The genetics of premature ovarian failure: current perspectives

Chevy Chapman
Lynsey Cree
Andrew N Shelling
Department of Obstetrics and Gynecology, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand

Abstract: Premature ovarian failure (POF) is a common cause of infertility in women, characterized by amenorrhea, hypoestrogenism, and elevated gonadotropin levels in women under the age of 40. Many genes have been identified over the past few years that contribute to the development of POF. However, few genes have been identified that can explain a substantial proportion of cases of POF. The unbiased approaches of genome-wide association studies and next-generation sequencing technologies have identified several novel genes implicated in POF. As only a small proportion of genes influencing idiopathic POF have been identified thus far, it remains to be determined how many genes and molecular pathways may influence idiopathic POF development. However, owing to POF’s diverse etiology and genetic heterogeneity, we expect to see the contribution of several new and novel molecular pathways that will greatly enhance our understanding of the regulation of ovarian function. Future genetic studies in large cohorts of well-defined, unrelated, idiopathic POF patients will provide a great opportunity to identify the missing heritability of idiopathic POF. The identification of several causative genes may allow for early detection and would provide better opportunity for early intervention, and furthermore, the identification of specific gene defects will help direct potential targets for future treatment.

Keywords: premature ovarian failure, POF, next-generation sequencing, genome-wide association studies, genetics, ovarian aging

Introduction

The end of a woman’s reproductive lifespan is marked by the occurrence of menopause, defined as being the last menstruation that occurs for a woman, but is caused by the exhaustion of the ovarian reserve. In the general female population, across many ethnicities and over recent human history, the average age of natural menopause has remained at 50–52 years. However, aberrations in ovulatory processes, as outlined in Figure 1, may cause a pathogenic early depletion of ovarian follicles, which may result in premature menopause. Menopause before the age of 40 years is defined as premature ovarian failure (POF), also known as primary ovarian insufficiency. POF is described as the premature cessation of ovarian function and is characterized by 4–6 months of amenorrhea, a rise in serum follicle-stimulating hormone (FSH) levels to greater than 40 mIU/L, and hypoestrogenism. POF is a common disease, occurring in 1% of all women and 0.1% of women below the age of 30 years. As a consequence of hypoestrogenism, POF is associated with a greater risk of osteoporosis, osteoarthritis, and cardiovascular disease. Women suffering from POF experience similar symptoms to natural menopause; however, these symptoms are also accompanied by an earlier loss of fertility. Therefore, women at risk of POF who delay childbearing until after their 30s may experience problems conceiving and carrying a pregnancy to full term. This loss of fertility can be due to an accelerated loss of follicles, an
inability of the remaining follicles to respond to ovulatory signals, an initially reduced ovarian reserve at the time of birth, or a combination of all. As the development of POF occurs without any early discernible signs or symptoms, the diagnosis of POF currently occurs at the end stage of the disease, with a large proportion of these patients having not previously conceived before the time of diagnosis.

POF is a heterogeneous disease that can develop secondarily to many factors such as infection, metabolic disease, autoimmune disorders, or iatrogenic causes such as radiation, chemotherapy, or physical damage to the ovary. Additionally, some POF cases are syndromic, where the POF phenotype accompanies a specific disease, such as Turner’s syndrome or blepharophimosis–ptosis–epicanthus inversus syndrome (BPES) type 1. However, 50%–90% of cases are idiopathic and likely involve a substantial genetic contribution. A genetic basis of menopause is supported by heritability estimates of menopausal age between mother–daughter pairs ranging from 44% to 65% and recent reports of several genetic loci with significant associations with age at natural menopause identified through genome-wide association studies (GWAS). The importance of genetics in the etiology of POF is supported by the observation that in approximately 10%–30% of idiopathic cases, a first-degree relative is also affected, as well as the observation that a woman with an affected mother is six times more likely to develop POF.

Genetic analyses of POF patients have identified several chromosomal abnormalities, single gene mutations, and genetic polymorphisms, from multiple different biological pathways, associated with POF development. However, the genetic defects investigated thus far have been shown to contribute only a small percentage toward the development of POF. The diverse etiology of POF is in line with these findings and suggests that the pathogenesis of nonsyndromic POF is unlikely to be caused by a single gene or genetic defect, but rather supports the view of POF being a heterogeneous genetic disease involving the interaction of multiple genetic defects and environmental factors. Therefore, current research is focused on increasing the understanding of the genetic basis of idiopathic POF. This information will provide a greater insight into POF pathogenesis and allow for the identification of biomarkers, which may have the ability to predict early onset, providing women at risk with the opportunity to plan their families earlier.

