Genetics in Medicine (2022) 24, 2274–2284

ELSEVIER

Genetics in Medicine
An Official Journal of the ACMG

www.journals.elsevier.com/genetics-in-medicine

ARTICLE

Heterozygous loss-of-function variants in *LHX8* cause female infertility characterized by oocyte maturation arrest

Lin Zhao[1,2], Qun Li[1,3], Yanping Kuang[4], Peng Xu[5], Xiaoxi Sun[6], Qingxia Meng[7], Wenjing Wang[1], Yang Zeng[1], Biaobang Chen[2], Jing Fu[6], Jie Dong[1], Jiawei Zhu[7], Yuxi Luo[1], Hao Gu[1], Caihong Li[8], Chunyi Li[8], Ling Wu[4], Xiaoyan Mao[4], Huizhen Fan[1], Ruyi Liu[1], Zhihua Zhang[1], Qiaoli Li[1], Jing Du[2], Lin He[9], Li Jin[10], Lei Wang[1,*], Qing Sang[1,*]

ARTICLE INFO

Article history:
Received 1 April 2022
Received in revised form
21 July 2022
Accepted 22 July 2022
Available online 27 August 2022

Keywords:
Female infertility
Gene-based burden test
LHX8
Loss-of-function variants
Oocyte maturation arrest

ABSTRACT

Purpose: The genetic causes of oocyte maturation arrest leading to female infertility are largely unknown, and no population-based genetic analysis has been applied in cohorts of patients with infertility. We aimed to identify novel pathogenic genes causing oocyte maturation arrest by using a gene-based burden test.

Methods: Through comparison of exome sequencing data from 716 females with infertility characterized by oocyte maturation arrest and 3539 controls, we performed a gene-based burden test and identified a novel pathogenic gene *LHX8*. Splicing event was evaluated using a minigene assay, expression of LHX8 protein was assessed in HeLa cells, and nuclear subcellular localization was determined in both HeLa cells and mouse oocytes.

Results: A total of 5 heterozygous loss-of-function *LHX8* variants were identified from 6 independent families (c.389+1G>T, c.412C>T [p.Arg138*], c.282C>A [p.Cys94*]; c.257dup [p.Tyr86*]; and c.180del, [p.Ser61Profs*30]). All the identified variants in *LHX8* produced truncated LHX8 protein and resulted in loss of LHX8 nuclear localization in both HeLa cells and mouse oocytes.

Conclusion: By combining genetic evidence and functional evaluations, we identified a novel pathogenic gene *LHX8* and established the causative relationship between LHX8 haploinsufficiency and female infertility characterized by oocyte maturation arrest.

Introduction

Infertility affects approximately 10% to 15% of reproductive-age couples worldwide and is estimated to affect approximately 186 million couples.[1,2] With the help of assisted reproductive technology, including in vitro fertilization (IVF) and intracytoplasmic sperm injection (ICSI), many couples with infertility can have their own babies.[3] However, a group of couples still suffer from recurrent failure of IVF/ICSI attempts because of oocyte

Lin Zhao, Qun Li, Yanping Kuang, Peng Xu, Xiaoxi Sun, Qingxia Meng, and Wenjing Wang contributed equally.

*Correspondence and requests for materials should be addressed to Lei Wang and Qing Sang, Fudan University, Institutes of Biomedical Sciences, Dong'an Road No 131, Shanghai, China 200032. *E-mail address:* wangleiwanglei@fudan.edu.cn; sangqing@fudan.edu.cn
Affiliations are at the end of the document.

doi: https://doi.org/10.1016/j.gim.2022.07.027

maturation arrest, oocyte death, fertilization failure, zona pellucida abnormality, zygotic cleavage failure, early embryonic arrest, etc. Oocyte maturation is a meiotic process in which germinal vesical (GV) oocytes undergo nuclear envelop breakdown and spindle assembly to reach metaphase I (MI) stage and then extrude the first polar body to form a mature egg.[4,5] Arrest at any of these stages will result in oocyte maturation defects and lead to female infertility. We previously showed that oocyte maturation arrest is a novel human Mendelian disease, and we identified pathogenic variants in *TUBB8* (OMIM 616768)[6] that are responsible for the disease.[6] With the wide and rapid application of exome sequencing, another 2 mutant genes have since been shown to be causative for oocyte maturation arrest, namely *PATL2* (OMIM 614661),[7] which is involved in oocyte GV arrest, and *TRIP13* (OMIM 604507),[8] which is involved in oocyte MI arrest. In addition, other factors associated with cytoplasmic maturation defects and morphologic defects were also identified.[9-13] These findings showed the importance of genetic contributions to oocyte maturation arrest. However, these mutant genes could only explain a limited number of affected individuals, and the underlying genetic factors in most cases remain unclear. Previous work suggested that the gene-based burden test is an effective approach to identify new genes in many disorders.[14-16] This test can overcome the lack of multiplex families and incomplete penetrance that often hinder conventional genetic analysis. To further explore the genetic etiology of the disease, we performed a gene-based burden test in a cohort of individuals with infertility with oocyte maturation arrest and mainly focused on rare and loss-of-function (LOF) variants that dramatically alter protein sequences.

LHX8 is a germ cell–specific transcription factor that is preferentially expressed in mammalian ovaries.[17-19] It controls the transcription of many genes that play critical roles in the maintenance and differentiation of oocytes during early oogenesis.[20] Previous studies of LHX8 have mainly been carried out in mice. Global knockout of *Lhx8* leads to female infertility due to loss of oocytes in newborn mice through disruption of oocyte-specific genes.[17] Conditional knockout of *Lhx8* in primordial follicles also resulted in female infertility in mice due to primary follicle death and decreased reserves in the secondary follicle pool.[20] Another study found that *Lhx8* ablation leads to large increases in autophagy in oocytes with DNA damage, leading to premature depletion of the ovarian reserve.[21] Despite the crucial role of *Lhx8* in mouse ovary development, the function of *LHX8* in human reproduction remains unclear, and *LHX8* has not been related to any human diseases.

In this study, we identified *LHX8* heterozygous LOF variants that were significantly enriched in a cohort of females with infertility with oocyte maturation arrest whereas no *LHX8* LOF variants were detected in the control group. In vitro studies showed that all the identified variants cause truncated LHX8 proteins and loss of nuclear localization, which indicates the severe impairment of LHX8 function. Our findings thus establish the causal relationship between *LHX8* variants and female infertility and expand our knowledge of the genetic etiology of human oocyte maturation arrest.

