380: 153-160.

216. Gupta PS, Folger JK, Rajput SK, Lv L, Yao J, et al. (2014) Regulation and regulatory role of WNT signaling in potentiating FSH action during bovine dominant follicle selection. PLoS One 2014 Jun 17;9(6):e100201. doi: 10.1371/journal.pone.0100201.

217. Huang X, Ding L, Pan R, Ma PF, Cheng PP, et al. (2013) WHAMM is required for meiotic spindle migration and asymmetric cytokinesis in mouse oocytes. Histochem Cell Biol 2013 Apr;139(4):525-34. doi: 10.1007/s00418-00012-01051-z.

218. Smith TH, Stedronsky K, Morgan B, McGowan RA (2006) Identification and isolation of a BTB-POZ-containing gene expressed in oocytes and early embryos of the zebrafish Danio rerio. Genome 49: 808-814.

219. Mahoney MG, Tang W, Xiang MM, Moss SB, Gerton GL, et al. (1998) Translocation of the zinc finger protein basonuclin from the mouse germ cell nucleus to the midpiece of the spermatozoon during spermiogenesis. Biol Reprod 59: 388-394.

220. Ma J, Zeng F, Schultz RM, Tseng H (2006) Basonuclin: a novel mammalian maternal-effect gene. Development: 2006 May;2133(2010):2053-2062.

221. Hirayama S, Bajari TM, Nimpf J, Schneider WJ (2003) Receptor-mediated chicken oocyte growth: differential expression of endophilin isoforms in developing follicles. Biol Reprod: 2003 May;2068(2005):1850-2060.

222. Awe JP, Byrne JA (2013) Identifying candidate oocyte reprogramming factors using cross-species global transcriptional analysis. Cell Reprogram 2013 Apr;15(2):126-33. doi: 10.1089/cell.2012.0060.

223. Xiong B, Sun SC, Lin SL, Li M, Xu BZ, et al. (2008) Involvement of Polo-like kinase 1 in MEK1/2-regulated spindle formation during mouse oocyte meiosis. Cell Cycle: 2008 Jun 2015;2007(2012):1804-2009.

224. Xiong B, Yu LZ, Wang Q, Ai JS, Yin S, et al. (2007) Regulation of intracellular MEK1/2 translocation in mouse oocytes: cytoplasmic dynein/dynactin-mediated poleward transport and cyclin B degradation-dependent release from spindle poles. Cell Cycle: 2007 Jun 2015;2006(2012):1521-2007.

225. Leonardsen L, Wiersma A, Baltsen M, Byskov AG, Andersen CY (2000) Regulation of spontaneous and induced resumption of meiosis in mouse oocytes by different intracellular pathways. J Reprod Fertil 120: 377-383.

226. Luo LL, Chen XC, Fu YC, Xu JJ, Li L, et al. (2012) The effects of caloric restriction and a high-fat diet on ovarian lifespan and the expression of SIRT1 and SIRT6 proteins in rats. Aging Clin Exp Res 2012 Apr;24(2):125-33. doi: 10.3275/7660.

227. Mostoslavsky R, Chua KF, Lombard DB, Pang WW, Fischer MR, et al. (2006) Genomic instability and aging-like phenotype in the absence of mammalian SIRT6. Cell 124: 315-329.

228. Muramatsu H, Zou P, Kurosawa N, Ichihara-Tanaka K, Maruyama K, et al. (2006) Female infertility in mice deficient in midkine and pleiotrophin, which form a distinct family of growth factors. Genes Cells 11: 1405-1417.

229. Ohyama Y, Miyamoto K, Minamino N, Matsuo H (1994) Isolation and identification of midkine and pleiotrophin in bovine follicular fluid. Mol Cell Endocrinol 105: 203-208.

230. Kleinjan DJ, van Heyningen V (1998) Position effect in human genetic disease. Hum Mol Genet 7: 1611-1618.

231. Kleinjan DA, van Heyningen V (2005) Long-range control of gene expression: emerging mechanisms and disruption in disease. Am J Hum Genet: 2005 Jan;2076(2001):2008-2032.

232. Nilsson EE, Larsen G, Skinner MK (2014) Roles of Gremlin 1 and Gremlin 2 in regulating ovarian primordial to primary follicle transition. Reproduction 2014 Jun;147(6):865-74. doi: 10.1530/REP-1514-0005.

233. Leader B, Lim H, Carabatsos MJ, Harrington A, Ecsedy J, et al. (2002) Formin-2, polyploidy, hypofertility and positioning of the meiotic spindle in mouse oocytes. Nat Cell Biol 4: 921-928.

234. Salmassi A, Lu S, Hedderich J, Oettinghaus C, Jonat W, et al. (2001) Interaction of interleukin-6 on human granulosa cell steroid secretion. J Endocrinol 170: 471-478.