In this review, we aim to provide an overview of the current knowledge of the genetic basis of POF development. To review current literature, we did a descriptive review of the literature based on our existing knowledge in the field and by
searching publication databases with the terms “premature ovarian failure”, “primary ovarian insufficiency”, “genetics”, and “mutation”. In addition, we will provide an update on the recent advances in identifying novel genetic factors associated with POF through means of GWAS and next-generation sequencing (NGS).

**X-chromosome defects**

X-chromosome abnormalities have long been recognized as the most common genetic cause of POF, estimated to account for as much as 12% of cases. These abnormalities usually involve either whole or partial deletions, duplications, or translocations of the X-chromosome. X monosomy (45X), known as Turner’s syndrome, has an incidence of 1:2,500. In phenotypic women, Turner’s syndrome is associated with short stature, gonadal dysgenesis, and primary amenorrhea. In women with Turner’s syndrome, oocyte loss usually begins early in childhood as a result of accelerated follicle atresia. Trisomy X (47XXX) is the most common form of aneuploidy, occurring in 1:1,000 females and may manifest in many mosaic forms, ie, 45X/47XXX, 46XX/47XXX, or 47XXX/48XXXX. Although most cases do not present with major medical problems, the 47XXX genotype has been associated with hypergonadotropic POF in some cases. Cytogenetic analysis in POF patients have identified a “critical region” on the long arm of the X-chromosome, necessary for ovarian development and function. Deletions are commonly localized at Xq21.3-q27 (POF1) in POF patients, whereas balanced X: autosomal translocations occur frequently at Xq13.3-q21.1 (POF2). Although no single gene has been shown to be consistently associated with POF within this critical region, the strong association of POF with X-chromosome abnormalities suggests that X-chromosome genes are critical for normal ovarian function, and any structural abnormality in this region may contribute toward POF development.

Aside from Turner’s syndrome, a premutation in the fragile-X mental retardation (FMR1) gene, located at Xq27.3, is the most commonly known congenital cause of POF. The 5’UTR region of the gene contains an expandable CGG repeat region, whereby <45 repeats is considered normal, while >200 repeats is considered a full mutation and is associated with fragile-X syndrome in male carriers. Repeats within the range of 55–199 are known as a premutation and are associated with an increased risk of POF in women. FMR1 premutations have been identified in 6% of sporadic POF cases and up to 13% in familial forms. The FMR1 protein is highly expressed in the fetal ovary germ cells, as well as in the granulosa cells of maturing follicles. There are two prevailing explanations for the association of the FMR1 premutation and POF. The FMR1 protein is a RNA-binding protein, which has a suppressive effect on the translation of a subset of mRNAs, the overexpression of which may impair the expression of oocyte development genes, resulting in a decrease in the size of the initial follicular pool. Alternatively, it has been proposed that an accumulation of mutated FMR1 mRNA may have a long-term toxic effect on the ovary and lead to excessive follicle atresia. Some recent studies have suggested that there is no association between CGG repeats of intermediate size and POF compared with controls; therefore, some of this data may need to be re-evaluated in light of recent findings.

**Candidate genes**

Given that the main phenotype of POF is infertility, identifying the causal gene in nonsyndromic POF families has proven difficult, as the majority of these cases have small or no family histories. Therefore, researchers have relied on candidate gene-driven studies in order to identify potentially causative genes involved in POF development. By screening genes known to play roles in folliculogenesis and ovarian function, as well as genes identified in the use of genetically modified mouse models, several POF candidate genes have been recognized. These genes are summarized in Table 1. However, the field has been limited by many underpowered studies, especially when only a single gene or a few gene variants have been studied at once, without providing adequate statistical analysis for the likelihood of the increased chance of identifying a false positive by not correcting for multiple testing. In addition, the identification of relevant candidate genes is restricted by our current understanding of disease, which in many cases is limited.