Materials and Methods

Clinical samples

We included 716 individuals affected by oocyte maturation arrest and 3539 control subjects. All individuals were recruited from 62 collaborating hospitals and reproductive centers in China from 2015. A complete medical evaluation was performed, including reproductive history, physical examination, laboratory blood work, and semen analysis. Inclusion criteria were (1) age younger than 45 years, (2) primary infertility without any pregnancy record, (3) patients had no other known causes of infertility, including male factors, chromosome anomalies, sexually transmitted infections, or surgical operations of the reproductive system, and (4) oocytes could be retrieved in IVF/ICSI attempt and more than half of the oocytes were defective, including morphologic abnormalities and arrested at the GV or MI stage in a single stimulation cycle. If only 1 cycle was performed and rare oocytes were retrieved, all the oocytes should be abnormal. The control subjects were fertile women. In this study, the affected subjects with *LHX8* variants were recruited from the Shanghai Ji'ai Genetics and IVF Institute, the Shanghai Ninth Hospital affiliated to Shanghai Jiao Tong University, the Affiliated Suzhou Hospital of Nanjing Medical University, and the Hainan Jinghua Hejing Hospital for Reproductive Medicine.

Exome sequencing

Genomic DNA was extracted from peripheral blood leukocytes using the ETP-300 Nucleic Acid Extractor (Enriching), library construction and capture were processed according to the kit (Twist Bioscience), and sequencing was performed on a Hiseq 3000 System (Illumina) platform. After sequencing, reads were converted to FASTQ format and mapped to reference genome GRCh37/hg19 using the Burrows-Wheeler Aligner. Variants were identified using the Genome Analysis Toolkit pipeline[22] and were checked using Mutalyzer, and all trios' samples passed kinship analysis using PLINK.[23] Variants were subsequently annotated using the ANNOVAR,[24] 1000 Genomes, single nucleotide polymorphism database, Haplotype Map, and Genome Aggregation Database (gnomAD). Qualifying variants were selected according to the following filtering criteria: (1) exclusion of low-complexity and duplication regions, (2) genotyping quality score was >40, (3) coverage

depth was at least 10× and mapping quality was >40, (4) alleles were LOF (defined as frameshift, stop-gain, or splicing variants), and (5) variants had low frequency (allele frequency < 0.0001) in gnomAD[25] and our in-house database (Supplemental Figure 1).

Gene-based burden test

Variants that disrupt protein-coding genes are thought to lead to abnormal gene function, and such variants have been identified in several diseases.[26-29] We conducted a burden test to quantify the enrichment of LOF variants in affected subjects compared with that in control subjects within every single gene. For each gene, the number of affected individuals and control subjects who carried at least 1 LOF variant were counted. After counting the numbers of cases and control subjects, a 2 × 2 contingency table was generated for each gene that contained the number of cases and control subjects carrying and not carrying the LOF variants. The *P* value was calculated using Fisher exact test in R, and the Bonferroni multiple-testing threshold was equal to 1.0×10^{-5} (0.05/4955 LOF genes). Subsequently, the quantile-quantile plot was visualized using the qqman package. Top significant genes ($P < .005$) were selected, and the candidate genes were evaluated according to the following criteria: (1) high gene expression level in oocytes and (2) functional relevance according to published references.

Plasmid construction and mutagenesis

All variants were referred to the same reference sequence (GenBank: NM_001001933). The full-length human *LHX8* coding sequence (GenBank: NM_001001933) was amplified from human oocyte complementary DNA and integrated into the pcDNA3.1 expression vector with an N-terminal FLAG-tag using a homologous recombination kit (Novoprotein). Point mutagenesis (c.389+1G>T, c. 412 C>T, c. 282 C>A, c.257dup, and c.180del) were created using the KOD-Plus Mutagenesis Kit (TOYOBO) according to the manufacturer's protocol. All clones were validated using Sanger sequencing. LHX8 plasmids were linearized using NotI, and the HiScribe T7 ARCA mRNA Kit (E2060S, New England Biolabs [UK] Ltd) was used to transcribe the wild-type and variant *LHX8* complementary RNAs according to the manufacturer's instructions.

Cell culture and transfection

HeLa cells were cultured in Dulbecco's Modified Eagle Medium supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin (Gibco) and maintained in a humidified incubator at 37 °C with 5% carbon dioxide. The empty plasmid with a FLAG-tag was used as the negative control. To evaluate the transfection efficiency of each variant, a GFP plasmid was cotransfected with the empty, wild-type, and variant plasmids of *LHX8* into HeLa cells using the PolyJet In Vitro DNA Transfection Reagent (SignaGen Laboratories) according to the manufacturer's instructions. After 36 h culture, the transfected cells were harvested for the following western blotting and cell immunostaining procedures. In total, 3 biological replicates were performed.

Mouse oocyte collection, microinjection, and cultivation

Mice were maintained and euthanized according to procedures approved by the Experimental Animal Ethics Committee of Fudan Medical College. Fully grown GV oocytes (diameter >60-80 μm with the GV located at the center) were collected from the ovaries of 7 to 8-week-old female Institute of Cancer Research mice (Beijing Vital River Laboratory Animal Technology Co Ltd) without any randomization or blinding. The empty plasmid with FLAG-tag was used as the negative control. The empty plasmid, wild-type, and variant *LHX8* complementary RNAs (1000 ng/μl) were injected into GV oocytes to express the corresponding protein. About 25 GV oocytes were used for each group, and the injected GV oocytes were cultured in M2 medium (Sigma-Aldrich) with milrinone (2.5 μM) (Sigma-Aldrich) for 12 h to express protein for the following oocyte immunostaining. In total, 3 biological replicates were performed.

Immunostaining

To evaluate the effects of the identified variants on LHX8 localization, immunostaining was performed in both HeLa cells and mouse oocytes. The transfected HeLa cells were fixed in 4% paraformaldehyde in phosphate-buffered saline (PBS) and blocked at room temperature for 1 h. For observing the subcellular localization of oocytes, the injected oocytes were fixed in 2% paraformaldehyde in PBS and permeabilized in PBS containing 0.5% Triton X-100 followed by incubation in blocking buffer (PBS, 1% bovine serum albumin, and 0.1% Tween-20) for 1 hour. For immunostaining in both transfected HeLa cells and injected mouse oocytes, anti-FLAG-Cy3 antibody (1:500 dilution; A9594, Sigma-Aldrich) was used for determining the LHX8-FLAG localization, and the DNA was labeled using Hoechst 33342 solution (1:700 dilution; 561908, BD). The HeLa cells and oocytes were observed, and images were captured on a confocal laser-scanning microscope (LSM880, Zeiss) using a 63× oil objective.