235. Webb RJ, Tinworth L, Thomas GM, Zaccolo M, Carroll J (2008) Developmentally acquired PKA localisation in mouse oocytes and embryos. Dev Biol 2008 May 1;317(1):36-45. doi: 10.1016/j.ydbio.2008.1001.1045.

236. Argov N, Moallem U, Sklan D (2004) Lipid transport in the developing bovine follicle: messenger RNA expression increases for selective uptake receptors and decreases for endocytosis receptors. Biol Reprod: 2004 Aug;2071(2002):2479-2085.

237. Deutsch GB, Zielonka EM, Coutandin D, Dotsch V (2011) Quality control in oocytes: domain-domain interactions regulate the activity of p63. Cell Cycle: 2011 Jun 2015;2010(2012):1884-2015.

238. Sala C, Arrigo G, Torri G, Martinazzi F, Riva P, et al. (1997) Eleven X chromosome breakpoints associated with premature ovarian failure (POF) map to a 15-Mb YAC contig spanning Xq21. Genomics 40: 123-131.

239. Carrel L, Willard HF (2005) X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature 434: 400-404.

240. Amelio I, Grespi F, Annicchiarico-Petruzzelli M, Melino G (2012) p63 the guardian of human reproduction. Cell Cycle 2012 Dec 15;11(24):4545-51. doi: 10.4161/cc.22819.

241. Tosca L, Crochet S, Ferre P, Foufelle F, Tesseraud S, et al. (2006) AMP-activated protein kinase activation modulates progesterone secretion in granulosa cells from hen preovulatory follicles. J Endocrinol 190: 85-97.

242. Geiman TM, Robertson KD (2002) Chromatin remodeling, histone modifications, and DNA methylation-how does it all fit together? J Cell Biochem 87: 117-125.

243. Murata M, Tamura A, Kodama H, Hirano H, Takahashi O, et al. (1998) Possible involvement of very low density lipoproteins in steroidogenesis in the human ovary. Mol Hum Reprod 4: 797-801.

244. Grummer RR, Carroll DJ (1988) A review of lipoprotein cholesterol metabolism: importance to ovarian function. J Anim Sci 66: 3160-3173.

245. Hu S, Liu H, Pan Z, Xia L, Dong X, et al. (2014) Molecular cloning, expression profile and transcriptional modulation of two splice variants of very low density lipoprotein receptor during ovarian follicle development in geese (Anser cygnoide). Anim Reprod Sci 2014 Oct;149(3-4):281-96. doi: 10.1016/j.anireprosci.2014.1006.1024.

246. Zhang G, Assadi AH, McNeil RS, Beffert U, Wynshaw-Boris A, et al. (2007) The Pafah1b complex interacts with the reelin receptor VLDLR. PLoS One 2.

247. Trommsdorff M, Gotthardt M, Hiesberger T, Shelton J, Stockinger W, et al. (1999) Reeler/Disabled-like disruption of neuronal migration in knockout mice lacking the VLDL receptor and ApoE receptor 2. Cell 97: 689-701.

248. Strasser V, Fasching D, Hauser C, Mayer H, Bock HH, et al. (2004) Receptor clustering is involved in Reelin signaling. Mol Cell Biol 24: 1378-1386.

249. May P, Herz J, Bock HH (2005) Molecular mechanisms of lipoprotein receptor signalling. Cell Mol Life Sci 62: 2325-2338.

250. Hiesberger T, Trommsdorff M, Howell BW, Goffinet A, Mumby MC, et al. (1999) Direct binding of Reelin to VLDL receptor and ApoE receptor 2 induces tyrosine phosphorylation of disabled-1 and modulates tau phosphorylation. Neuron 24: 481-489.

251. Arnaud L, Ballif BA, Forster E, Cooper JA (2003) Fyn tyrosine kinase is a critical regulator of disabled-1 during brain development. Curr Biol 13: 9-17.

252. Ikeda Y, Terashima T (1997) Expression of reelin, the gene responsible for the reeler mutation, in embryonic development and adulthood in the mouse. Dev Dyn 210: 157-172.

253. Fayad T, Lefebvre R, Nimpf J, Silversides DW, Lussier JG (2007) Low-density lipoprotein receptor-related protein 8 (LRP8) is upregulated in granulosa cells of bovine dominant follicle: molecular characterization and spatio-temporal expression studies. Biol Reprod: 2007 Mar;2076(2003):2466-2075.

254. Streiter S, Fisch B, Sabbah B, Ao A, Abir R (2015) The Importance of Neuronal Growth Factors in the Ovary. Mol Hum Reprod: 2015 Oct 2020.