**Oocyte-specific transcription factors**

The initiation of primordial follicle formation and their subsequent differentiation into primary follicles is dependent on the expression of oocyte-specific genes, which are regulated by specific transcription factors. The screening of transcription factors regulating folliculogenesis in multiple POF populations has led to the identification of several causative mutations in the NR5A1, NOBOX, FIGLA, and FOXL2 genes. The nuclear receptor subfamily five group A member 1 (NR5A1), also known as steroidogenic factor 1, is a nuclear receptor located at 9q33.3 and is implicated in early gonadal differentiation. NR5A1 modulates steroidogenesis via the regulation of genes involved in the hypothalamic–pituitary–steroidogenic axis, such as STAR, CYP11A1, CYP17A1,
CYP19A1, LH, and INHA.\textsuperscript{28} The conditional knockout of NR5A1 in mouse granulosa cells causes infertility, as a result of hypoplastic ovaries and a reduced number of oocytes.\textsuperscript{29} To date, several NR5A1 mutations have been identified in four families with a history of 46XX POF, as well as in 2/25 idiopathic POF patients, which were absent in the 700 controls.\textsuperscript{27} All of these mutations resulted in impaired transactivational activity of gonadal promoters regulating follicle growth, thus leading to altered folliculogenesis and subsequent POF.\textsuperscript{30}

Both newborn ovary homeobox (NOBOX) and factor in germline alpha (FIGLA) encode oocyte-specific transcription factors, which regulate oocyte-specific genes necessary for the early stages of folliculogenesis.\textsuperscript{31} Through the use of knockout mouse models, NOBOX is thought to regulate the expression of a variety of oocyte-specific genes including BMP-15 and GDF-9.\textsuperscript{31} The deletion of NOBOX expression in these mice results in an accelerated loss of oocytes postnatally, representing a phenotype similar to nonsyndromic POF in humans.\textsuperscript{28} Two separate studies in Caucasian POF patients have identified several NOBOX mutations, which were absent in the control populations.\textsuperscript{26} Functional analysis of these mutations revealed a significant decrease in NOBOX DNA binding activity, resulting in the impaired regulation of genes critical for folliculogenesis.\textsuperscript{26} However, no NOBOX mutations were identified in studies of Chinese and Japanese POF patients, suggesting that NOBOX mutations may be rare in these populations.\textsuperscript{32,33} FIGLA regulates the expression of zona pellucida genes, the knockout of which in mice also mimics the phenotype of nonsyndromic POF, with an accelerated postnatal loss of primordial follicles.\textsuperscript{34} A study by Zhao et al\textsuperscript{35} in 2008 identified two novel FIGLA coding variants in a cohort of 100 Chinese POF women. These variants were absent in the 304 Chinese controls.\textsuperscript{35} The suggested mechanisms through which these variants led to POF are through haploinsufficiency (p.G6sfsX66), and a dominant negative effect, through altered heterodimerization (p.140delN).\textsuperscript{35}

The forkhead box L2 (FOXL2) gene, located at 3q23, is constitutively expressed primarily in granulosa cells in the
ovaries from embryogenesis through to adult life. Mutations in the FOXL2 gene are responsible for BPES type 1, of which female carriers also present with POF. FOXL2 is believed to play a role in granulosa cell differentiation, which is supported by in vitro studies localizing FOXL2 expression to undifferentiated granulosa cells, and through knockout mouse models that show the failure of granulosa cell differentiation, leading to the premature activation and depletion of primordial follicles. In addition, FOXL2 mutations have been identified in ~5% of nonsyndromic POF patients (without BPES), suggesting that FOXL2 mutations may also be associated with idiopathic POF.

More recent studies have highlighted the involvement of the spermatogenesis and oogenesis-specific basic helix-loop–helix (SOHLH) 1 and 2 transcription factors in POF development. SOHLH 1 and 2 are expressed exclusively in primordial follicles up until the primary follicle stage, and are master regulators of oocyte-specific genes critical for early follicle growth and differentiation, including NOBOX, FIGLA, BMP-15, and GDF-9. Both SOHLH 1 and 2 knockout female mouse models present with infertility and accelerated follicle loss, due to defects in primordial cell differentiation, with both models exhibiting atrophied ovaries devoid of oocytes by 6 weeks of age. A recent study by Qin et al identified four novel potentially deleterious heterogeneous SOHLH2 variants in a large cohort of Chinese and Serbian patients with POF. The study suggested that as two of these variants, —210G > T and 530+6T > G, resulted in reduced mature protein production, the primordial to primary follicle transition will be altered, leading to accelerated follicle loss and contributing to the POF phenotype in these patients. Additionally, a study conducted by Zhao et al identified three novel SOHLH1 variants potentially causative to POF in 364 Chinese that were absent in 400 controls. However, replication studies in larger populations, especially those in different ethnicities, are required to validate a causative association of SOHLH1 and 2 variants with POF.