Western blotting

To evaluate the effects of the identified variants in *LHX8*, HeLa cells were harvested after being transfected with *LHX8* wild-type and variant vectors. Total protein was extracted using radioimmunoprecipitation assay lysis buffer (Shanghai Wei Ao Biological Technology) containing a

protease inhibitor cocktail (Bimake) and centrifuged (12,000*g* for 20 min) at 4 °C to yield the supernatant. Equal amounts of protein samples were electrophoresed on sodium dodecyl sulfate polyacrylamide gel electrophoresis gels and transferred to nitrocellulose filter membranes. Membranes were blocked using 5% nonfat milk for 1 hour and then, incubated with primary antibodies overnight at 4 °C. Mouse anti-FLAG antibody (1:3000 dilution; F7425, Sigma-Aldrich) was used to detect the protein expression level of LHX8-FLAG, and rabbit antivinculin antibody (OMIM 193065) (1:5000 dilution; 13901, CST) was used as the internal control. Goat antirabbit IgG (1:5000 dilution; M21002, Abmart) and goat antimouse IgG (1:5,000 dilution; M21001, Abmart) were used as the secondary antibodies to detect the primary antibodies. The blots were finally captured using Tanon ECL Western Blotting Substrate (ABclonal) after incubation with the secondary antibodies.

Minigene assay

The c.389+1G>T variant was located at the donor splice site of intron 4. Owing to the large size of the introns, we failed to clone the full sequence of exon 3, intron 3, exon 4, intron 4, and exon 5 into the minigene vector. Therefore, we integrated exon 3, exon 4, exon 5, and the 300 to 500 bp sequences before and after the 3 exons through polymerase chain reaction (PCR) from genomic DNA of proband (II-1) in family 1. The wild-type allele and c.389+1G>T variant allele were obtained through PCR because of the heterozygous status of the proband. The PCR products were cloned into the pcDNA3 vector, and the wild-type and c.389+1G>T variant plasmids were transfected into HeLa cells to determine the splicing effect of the variant. After incubation for 36 hours, total RNA was extracted using the RNeasy Mini Kit (Qiagen), and complementary DNA was obtained using PrimeScript RT Reagent Kit (Takara). The primers for constructing the minigene vectors and for detecting alternative splice sites are listed in the Supplemental Table 1.

Real-time quantitative PCR

Tissues were from aborted fetuses provided by the Obstetrics and Gynecology Hospital of Fudan University, and the donor's parents signed an informed consent form. Total RNA was extracted using the RNeasy Mini Kit. Genomic DNA was depleted, and reverse transcription was performed using the PrimeScript RT Reagent Kit with the gDNA Eraser (Takara). The expression level of *LHX8* was determined using specific primers (Supplemental Table 1) and was normalized to the expression level of an internal *GAPDH* (OMIM 138400) control. Real-time quantitative PCRs were performed on a QuantStudio 6 Flex Real-time PCR System (Applied Biosystems). In total, 3 biological replicates were performed.

Results

The gene-based burden test identified *LHX8* as a novel gene for oocyte maturation arrest

In general, LOF variants in key proteins can lead to severe impairments and are likely to be causative for many human diseases.[26] To explore the role of LOF variants in females with infertility with oocyte maturation arrest, we developed an analysis model of rare and highly deleterious LOF variants to identify candidate pathogenic genes in 716 case subjects in comparison to 3539 control subjects (Supplemental Figure 1). The analysis of LOF variants yielded 55 genes with a Fisher exact test *P* value of <.005 and minimal genomic inflation as shown in Supplemental Table 2. As a positive control, the well-studied gene *PATL2* showed significant enrichment in the case group, suggesting the feasibility of the sample size and the genetic analysis method. A stronger signal than that of *PATL2* was seen for the novel gene *LHX8*, which was observed in 7 LOF heterozygotes among the 716 cases and in none of the 3539 controls ($P = 3.73 \times 10^{-6}$) (Figure 1). The proband (II-1) and her sister (II-2) in family 1 carried the same splicing variant c.389+1G>T, whereas the individuals with infertility in family 2 and family 3 carried the recurrent variant c. 412C>T (p.Arg138*). The affected individuals in family 4, family 5, and family 6 carried the variants c.282C>A (p.Cys94*); c.257dup (p.Tyr86*); and c.180del, (p.Ser61Profs*30), respectively (Figure 2A). Taken together, we identified 5 LOF variants of *LHX8* in 7 affected individuals. All the variants were verified using Sanger sequencing and were not found in public databases, including the single nucleotide polymorphism database and gnomAD

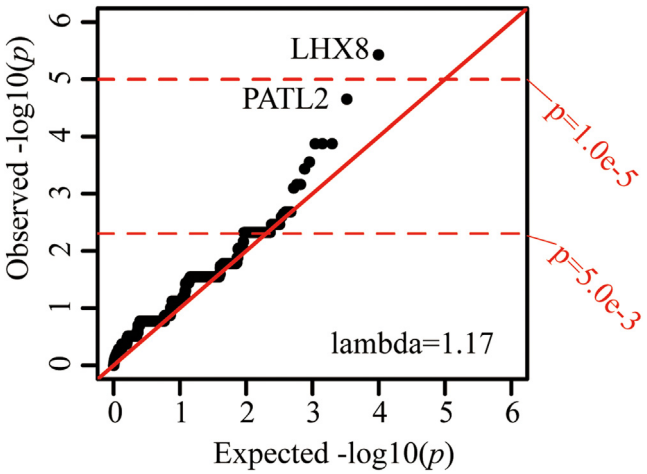

Figure 1 Top signals were identified using a gene-based burden test. Expected vs observed *P* values of rare loss-of-function (LOF) variants in the exome sequencing gene-based burden test in 716 female patients with infertility and 3539 controls. The signals for *LHX8* and *PATL2* are indicated. The significance of the difference between the observed and expected number of LOF variants was calculated using Fisher exact test.