255. Novoa I, Gallego J, Ferreira PG, Mendez R (2010) Mitotic cell-cycle progression is regulated by CPEB1 and CPEB4-dependent translational control. Nat Cell Biol 2010 May;12(5):447-56. doi: 10.1038/ncb2046.

256. Racki WJ, Richter JD (2006) CPEB controls oocyte growth and follicle development in the mouse. Development: 2006 Nov;2133(2022):4527-2037.

257. Kim H, Heo K, Kim JH, Kim K, Choi J, et al. (2009) Requirement of histone methyltransferase SMYD3 for estrogen receptor-mediated transcription. J Biol Chem 2009 Jul 24;284(30):19867-77. doi: 10.1074/jbc.M1109.021485.

258. Richards S, Aziz N, Bale S, Bick D, Das S, et al. (2015) Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015 May;17(5):405-24. doi: 10.1038/gim.2015.1030.

Website List:

DGV, Database of Genomic Variants, http://projects.tcag.ca/variation/

UCSC Genome Browser, http://genome.ucsc.edu/cgi-bin/hgGateway

Entrez gene NCBI, http://www.ncbi.nlm.nih.gov/gene

GeneCards, http://www.genecards.org/

DECIPHER,Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources, http://decipher.sanger.ac.uk/

ISCA, International Standards for Cytogenomics Arrat Consortium, http://www.iscaconsortium.org

OMIM, Online Mendelian Inheritance in Men, http://www.ncbi.nlm.nih.gov/omim

PubMed, http://www.ncbi.nlm.nih.gov/pubmed

OKdb, Ovarian Kaleidoscope Database, http://okdb.appliedbioinfo.net

STRING, http://string-db.org/

ToppGene, https://toppgene.cchmc.org/

geneMania, http://www.genemania.org/

Primer3, http://frodo.wi.mit.edu/primer3/

Optimase Protocol Writer,
http://www.mutationdiscovery.com/md/MD.com/screens/optimase/OptimaseInput.html

wANNOVAR, http://wannovar.usc.edu/index.php

1000 Genomes, http://www.1000genomes.org/

ExAC browser, Exome Aggregation Consortium, http://exac.broadinstitute.org/

EVS, Exome Variant Server, http://evs.gs.washington.edu/EVS/

dbSNP, http://www.ncbi.nlm.nih.gov/SNP/

Ensembl, http://www.ensembl.org/index.html

SIFT, http://sift.bii.a-star.edu.sg/

PolyPhen-2, http://genetics.bwh.harvard.edu/pph2/

MutationTaster, http://www.mutationtaster.org/

ClinVar, http://www.ncbi.nlm.nih.gov/clinvar/

HGMD, Human Gene Mutation Database, http://www.hgmd.cf.ac.uk/ac/index.php

Expasy translate tool, http://web.expasy.org/translate/

Concluso questo ulteriore traguardo desidero ringraziare:

- *Prof.ssa Palma Finelli per avermi dato la possibilità di svolgere il Dottorato di ricerca nel suo laboratorio, per aver sempre creduto in me e per avermi fatto appassionare al mondo della ricerca e della citogenetica molecolare. Grazie per tutti i preziosi consigli e insegnamenti.*
- *Prof.ssa Anna Marozzi, per avermi dato la possibilità di svolgere questo percorso, per aver raccolto parte delle pazienti analizzate, e per aver contribuito sperimentalmente insieme a Cecilia alla realizzazione di questo progetto;*
- *Prof.ssa Daniela Toniolo (Genetica delle malattie comuni, Istituto Scientifico San Raffaele) e le sue collaboratrici Caterina e Cinzia per aver contribuito alla raccolta delle pazienti oggetto di questo studio e per avermi aiutato nell'analisi dei dati WES;*
- *Dott. Davide Gentilini (Lab. di Biologia Molecolare, Istituto Auxologico Italiano) per il supporto informatico durante l'analisi WES;*
- *Prof. Luca Persani e le sue collaboratrici Raffaella e Ilaria (Lab. di Ricerche Endocrino-Metaboliche, Istituto Auxologico Italiano) per la raccolta e diagnosi delle pazienti;*
- *il gruppo diagnostica del Lab. di Citogenetica Medica e Genetica Molecolare (Istituto Auxologico Italiano) per la condivisione del lavoro giornaliero e per essere state sempre disponibili nei miei confronti;*
- *le mie colleghe di laboratorio Chiara, Milena, Alessandra e Maria, per avermi aiutata durante lo svolgimento di questo progetto di ricerca sia sperimentalmente che nella stesura della tesi. Grazie per tutti i consigli, gli insegnamenti che ognuna di voi ha saputo darmi durante questo periodo passato insieme; grazie inoltre per la complicità che si è creata, sia lavorativa che personale e per esserci state sempre.*

GRAZIE.