**Folliculogenesis growth factors**

Bone morphogenetic protein 15 (BMP-15), located at Xq11.2, is a member of the transforming growth factor β (TGF-β) superfamily and encodes an oocyte-specific growth and differentiation factor. BMP-15 is involved in stimulating folliculogenesis and promoting follicle maturation by regulating granulosa cell differentiation and proliferation. The essential role of BMP-15 in folliculogenesis is illustrated by knockout mice models that present with subfertility due to defective ovulation processes. Many heterozygous variants in the BMP-15 gene prorogation, essential for proprotein dimerization, have been identified in several populations of POF patients with variable frequencies. These variants result in either a reduced production of mature BMP-15 protein or the production of variants with impaired activity. Although the mechanism through which BMP-15 variants influence POF development is presently unknown, it is proposed that BMP-15 haploinsufficiency may lead to defective granulosa cell signaling and impaired antiapoptotic mechanisms resulting in increased follicle atresia. However, an association of BMP-15 variants and POF has not been found in all studies. Additionally, the variants have also been identified in a small number of control participants, which may diminish their significance. Conversely, it may suggest that these are relatively common variants that contribute to an individual’s predisposition toward POF in combination with other predisposing factors.

The growth differentiation factor 9 (GDF-9) gene, a homologue of BMP-15, is also a member of the TGF-β superfamily and is located at 5q23.2. Like BMP-15, GDF-9 is oocyte specific and regulates primordial follicle development and stimulates granulosa cell proliferation and follicle maturation. GDF-9 is also thought to play a role in steroidogenesis and the modulation of FSH sensitivity in granulosa cells. GDF-9 null mice are completely infertile, suggesting that GDF-9 stimulation is essential for primordial follicles to progress. To date, several GDF-9 variants have been identified in POF cohorts of varying ethnicities, all of which are heterozygous and located in the prorogation. However, several other studies have found no association between GDF-9 variants and POF, suggesting GDF-9 variants may be a rare contributor to POF.

Inhibin alpha (INHA) encodes the alpha subunit of inhibin A and inhibin B, which together with their corresponding beta units (INHBA and INHBB), a class of dimeric glycoproteins, also belonging to the TGF-β superfamily. Inhibins are produced primarily by granulosa cells and act on the pituitary by inhibiting FSH production. Decreased inhibin production is associated with increased FSH production, leading to increased follicle recruitment and an increased depletion of the follicular pool. Indeed, women with idiopathic POF have lower serum levels of inhibins and higher serum FSH than that of age-matched fertile controls, suggesting that inhibins play an important role in normal ovarian function. Given the important biological role of inhibins in regulating folliculogenesis, INHA has long been a candidate gene for POF. Genetic studies on INHA variants have identified three polymorphisms associated with POF. One such polymorphism, a missense 769G > A variation, has...
shown a significant association in multiple POF populations across several ethnicities.\(^{13,54,55}\) Functional studies show this polymorphism results in a decrease in inhibin activity, suggesting that women carrying this variant have reduced FSH inhibition and an increased rate of follicular depletion, predisposing these women to POF.\(^{56}\) However, the association of \textit{INHA} variants with POF is controversial due to a lack of association found in a larger Italian and German cohort consisting of 611 POF patients and 1,084 controls, and in recent meta-analysis.\(^{56,57}\) As the 769G \(\rightarrow\) A variant is rare in some ethnicities, it is possible this variant may increase the risk of POF in some ethnicities more than others, thus contributing toward the discrepancies in the literature.\(^{49,52}\)

**Ovarian steroidogenesis**

The FSH and luteinizing hormone receptors (FSHR and LHR) and their corresponding ligands, FSH and LH, play a critical role during the gonadotropic-dependent stages of follicle growth and development. FSH and LH are secreted by the pituitary gland in response to the pulsatile stimulation of gonadotropin-releasing hormone from the hypothalamus.\(^{58}\) The binding of FSH to its receptor, located on the granulosa cells of small antral follicles, causes the upregulation of granulosa cell estrogen production, increased follicle growth, and the recruitment of follicles into a pool of growing preovulatory follicles.\(^{59}\) LHR activation is involved in oocyte maturation, the promotion of ovulation, and ovarian steroidogenesis.\(^{60}\) The hypothalamic–pituitary–gonadal loop is closed by the negative feedback of estradiol and inhibins on the pituitary.\(^{61}\) Because of its central role in follicle recruitment, FSHR has long been considered a candidate gene for POF.\(^{62}\) The first FSHR mutation, a homozygous C566T variant, was reported by Aittomäki et al\(^{62}\) which was found to be causative for primary amenorrhea in six Finnish POF families.\(^{62}\) However, the FSHR C566T variant is rare outside of the Finnish population.\(^{53,64}\) Mutations in the FSHR gene are associated with FSH resistance resulting in a raised basal serum FSH and estrogen deficiency, indicative of a reduced ovarian reserve. Women with complete FSH resistance have hypoplastic ovaries and primary amenorrhea, similar to FSHR knockout mice, which are sterile due to a lack of puberty and incomplete follicular development.\(^{65}\) Several other FSHR variants have been identified in POF women with secondary amenorrhea, including a novel heterozygous substitution found recently in a New Zealand woman with mild ovarian dysfunction and a family history of POF.\(^{66}\) However, FSHR variants are rare and have not been shown to have a direct role in POF development.\(^{6,11}\) LH resistance is characterized by oligomenorrhea or secondary amenorrhea due to the failure of ovulation.\(^{27,67}\) However, the occurrence of inactivating LHR variants in POF women is rare, probably because the degree of LH resistance must be severe in order to produce the POF phenotype.\(^{28}\)