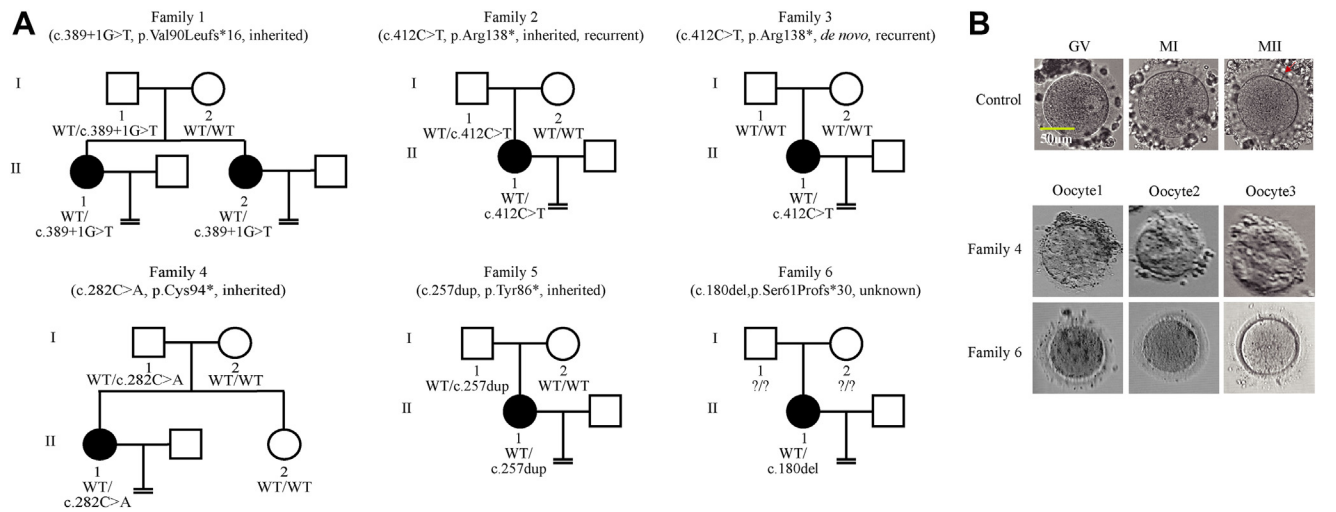

Figure 2 Identification of LOF variants in *LHX8*. A. Six pedigrees affected by oocyte maturation arrest. All 7 females with infertility carried LOF variants in *LHX8*. Squares denote male members, circles denote female members, solid symbols represent affected members, question marks indicate unavailable DNA samples, and the "=" sign indicates infertility. B. The retrieved oocytes from the proband in family 4 (II-1) had abnormal morphology compared with normal control oocytes (GV, MI, and MII). Owing to the incomplete clinical information, the exact stage of retrieved oocytes was unclear. The oocytes from the proband in family 6 (II-1) were arrested at the MI stage without polar body extrusion. The polar body of the control MII oocyte is indicated with a red arrow. The scale bar is 50 μm. GV, germinal vesicle; LOF, loss-of-function; MI, metaphase I; MII, metaphase II; WT, wild type.

(Supplemental Figure 2; Table 1). The variants in families 1, 2, 4, and 5 were inherited paternally in an autosomal dominant pattern whereas the variant in family 3 was de novo. The inheritance pattern in family 6 was unknown owing to the unavailability of DNA samples from her parents. Besides the LOF variants, we also examined the missense variants in our cohort through gene-based burden test. The detailed information of missense variants, including the dominant or recessive inheritance pattern, is listed in Supplemental Table 3. However, no other positive signal ($P < 1.0 \times 10^{-5}$) of novel genes were found for the missense variants.

Clinical characteristics of the individuals with infertility

The 7 individuals with infertility with *LHX8* variants shared similar phenotypes of oocyte maturation arrest according to their clinical information (Supplemental Table 4). The proband (II-1) in family 1 had undergone 1 IVF/ICSI attempt, and 2 oocytes were retrieved and were arrested at the GV and MI stage, respectively. Her sister (family 1, II-2) had undergone 2 failed IVF attempts and 1 failed ICSI attempt with unknown information for stages of arrested oocytes. The proband in family 2 was age 33 years at examination and was diagnosed with primary infertility for 5 years. She had undergone 2 failed IVF attempts and 1 failed ICSI attempt. In her ICSI attempt, 7 oocytes were retrieved, of which 4 oocytes were arrested at the GV stage, 2 oocytes were arrested at the MI stage, and only 1 oocyte was successfully fertilized. The proband in family 3 had undergone 2 ICSI attempts. In her first attempt, 5 oocytes were retrieved, of which 1 oocyte was arrested at the GV stage and 2 oocytes were arrested at the MI stage. The probands in family 4 and family 5 had a more severe oocyte maturation arrest phenotype, and >75% of their retrieved oocytes had abnormal morphology (Figure 2B, family 4). The proband in family 6 had been diagnosed with primary infertility for 5 years and had undergone 2 failed IVF attempts. In her 2 attempts, more than half of the retrieved oocytes were immature and showed no polar body extrusion (Figure 2B, family 6). In brief, the oocytes from the 7 affected individuals from 6 families had a similar phenotype of oocyte maturation arrest. The detailed clinical information and a full description of these individuals are provided in Supplemental Table 4.

LHX8 is evolutionarily conserved across species and is highly expressed in human oocytes

A multiple sequence alignment indicated that the amino acids of LHX8 are highly conserved among different mammalian species (Figure 3A), indicating an evolutionarily conserved function of LHX8. It has been reported that *Lhx8* is preferentially expressed in oocytes and germ cells within the mouse ovary.[33] To determine the spatial expression of human *LHX8*, we detected the messenger RNA (mRNA) of *LHX8* in human oocytes, early embryos, and other somatic tissues, such as the heart, liver, spleen, lung, kidney, brain, spine, and granulosa cells. As shown in Figure 3B, *LHX8* was highly expressed in human oocytes and weakly expressed in other somatic tissues, which implies its important role in human oocyte maturation.

Table 1 Overview of the *LHX8* variants identified in 6 families

Family	Position on Chr1 (bp)	cDNA Change	Protein Change	Variant Type	ACMG/AMP Classification (Evidence)	Inheritance	dbSNP eas	gnomAD eas	dbSNP European	gnomAD European
1(II-1)	75606792	c.389+1G>T	p.Val90Leufs*16	Splicing	Pathogenic (PVS1+PS3+PS4)	Inherited	NA	NA	NA	NA
1(II-2)	75606792	c.389+1G>T	p.Val90Leufs*16	Splicing	Pathogenic (PVS1+PS3+PS4)	Inherited	NA	NA	NA	NA
2	75608825	c. 412C>T	p.Arg138*	Stop-gain	Pathogenic (PVS1+PS3+PS4)	Inherited	NA	NA	NA	NA
3	75608825	c.412C>T	p.Arg138*	Stop-gain	Pathogenic (PVS1+PS2+PS3+PS4)	De novo	NA	NA	NA	NA
4	75606684	c.282C>A	p.Cys94*	Stop-gain	Pathogenic (PVS1+PS3+PS4)	Inherited	NA	NA	NA	NA
5	75602935	c.257dup	p.Tyr86*	Stop-gain	Pathogenic (PVS1+PS3+PS4)	Inherited	NA	NA	NA	NA
6	75602858	c.180del	p.Ser61Profs*30	Frameshift	Pathogenic (PVS1+PS3+PS4)	Unknown	NA	NA	NA	NA

Frequency of corresponding variants in the East Asian population of the dbSNP and gnomAD database.
ACMG/AMP, American College of Medical Genetics and Genomics/Association for Molecular Pathology; *bp*, base pair; *cDNA*, complementary DNA; *Chr*, chromosome; *dbSNP*, single nucleotide polymorphism database; *eas*, East Asian; *gnomAD*, Genome Aggregation Database; *NA*, not available.

LOF variants lacked key functional domains of LHX8

LHX8 consists of 1 homeobox domain, 2 LIM domains, 1 disordered region, and 1 nuclear localization signal (NLS) (Figure 3A). The LIM domains of the LHX family of transcription factors are protein–protein interaction domains that are necessary for binding with the cofactors needed for transcriptional activity,[34] and the NLS is a short stretch of amino acids that mediates the transport of proteins into the nucleus.[35] As shown in Figure 3A, 4 of the variants (p.Ser61Profs*30, p.Tyr86*, p.Val90Lfs*16, and p.Cys94*) were located before or within the first conserved LIM domain, which resulted in the complete loss of both LIM domains and the NLS. For the variant p.Arg138*, the first LIM domain was retained but the second LIM domain and the NLS were lost, disrupting the ability of LHX8 to bind cofactors for transcriptional function in the nucleus.