Estrogen receptor 1 (\textit{ESR1}), located at 6q25.1, encodes ER-\(\alpha\) and is one of the two estrogen receptor subtypes; the other of which is ER-\(\beta\), encoded by \textit{ESR2}.\(^{68}\) The estrogen receptors are ligand-dependent transcription factors located on thecal and granulosa cells, as well as a variety of other tissues, which regulate the expression of a variety of genes involved in cellular growth and proliferation.\(^{69}\) Estrogenic signaling plays a critical role in follicular growth and maturation by maintaining FSH-induced follicular growth, regulating FSH activity, and inhibiting follicle apoptosis.\(^{70}\) Indeed, \textit{ESR1} knockout mice models are born fertile, but gradually lose their fertility due to the inability of follicles to progress to the antral stage, resulting in the presence of multiple large atretic follicles around 20–22 days of age.\(^{71}\) Several studies have investigated the relationship between the two most common ESR1 variants, PvuII (397T \(\rightarrow\) C) and XbaI (351A \(\rightarrow\) G), and POF development; however, the results are conflicting. Some studies reported an association between these variants and an increased risk of POF,\(^{70,72,73}\) while others have found either no association\(^{74}\) or that the XbaI variant in particular is associated with a reduced risk of POF.\(^{69,75}\) Additionally, the data on PvuII/XbaI haplotypes are also inconsistent, for instance, Cordts et al\(^{73}\) found the C/A haplotype to be associated with an increased risk of POF, whereas a study by Yang et al\(^{7}\) identified an association with a decrease in risk. In an unpublished New Zealand study conducted in a predominantly Caucasian cohort of 245 women, our team identified an association between the C/A haplotype and a reduced risk of low ovarian reserve, supporting a protective effect against POF development.

**Genome-wide association studies**

The genes mentioned previously were chosen as candidate genes for POF due to their known roles in folliculogenesis or ovarian function. Although the candidate gene approach has yielded some significant findings, it fails to identify novel genes from other pathways, which may be associated with POF development.\(^{76}\) A more recent strategy for identifying novel causative genes is to perform GWAS. In contrast to the candidate gene approach, which assesses one or a few genetic regions, GWAS studies employ an unbiased approach by investigating the entire genome, allowing for the identification of genetic variants without any assumptions as to their
underlying mechanisms or biological pathways.\textsuperscript{77} GWAS identify potential causal genes by analyzing the association of genetic markers or single nucleotide polymorphisms (SNPs), between large case-control groups of unrelated individuals.\textsuperscript{80} GWAS have identified novel gene–disease associations, opening up new areas of research, often where previously there had been limited understanding of the disease etiology.\textsuperscript{24} A significant amount of interest has developed around identifying the genes governing menopausal age, with several loci significantly associated with age at natural menopause identified through multiple large-scale GWAS studies.\textsuperscript{8–10} These studies identified 17 novel susceptibility loci, over several chromosomes, involved in hormonal regulation, immune function, and DNA repair, which account for 2.5%–4.1% of the variation in the age at natural menopause.\textsuperscript{8–10,76} It has always been assumed that genes that regulate the age at natural menopause would also be involved in the more pronounced POF phenotype. To date, only a few GWAS studies have been conducted in POF patients, and these have often been in only small populations. The results of these studies will need to be replicated in larger studies, and across different ethnicities, to confirm their importance and validity.

Kang et al\textsuperscript{81} reported the first two-stage POF GWAS study in 101 cases and 87 controls of Korean descent. This study identified a strong association between the parathyroid hormone-responsive B1 (\textit{PTHBI}) gene and POF. The physiological function of \textit{PTHBI} and its role in ovarian function remain unknown.\textsuperscript{80} However, \textit{PTHBI} variants have been identified in patients with Bardet–Biedl syndrome, a rare multisystemic genetic condition that can sometimes be associated with POF, suggesting that \textit{PTHBI} may be a novel candidate gene for POF.\textsuperscript{80} Another GWAS study by Knauff et al\textsuperscript{81} in 99 unrelated idiopathic POF patients and 235 unrelated controls of Caucasian descent identified a significant association between the \textit{ADAMTS19} gene and POF.\textsuperscript{81} \textit{ADAMTS19} encodes a zinc-dependent metalloprotease, which has been shown to be upregulated in female mice gonads during sexual differentiation, suggesting that \textit{ADAMTS19} is a biologically plausible POF candidate gene.\textsuperscript{81} However, this finding was not confirmed in an independent replication study of 60 POF patients and 90 controls,\textsuperscript{77} and so its role is currently unclear.

Following these initial POF GWAS studies, researchers have investigated the risk of early menopause (<45 years of age) and POF, for the 17 genetic variants known to be associated with the natural age of menopause.\textsuperscript{8–10} Genetic variants in four loci, 20q12.3 (\textit{MCM8}), 19q13.42 (\textit{BRSK1}/\textit{TMEM150B}), 5q35.2 (\textit{UMCI}/\textit{HK3}), and 6q24.2 (\textit{SYCP2L}), previously shown to be associated with variations in the age at natural menopause, were shown to increase the risk of both early menopause and POF.\textsuperscript{1,82,83} These findings suggest that normal menopause and POF are interrelated traits, which are influenced by common underlying genetic variants, with POF representing the extreme end of the normal distribution of the age at natural menopause.\textsuperscript{1} In support of this hypothesis, our team identified an association between a \textit{TMEM150B} variant and a \textit{BRSK1}/\textit{TMEM150B} haplotype, with an increased risk of low ovarian reserve in an unpublished study of 245 New Zealand women, hence predisposing carriers toward POF.

Little is known about the physiological functions of these newly identified variants, which is a common problem associated with GWAS, as genes are identified based on their statistical significance, and not their biological relevance. However, it is believed \textit{MCM8} is involved in the initiation of DNA replication and cellular proliferation, \textit{BRSK1} in oocyte polarity and growth, and \textit{TMEM150B} in damage-induced apoptosis.\textsuperscript{8,9} The overlap of the association of natural menopause variants with POF risk highlights a significant influence of DNA damage and repair pathways in the natural process of folliculogenesis. As an accumulation of DNA damage is a major contributor to the aging process, the association of these ubiquitously expressed DNA repair genes with earlier onset menopause aligns with the hypothesis that early menopause and POF are symptoms of an individual who suffers from an overall premature aging phenotype.\textsuperscript{10,22} However, the lack of functional characterization of these novel variants limits the ability to interpret GWAS findings, as they provide no biological insight into disease pathogenesis. Therefore, further studies, including replication in larger independent populations and the development of animal models, are required to provide functional support to the GWAS candidate gene associations and further improve our knowledge of the mechanisms regulating ovarian aging and possibly aging in general.\textsuperscript{52} Currently, the ReproGen Consortium is undertaking the largest GWAS study on early menopause and POF to date. The results of this project will help illuminate whether common genetic variants regulating normal menopausal age variation also influence the phenotypic extremes of the menopausal age distribution, as well as provide novel insights into the mechanisms involved in regulating ovarian aging.\textsuperscript{76}

While GWAS studies have identified some important and relevant genes, future studies will require larger datasets and meta-analysis of combined GWAS to provide even greater statistical power to identify smaller genetic associations. To date, for other diseases, GWAS studies have only been able...
to explain a small amount of the genetic variation related to that disease. This suggests that there are either more genes to be identified, other genetic mechanisms involved, or the GWAS technology is unable to identify all the relevant variants.\textsuperscript{24} It is unlikely that any really significant common SNPs will be present in POF; therefore, it is more likely that there will be more importance placed on individual rare SNPs that will require new sequencing technology to identify.