The LOF variants produced truncated LHX8 proteins

To confirm the effect of the c.389+1G>T variant on splicing events, minigene assay was performed using agarose gel electrophoresis. Compared with the wild-type vector, the c.389+1G>T variant resulted in a shifted and smaller band size, indicating abnormal splicing isoforms of this variant (Figure 4A). Sequencing analysis of the 2 bands also showed that the c.389+1G>T variant caused skipping of *LHX8* exon 4 (Figure 4B). This result indicates that the c.389+1G>T splicing variant led to an abnormal transcript and resulted in premature termination of the LHX8 protein. To further evaluate the effect of the identified variants on LHX8 protein, we examined the protein level of LHX8 in HeLa cells (Figure 4C). Compared with wild-type *LHX8*, the 3 nonsense variants (p.Tyr86*, p.Cys94*, and p.Arg138*) led to truncated proteins, and 2 of them (p.Cys94* and p.Arg138*) also caused decreased expression levels. The frameshift variant p.Ser61Profs*30 also caused a significant reduced expression level. The protein levels of the splicing variant p.Val90Leufs*16 was undetectable, which might be because of premature termination product.

The LOF variants resulted in the loss of LHX8 nuclear localization

To confirm the predicted loss of the NLS domain in the *LHX8* LOF variants, we determined the subcellular localization of wild-type and mutant LHX8 in HeLa cells. Similar to our prediction and western blotting results, the splicing variant c.389+1G>T, p.Val90Leufs*16 and the frameshift variant p.Ser61Profs*30 led to nearly undetectable cellular signal for LHX8. Although, the other 3 nonsense variants (p.Arg138*, p.Cys94*, and p.Tyr86*) showed visible signals, most of the signal was found in the cytoplasm, thus confirming loss of nuclear localization (Figure 4D). Considering the high expression of LHX8 specifically in human oocytes, we further confirmed the effects of LOF variants on subcellular

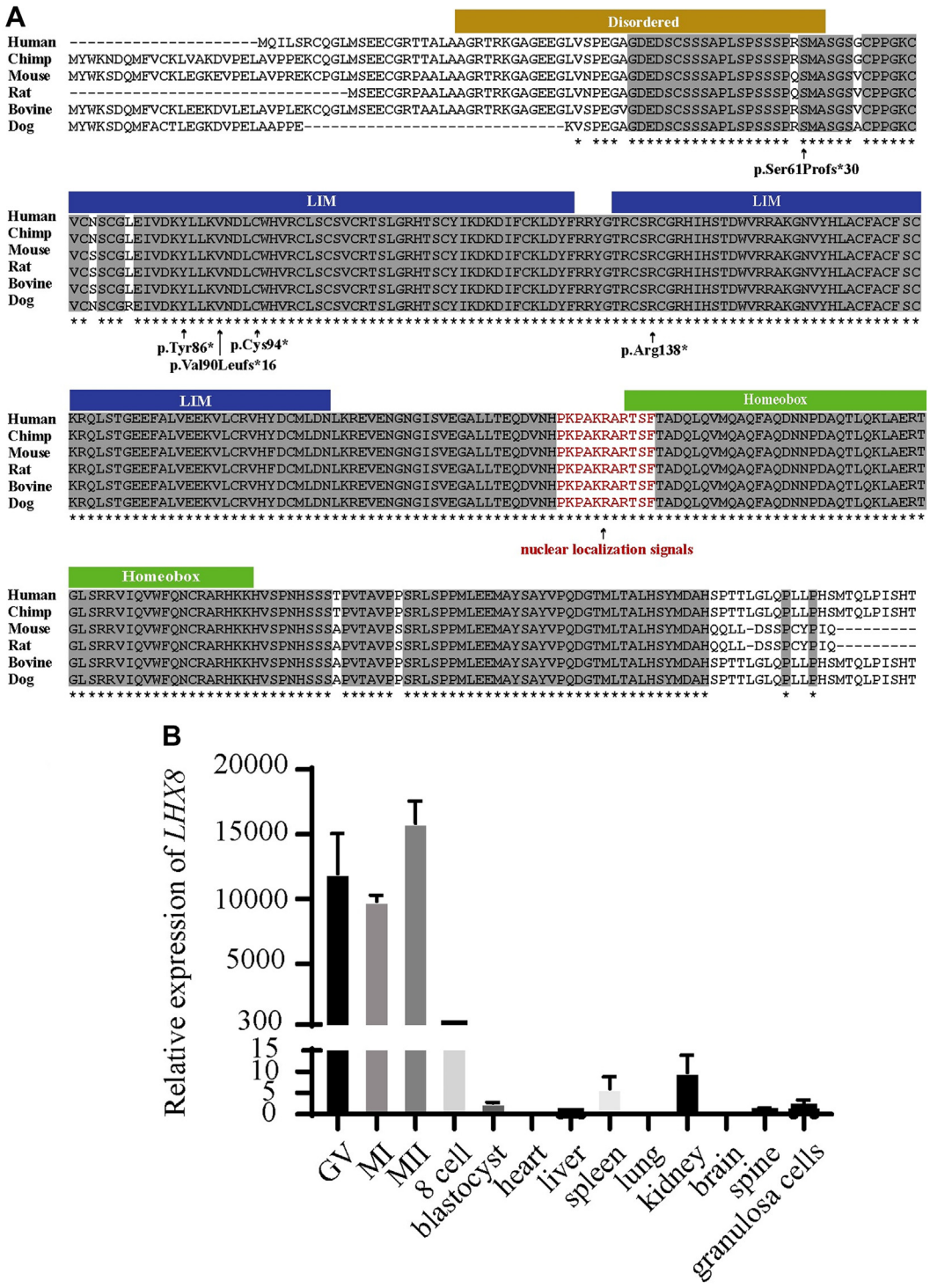

Figure 3 Conservation and distribution of LOF variants in *LHX8*. A. The conservation of the *LHX8* sequence is indicated among 6 mammalian species, and the locations of the LOF variants are indicated. Sequence alignment was performed using Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/),[30] and the functional regions and domains were determined by searching the UniProt (https://www.uniprot.org/) and Conserved Domain Database (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).[31] The disordered region is indicated by the yellow box. The LIM and homeobox domains are indicated by blue and green boxes, respectively. The nuclear localization signal was predicted using cNLS Mapper (https://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi)[32] and is indicated in red. The conserved amino acids are highlight in gray with asterisks below. B. The relative expression of *LHX8* messenger RNA (mRNA) in human GV, MI, and MII oocytes, 8-cell embryos, blastocysts, and heart, liver, spleen, lung, kidney, brain, spine, and granulosa cells as measured through quantitative reverse transcriptase polymerase chain reaction and normalized to *GAPDH* mRNA. The bars show the mean of 3 biological separate measurements, and error bars denote SDs. GV, germinal vesicle; LOF, loss-of-function; MI, metaphase I; MII, metaphase II.