**Copy number variants**

In addition to candidate gene and SNP-based GWAS studies, genome-wide structural variation technologies, such as array comparative genomic hybridization, have been applied to complex diseases such as POF for the detection of genetic imbalances, such as deletions and duplications, also known as copy number variations (CNVs).\textsuperscript{84} CNVs are DNA regions $>$ 1 kb that vary in number between individuals and contribute toward phenotypic variation and disease susceptibility by altering transcriptional and translational levels, disrupting regulatory elements, or influencing gene dosage levels of adjacent or nearby genes.\textsuperscript{85} Aboura et al\textsuperscript{86} conducted the first study to assess the presence and prevalence of CNV in both sporadic and familial POF cases. The researchers identified five statistically significant potential candidate CNVs in genes involved in reproductive disease (DNAH5 and N4IP), reproductive endocrinology (DUSP22 and NUPRI), and folliculogenesis (AKTI).\textsuperscript{87} Another CNV study, conducted with higher resolution, identified a further 44 potentially causative microdeletions and microduplications in 74 POF patients.\textsuperscript{87} Interestingly, the rearrangements affected genes involved in meiosis (PLCB1, RB1CC1, and MAP4K4), DNA repair (RRBP8), and folliculogenesis (IMMP2L, FERIL6, and MEJ1),\textsuperscript{87} which are biological processes already implicated in POF pathogenesis, as identified in other POF genetic studies. McGuire et al\textsuperscript{88} conducted a SNP array-based study in 88 POF women and discovered 25 novel microdeletions and microduplications. The large majority of these CNVs were located on autosomes, suggesting that autosomal gene defects may have a larger contribution toward POF development than X-chromosome abnormalities, which coincides with the genetic heterogeneity and complex etiology of POF.\textsuperscript{77,88} Two of the genes associated with the autosomal microdeletions are involved in DNA double-strand break repair during meiosis (SYCE1) and mitotic cell progression (CPEB1). SYCE1 disruption leads to germ cell apoptosis in knockout mice, whereas CPEB1-deficient mice lack follicles as a result of oocyte loss.\textsuperscript{88} Therefore, these two genes are possible candidate POF genes, whereby reduced copy number/haploinsufficiency may lead to accelerated follicle loss and POF in humans.\textsuperscript{88} The consistent association of genetic defects in DNA repair and replication genes with POF suggests that this pathway significantly influences POF development and highlights a potential area to focus on for possible targeted treatments. Further studies in much larger POF cohorts are warranted to validate these findings and to identify novel CNVs. Animal studies should also be utilized to elucidate the role of these novel variants in ovarian development and aging; however, creating animal models of gene duplications is more technically challenging to perform than simple gene deletions.

**MicroRNAs**

The analysis of microRNAs (miRNAs) is another promising strategy recently applied to POF gene research. miRNAs are small noncoding single-stranded RNA molecules, approximately 22–24 nucleotides in size, that are involved in posttranslational gene regulation through repression or cleavage of specific RNA targets.\textsuperscript{89} miRNAs regulate cell differentiation, cell cycle progression, and apoptosis and, through the use of mouse models, have been shown to play a role in oocyte maturation and folliculogenesis.\textsuperscript{90} The first POF miRNA microarray analysis, by Zhaou et al demonstrated a differential miRNA expression profile in a very small cohort of three POF patients, compared to three normal women. One of the upregulated miRNAs, mir-23a, was demonstrated to promote granulosa cell apoptosis via the repression of X-linked inhibitor of apoptosis (XIAP) expression, suggesting the differential expression of mir-23a may be a potential contributor to POF development.\textsuperscript{91} However, because of the extremely small sample size in this study, the role of miRNAs in the etiology of POF is unclear. More recently, Dang et al\textsuperscript{92} investigated the differential expression of miRNAs in a large cohort of Chinese women, and identified 22 significantly upregulated and 29 significantly downregulated miRNAs in 140 POF patients, compared to 140 controls. Among them, mir-22-3p was significantly downregulated in POF patients, and a negative association between serum mir-22-3p and FSH levels was identified.\textsuperscript{92} The researchers suggest mir-22-3p may regulate pituitary FSH secretion, as its expression has been identified in the pig pituitary, whereby the decreased expression may contribute toward POF pathogenesis.\textsuperscript{93} These findings highlight how defects in the miRNA regulatory networks can impact on folliculogenesis. Further studies with larger sample sizes and different ethnicities are warranted to investigate the role of miRNAs in POF pathogenesis.
NGS and future directions