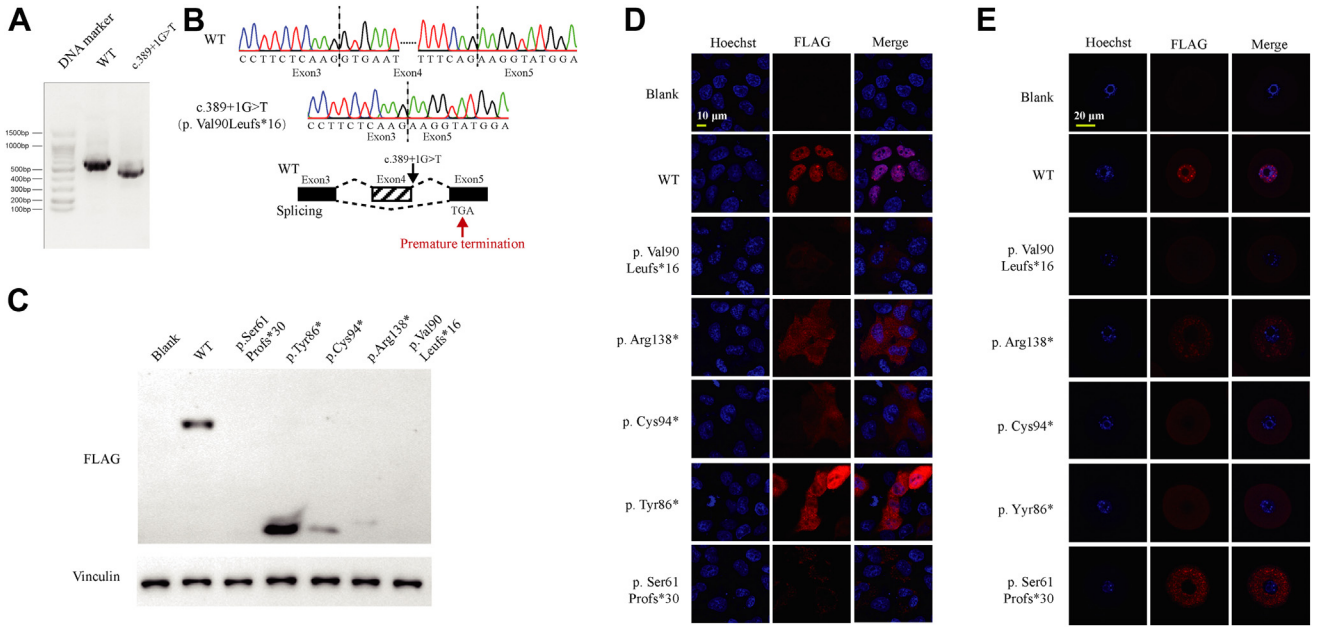

Figure 4 Effects of the LOF variants on nuclear localization in HeLa cells and mouse GV oocytes. A. Agarose gel electrophoresis showing the effect of the *LHX8* c.389+1G>T variant. Compared with the WT lane, the c.389+1G>T lane shows a band of smaller size. B. A scheme showing the effect of the *LHX8* c.389+1G>T variant. The *LHX8* c.389+1G>T splicing variant causes exon 4 to be skipped leading to the premature termination of the LHX8 protein. C. The effects of the LOF variants on LHX8-FLAG protein level as determined through western blotting in HeLa cells. D, E. Localization of WT and variant FLAG-tagged *LHX8* in human HeLa cells and in mouse GV oocytes. Images were captured using confocal microscopy (LSM880, Zeiss). Hoechst staining was used to label the DNA. The scale bars are 10 μm and 20 μm in HeLa cells and mouse oocytes, respectively. GV, germinal vesicle; LOF, loss-of-function; WT, wild type.

localization in mouse oocytes. Wild-type and variant *LHX8* were overexpressed in mouse oocytes. Similar results were obtained in mouse oocytes in which the wild-type LHX8 localized in the nucleus, whereas all the identified variants resulted in abnormal localization in the cytoplasm (Figure 4E). All these results provide strong evidence that these identified LOF variants result in severe functional impairment of the LHX8 protein.

Discussion

In this study, we identified 5 heterozygous LOF variants in *LHX8* in 7 female individuals with infertility from 6 independent families, and these patients shared similar phenotype of oocyte maturation arrest. We provide evidence that LOF variants resulted in abnormal splicing, truncated proteins, and abnormal cytoplasm localization in both HeLa cells and mouse oocytes. Taken together, these results suggest that all the identified variants were LOF alterations.

In mice, only homozygous deletion of *LHX8* causes female infertility, and heterozygous deletion in mice have normal fertility.[17] However, all the females with infertility with *LHX8* variants in this study were heterozygotes, which is inconsistent with observations in mice. This difference might be explained by the possibility that different thresholds of LHX8 protein are required for oocyte development in humans and mice, and it is possible that the baseline dosage to maintain normal function of LHX8 in humans is much higher than in mice. In addition, homozygous knockout female mice showed a severe phenotype of rapid oocyte loss and total depletion of ovarian reserves.[21] However, the individuals with infertility with *LHX8* LOF variants in this study showed a slightly different phenotype in which several oocytes could be obtained but most were immature. The phenotypic difference may be explained by the expression of half wild-type LHX8 protein in the patients' oocytes that is sufficient to maintain the survival of oocytes but cannot support normal meiotic maturation.

Genetically, the identified *LHX8* variants in the probands were inherited from their fathers in families 1, 2, 4, and 5, suggesting that the *LHX8* LOF variants had no effect on male fertility. A similar phenomenon was observed in *Lhx8* knockout mice in which *Lhx8* deficiency resulted in female infertility due to oocyte loss but had no effect on male fertility.[20] The phenotypic difference between males and females implies that LHX8 plays a dispensable role in male reproduction.

Previously, 3 mutant genes were reported to be responsible for human oocyte maturation arrest, namely *TUBB8*, *PATL2*, and *TRIP13*.[6-8] In this study, we showed that *LHX8* is a novel mutant gene that is also responsible for human oocyte maturation arrest. For *TUBB8*, most of the variants were missense pathogenic variants and exhibited dominant-negative effects, whereas homozygous or compound heterozygous variants in *PATL2* and *TRIP13* produced defective proteins. As for *LHX8*, the LOF variants resulted in complete functional loss in 1 allele, showing a different causative effect, referred to as haploinsufficiency,[36] compared with the

previously identified genes. This indicated that various kinds of pathogenic patterns should be taken into consideration when identifying new genes in future studies.