Over the last two decades, considerable efforts have been made to identify genetic causes and risk factors of idiopathic POF; however, only a limited number of genes have been validated as causative. The failure thus far to identify the genetic basis of idiopathic POF is due, in part, to the lack of familial cases large enough to allow causal gene detection through traditional genetic linkage analysis approaches. Additionally, ovarian development and function is influenced by several different molecular pathways, each regulated by many genes, making candidate gene selection difficult. The selection of relevant candidate genes is limited by our current understanding of POF, and it is likely many biologically relevant molecular pathways are underrepresented in our current genetic studies. Although the application of GWAS in POF research has accelerated the identification of novel candidate genes and genetic risk factors through unbiased genome analysis, it has not identified any genes that have a significant role in the development of POF. Sequencing-based approaches to gene identification have previously been limited due to its expense, limited throughput, and also due to technical limitations, as traditional Sanger sequencing can only analyze ~700 base pairs per reaction.\(^1\) Recently, significant advances in NGS technologies have provided a more efficient means of large-scale analysis of the entire human genome.\(^2\) NGS, such as whole-exome sequencing (WES), targeted NGS and whole-genome sequencing are high throughput DNA sequencing technologies that allow for the analysis of millions of base pairs in parallel across the entire exome or genome in a single reaction.\(^3\) Since the initial introduction of NGS in 2005, the cost of sequencing entire genomes has greatly reduced, resulting in NGS being used for both the discovery of novel disease-related genes and improving diagnostic genetic screening for both monogenic and polygenic diseases.\(^4\)

To date, four studies using WES have mainly been conducted in consanguineous families with inherited POF, of which pathogenic variants in the stromal antigen 3 (STAG3), HFM1, MCM8, and MCM9 genes were identified.\(^5-8\) Interestingly, these genes are implicated in DNA replication and repair, meiosis, and chromosome stability, further supporting the importance of these pathways in idiopathic POF pathogenesis. Although the exact mechanism of how defects in DNA repair pathways and genomic instability contribute toward POF development is not yet known, it is likely that an accumulation of DNA damage and chromosomal instability in the ovary would lead to accelerated follicle atresia, therefore predisposing women to POF.\(^9\) Indeed, these findings coincide with several knockout mouse models of meiotic and DNA repair genes, such as DMC1, MSH5, STAG3, and SYCE1, which present with accelerated follicle loss, due in part to increased double-stranded DNA breaks.\(^10\) Additionally, a more recent targeted NGS study of 70 candidate POF genes by Fonseca et al\(^9\) reported an association between mutations in the ADAMTS19 and BMP receptor 2 (BMPR2) genes with POF pathogenesis, in 12 unrelated idiopathic POF women. Both of these genes are plausible candidate genes for POF. ADAMTS19 has previously been reported as associated with POF,\(^11\) and defects in BMPR2 may disrupt the downstream signaling of its ligands, BMP-15 and GDF-9, therefore disrupting folliculogenesis and contributing to POF pathogenesis.\(^9\)

Most NGS studies conducted thus far have only been performed in small cohorts and are usually consanguineous POF families. This focus on consanguineous families has been partially to ensure a high likelihood for success, as the identified genes would be recessive, and therefore would be homozygously mutant in the affected individuals. As NGS technology gets cheaper, it is expected that there will be more studies in larger cohorts of unrelated idiopathic POF patients, which will identify novel rare gene variants and will identify further genetic regions contributing to POF development. Any new identified gene variants will produce a larger suite of candidate genes for more thorough analysis. Additionally, as the role of these newly identified variants play in POF pathogenesis is not known, future functional studies, both in in vitro human cell lines and in vivo mouse models, are necessary to elucidate the mechanisms through which these variants influence POF risk. Nevertheless, these preliminary findings suggest NGS technologies are powerful new tools, useful for identifying novel genetic regions and pathways involved in POF pathogenesis. The continued discovery and validation of POF-related genetic variants will allow for the development of new diagnostic and predictive tests, as well as the potential for novel targeted treatments.

Conclusion

As a result of recent technological advances, knowledge of the molecular basis and pathophysiology of idiopathic POF is rapidly growing. The unbiased approaches of GWAS and NGS technologies have highlighted defects in the folliculogenesis and DNA damage and repair pathways as significant contributors toward POF pathogenesis, providing a tangible focus for potential targeted treatments. However, it is unclear at present how we might design targeted therapies to defects affecting molecular mechanisms such as DNA damage.
and repair. Because of POF’s diverse etiology and genetic heterogeneity, we expect to see the contribution of many new and novel molecular pathways that will greatly enhance our understanding of the regulation of ovarian function. As only a small proportion of genes influencing idiopathic POF have been identified thus far, it remains to be determined how many genes and molecular pathways may influence idiopathic POF development. Future NGS studies in large cohorts of well-defined, unrelated, idiopathic POF patients will provide a great opportunity to identify the missing heritability of idiopathic POF.

Disclosure

The authors report no conflicts of interest in this work.

References