In this research, we identified LOF variants in *LHX8* by gene-based burden test and showed that these variants were pathogenic for female infertility, characterized by oocyte maturation arrest, suggesting a haploinsufficiency effect. For human genetic disease, there are some other genetic inheritance patterns, including digenic, oligogenic, and polygenic conditions. It is possible that other ovarian genes might be involved in combination with *LHX8*. Thus, additional genomic variants need to be further screened and studied with construction of the corresponding variants in mouse models for functional assays. Such effort would increase our understanding of the interplay between different genetic factors and reproductive health.

In conclusion, we have identified heterozygous LOF variants in *LHX8* that cause female infertility with oocyte maturation arrest via a gene-based burden test in a large cohort. In vitro experiments showed the functional impairment of the identified LOF variants. Our findings bring new insights into the understanding of human oocyte maturation and provide a novel diagnostic marker for individuals with infertility, characterized by oocyte maturation arrest.

Data Availability

Code and data are available on request. The main data relevant to the study are included in the article or uploaded as Supplemental Information.

Acknowledgments

This work was supported by the National Key Research and Development Program of China, China (2021YFC2700100), the National Natural Science Foundation of China, China (81725006, 32130029, 82171643, 81971450, 81971382, 82071642, 82001538), the Project of the Shanghai Municipal Science and Technology Commission, China (21XD1420300, 19JC1411001), the Natural Science Foundation of Shanghai, China (19ZR1444500, 21ZR1404800), the Capacity Building Planning Program for Shanghai Women and Children's Health Service, the collaborative innovation center project construction for Shanghai Women and Children's Health, the Guangdong Science and Technology Department, China, Guangdong-Hong Kong-Macao Joint Innovation Project (2020A0505140003), and Medical Engineering Fund of Fudan University, China (yg2021-012).

Author Information

L.Z., Qu.L., Y.K., P.X., X.S., and Q.M. contributed equally to this work. L.Z., Qu.L., Q.S., and Le.W. conceived and designed this study. Y.K., P.X., X.S., Q.M., and J.F. collected the clinical samples. J.Z., Ca.L., Ch.L., Li.W., and X.M. collected the clinical records. B.C., Jie.D., Y.L., H.G., and H.F. organized the medical records. Qu.L. analyzed the exome sequencing data and constructed the plasmids. L.Z. performed the in vitro experiments. Y.Z. performed the microinjection. R.L., Z.Z., W.W., Qi.L., and Jin.D. analyzed the experimental data. L.Z., Qu.L., Le.W., and Q.S. wrote the manuscript. L.H. and L.J. provided important guidance for the study. All authors reviewed and approved the manuscript.

Ethics Declaration

Our study was approved by the Ethics Committee of the Biomedical Research Institute of Fudan University. Written informed consent was obtained from all human participants. Mice were maintained and euthanized according to the procedures approved by the Experimental Animal Ethics Committee of Fudan Medical College.

Conflict of Interest

The authors declare no conflict of interest.

Additional Information

The online version of this article (https://doi.org/10.1016/j.gim.2022.07.027) contains supplementary material, which is available to authorized users.

Affiliations

[1]Institute of Pediatrics, Children's Hospital of Fudan University, the Institutes of Biomedical Sciences, the State Key Laboratory of Genetic Engineering, Fudan University, Shanghai, China; [2]NHC Key Lab of Reproduction Regulation, Shanghai Institute for Biomedical and Pharmaceutical Technologies, Fudan University, Shanghai, China; [3]Human Phenome Institute, Fudan University, Shanghai, China; [4]Reproductive Medicine Center, Shanghai Ninth Hospital, Shanghai Jiao Tong University, Shanghai, China; [5]Hainan Jinghua Hejing Hospital for Reproductive Medicine, Haikou, China; [6]Shanghai Ji Ai Genetics and IVF Institute, Obstetrics and Gynecology Hospital, Fudan University, Shanghai, China; [7]Center for Reproduction and Genetics, The Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou, China; [8]Shenyang Jinghua Hospital, Liaoning, China; [9]Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders, Ministry of Education, Shanghai Jiao Tong University, Shanghai, China; [10]State Key Laboratory of Genetic

Engineering and Collaborative Innovation Center for Genetics and Development, School of Life Sciences, Fudan University, Shanghai, China

Web Resources

EMBL-EBI.1000 Genomes. Accessed on December 24, 2020. http://www.internationalgenome.org/

Broad Institute of MIT and Harvard. gnomAD. Accessed on December 24, 2020. https://gnomad.broadinstitute.org/

Wang K, Li M, Hakonarson H. ANNOVAR. Accessed on December 24, 2020. https://annovar.openbioinformatics.org/

National Center for Biotechnology Information. dbSNP. Accessed on December 24, 2020. https://www.ncbi.nlm.nih.gov/projects/SNP/

Leids Universitair Medisch Centrum. Mutalyzer. Accessed on June 01, 2021. https://www.mutalyzer.nl/

National Center for Biotechnology Information. GenBank. Accessed on September 10, 2020. https://www.ncbi.nlm.nih.gov/genbank/

Johns Hopkins University. OMIM. Accessed on June 01, 2021. http://www.omim.org/

Adzhubei IA, Schmidt S, Peshkin L, et al. PolyPhen-2. Accessed on December 24, 2020. http://genetics.bwh.harvard.edu/pph2/

National Center for Biotechnology Information. RefSeq. Accessed on September 10, 2020. http://www.ncbi.nlm.nih.gov/RefSeq

Paulin Ng. SIFT. Accessed on December 24, 2020. http://sift.bii.a-star.edu.sg/

References

1. Inhorn MC, Patrizio P. Infertility around the globe: new thinking on gender, reproductive technologies and global movements in the 21st century. *Hum Reprod Update*. 2015;21(4):411-426. http://doi.org/10.1093/humupd/dmv016
2. Tamrakar SR, Bastakoti R. Determinants of infertility in couples. *J Nepal Health Res Counc*. 2019;17(1):85-89. http://doi.org/10.33314/jnhrc.1827
3. DeAngelis AM, Martini AE, Owen CM. Assisted reproductive technology and epigenetics. *Semin Reprod Med*. 2018;36(3-4):221-232. http://doi.org/10.1055/s-0038-1675780
4. Fan H-Y, Sun Q-Y. Chapter 12 - Oocyte meiotic maturation. In: Leung PCK, Adashi EY, eds. *The Ovary*. 3rd ed. Academic Press; 2019:181-203.
5. Mehlmann LM. Stops and starts in mammalian oocytes: recent advances in understanding the regulation of meiotic arrest and oocyte maturation. *Reproduction*. 2005;130(6):791-799. http://doi.org/10.1530/rep.1.00793
6. Feng R, Sang Q, Kuang Y, et al. Mutations in TUBB8 and human oocyte meiotic arrest. *N Engl J Med*. 2016;374(3):223-232. http://doi.org/10.1056/NEJMoa1510791
7. Chen B, Zhang Z, Sun X, et al. Biallelic mutations in PATL2 cause female infertility characterized by oocyte maturation arrest. *Am J Hum Genet*. 2017;101(4):609-615. http://doi.org/10.1016/j.ajhg.2017.08.018
8. Zhang Z, Li B, Fu J, et al. Bi-allelic missense pathogenic variants in TRIP13 cause female infertility characterized by oocyte maturation arrest. *Am J Hum Genet*. 2020;107(1):15-23. http://doi.org/10.1016/j.ajhg.2020.05.001
9. Sang Q, Li B, Kuang Y, et al. Homozygous mutations in WEE2 cause fertilization failure and female infertility. *Am J Hum Genet*. 2018;102(4):649-657. http://doi.org/10.1016/j.ajhg.2018.02.015
10. Huang HL, Lv C, Zhao YC, et al. Mutant ZP1 in familial infertility. *N Engl J Med*. 2014;370(13):1220-1226. http://doi.org/10.1056/NEJMoa1308851
11. Li Y, Liang Z, Liang Z, Yang X, Xia H, Yu H. Abnormal PIWI-interacting RNA profile and its association with the deformed extracellular matrix of oocytes from recurrent oocyte maturation arrest patients. *Fertil Steril*. 2021;115(5):1318-1326. http://doi.org/10.1016/j.fertnstert.2020.11.037
12. Zhao L, Xue S, Yao Z, et al. Biallelic mutations in CDC20 cause female infertility characterized by abnormalities in oocyte maturation and early embryonic development. *Protein Cell*. 2020;11(12):921-927. http://doi.org/10.1007/s13238-020-00756-0
13. Zheng W, Zhou Z, Sha Q, et al. Homozygous mutations in BTG4 cause zygotic cleavage failure and female infertility. *Am J Hum Genet*. 2020;107(1):24-33. http://doi.org/10.1016/j.ajhg.2020.05.010
14. Guo MH, Plummer L, Chan YM, Hirschhorn JN, Lippincott MF. Burden testing of rare variants identified through exome sequencing via publicly available control data. *Am J Hum Genet*. 2018;103(4):522-534. http://doi.org/10.1016/j.ajhg.2018.08.016
15. Povysil G, Chazara O, Carss KJ, et al. Assessing the role of rare genetic variation in patients with heart failure. *JAMA Cardiol*. 2021;6(4):379-386. http://doi.org/10.1001/jamacardio.2020.6500
16. Bestetti I, Barbieri C, Sironi A, et al. Targeted whole exome sequencing and Drosophila modelling to unveil the molecular basis of primary ovarian insufficiency. *Hum Reprod*. 2021;36(11):2975-2991. http://doi.org/10.1093/humrep/deab192
17. Choi Y, Ballow DJ, Xin Y, Rajkovic A. Lim homeobox gene, lhx8, is essential for mouse oocyte differentiation and survival. *Biol Reprod*. 2008;79(3):442-449. http://doi.org/10.1095/biolreprod.108.069393
18. Wan Q, Wang Y, Wang H. Identification and analysis of regulatory elements in porcine bone morphogenetic protein 15 gene promoter. *Int J Mol Sci*. 2015;16(10):25759-25772. http://doi.org/10.3390/ijms161025759
19. White YA, Woods DC, Takai Y, Ishihara O, Seki H, Tilly JL. Oocyte formation by mitotically active germ cells purified from ovaries of reproductive-age women. *Nat Med*. 2012;18(3):413-421. http://doi.org/10.1038/nm.2669
20. Ren Y, Suzuki H, Jagarlamudi K, et al. Lhx8 regulates primordial follicle activation and postnatal folliculogenesis. *BMC Biol*. 2015;13:39. http://doi.org/10.1186/s12915-015-0151-3
21. D'Ignazio L, Michel M, Beyer M, et al. Lhx8 ablation leads to massive autophagy of mouse oocytes associated with DNA damage. *Biol Reprod*. 2018;98(4):532-542. http://doi.org/10.1093/biolre/iox184
22. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297-1303. http://doi.org/10.1101/gr.107524.110
23. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559-575. http://doi.org/10.1086/519795
24. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):e164. http://doi.org/10.1093/nar/gkq603
25. Gudmundsson S, Singer-Berk M, Watts NA, et al. Variant interpretation using population databases: lessons from gnomAD. *Hum Mutat*. 2022;43(8):1012-1030. http://doi.org/10.1002/humu.24309
26. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536(7616):285-291. http://doi.org/10.1038/nature19057
27. Prudente S, Jungtrakoon P, Marucci A, et al. Loss-of-function mutations in APPL1 in familial diabetes mellitus. *Am J Hum Genet*. 2015;97(1):177-185. http://doi.org/10.1016/j.ajhg.2015.05.011

28. Hengel H, Hannan SB, Dyack S, et al. Bi-allelic loss-of-function variants in BCAS3 cause a syndromic neurodevelopmental disorder. *Am J Hum Genet*. 2021;108(6):1069-1082. http://doi.org/10.1016/j.ajhg.2021.04.024
29. Jiao SY, Yang YH, Chen SR. Molecular genetics of infertility: loss-of-function mutations in humans and corresponding knockout/mutated mice. *Hum Reprod Update*. 2021;27(1):154-189. http://doi.org/10.1093/humupd/dmaa034
30. Madeira F, Park YM, Lee J, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res*. 2019;47(W1):W636-W641. http://doi.org/10.1093/nar/gkz268
31. Lu S, Wang J, Chitsaz F, et al. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res*. 2020;48(D1):D265-D268. http://doi.org/10.1093/nar/gkz991
32. Kosugi S, Hasebe M, Tomita M, Yanagawa H. Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc Natl Acad Sci U S A*. 2009;106(25):10171-10176. Published correction appears in *Proc Natl Acad Sci U S A*. 2009;106(31):13142. https://doi.org/10.1073/pnas.0900604106
33. Pangas SA, Choi Y, Ballow DJ, et al. Oogenesis requires germ cell-specific transcriptional regulators Sohlh1 and Lhx8. *Proc Natl Acad Sci U S A*. 2006;103(21):8090-8095. http://doi.org/10.1073/pnas.0601083103
34. Dawid IB, Breen JJ, Toyama R. LIM domains: multiple roles as adapters and functional modifiers in protein interactions. *Trends Genet*. 1998;14(4):156-162. http://doi.org/10.1016/s0168-9525(98)01424-3
35. Cokol M, Nair R, Rost B. Finding nuclear localization signals. *EMBO Rep*. 2000;1(5):411-415. http://doi.org/10.1093/embo-reports/kvd092
36. Cleary MA. Haploinsufficiency. In: Brenner S, Miller JH, eds. *Brenner's Encyclopedia of Genetics*. Academic Press; 2001:911.