

# Genetic testing of sperm donors at a human sperm bank in China

**Chuan Huang**

Reproductive and Genetic Hospital of CITIC-Xiangya

**Hui-Lan Wu**

Reproductive and Genetic Hospital of CITIC-Xiangya

**Wen-Jun Zhou**

Basic Medicine College, Central South University

**Zeng-Hui Huang**

Reproductive and Genetic Hospital of CITIC-Xiangya

**Xue-Feng Luo**

Basic Medicine College, Central South University

**Yu-Ling Tang**

Basic Medicine College, Central South University

**Qian Liu**

Basic Medicine College, Central South University

**Li-Qing Fan**

Reproductive and Genetic Hospital of CITIC-Xiangya

**Hong-Chuan Nie**

Reproductive and Genetic Hospital of CITIC-Xiangya

**Wen-Bing Zhu (✉ [zhuwenbing0971@sina.com](mailto:zhuwenbing0971@sina.com))**

Reproductive and Genetic Hospital of CITIC-Xiangya

---

**Research Article**

**Keywords:** Sperm bank, sperm donor, genetic testing, whole exom sequencing, specific gene test

**Posted Date:** May 13th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1639899/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

## Abstract

## Background

In China, numerous human sperm banks only perform three-generation family history analysis to exclude genetic diseases with clinical symptoms; therefore, many inherited risks cannot be detected before donor qualification even when a thorough genetic family history evaluation has been performed. Hence, the risk of recessive disease inheritance persists with the current eligibility guidelines in China regarding the donor selection process.

## Materials and Methods

Retrospective study that reviewed the genetic test analyses and clinical outcomes of young adult men who were qualified sperm donors at the Hunan Province Human Sperm Bank of China from January 1, 2018, to May 1, 2021. We included a total of 3231 qualified sperm donors: all donors underwent primary screening for thalassemia and glucose-6-phosphate dehydrogenase (G6PD) deficiency. Whereafter, 278 of donors underwent genetic testing for specific genes, and 43 donors underwent whole exome sequencing.

## Results

2.4% of 3231 qualified sperm donors might have thalassemia and 1.4% might have G6PD deficiency. Sperm donors with thalassemia and G6PD deficiency would be eliminated. Specific gene testing identified 7 of the 278 donors (2.5%) as carriers of at least one pathogenic or likely pathogenic variant in a gene, including 1.9% of 154 donors (3/154) as carrier variants in  $\alpha$ -Like or  $\beta$ -Like globin genes, 17.6% of 17 donors (3/17) as carrier variants in GJB2, 12.5% of 8 donors (1/8) as carrier variants in SMN1. In addition, among the 43 sperm donors carrying the 111 pathogenic/likely pathogenic variants, eight (18.6%) were carriers of pathogenic variants of the GJB2 gene. The frequency, therefore, was approximately 1 in 5.

## Conclusions

The data suggest that used blood routine and RDT can make a preliminary screening of sperm donors, and special gene testing should be performed for sperm donors according to the regional incidence of specific genetic diseases. Meanwhile, whole exome sequencing can be used as a supplementary application in sperm donor genetic testing, and aid a successful and healthy pregnancy. However, industry guidelines must be modified to incorporate its use.

## Background

Like all prospective parents, people who rely on donor sperm want to conceive healthy children. However, a retrospective study of birth defects and the medical risks of donor-conceived offspring found that approximately one-third of the birth defects in such offspring involved autosomal recessive (AR) disorders [1]. In fact, some autosomal dominant (AD) disorders cannot always be detected at a donor's eligibility assessment owing to reduced penetrance, variable expressivity, or temporal factors. Today, China has 27 sperm banks, the largest of which screens 4000 potential donors each year [2]. Sperm donor screening is conducted strictly in accordance with standard guidelines published by the Chinese Ministry of Health in 2003 [3]. One of the guidelines states that potential donors must undergo laboratory testing to exclude genetic diseases, and undergo karyotype analysis. However, numerous human sperm banks in China only perform three-generation family history analysis to exclude genetic diseases with clinical symptoms. Therefore, many inherited risks cannot be detected before donor qualification even when a thorough genetic family history evaluation has been performed. Hence, not only is there a risk for autosomal recessive disease, but also risk for autosomal dominant variants with reduced penetrance, variable expressivity, or other complex modes of inheritance with the current eligibility guidelines in China for the donor selection process.

The American Society for Reproductive Medicine (ASRM) has recommended that sperm donors undergo appropriate genetic evaluation, e.g., genetic carrier screening [4]. Fifteen of the 17 sperm banks in the United States reported that they would attempt to accommodate special requests from donor semen recipients to perform a particular genetic test on a specific anonymous sperm donor [5]. In recent years, the cost of molecular genetic studies has been greatly reduced due to the emergence of next-generation sequencing (NGS) technologies, allowing the analysis of numerous genes and the concurrent evaluation of hundreds of variants [6]. Some reproductive centers use NGS to ensure that donors and recipients do not have mutations in the same genes before assisted reproductive therapy (ART) is initiated, thereby avoiding high reproductive risk for transmitting a severe AR genetic condition to the offspring [7]. Therefore, it is necessary to perform genetic testing of sperm before donor qualification.

In the present study, we performed genetic testing on donors to a human sperm bank in China. We used whole exome sequencing to analyze variants to determine if they are potentially pathogenic in the donor and recipient and in preassigned donor–recipient matches. We provide evidence that donors have a significant risk of passing on an inherited susceptibility for AR and undiagnosed AD disorders, and we recommend performing more genetic testing at human sperm banks.

## Materials And Methods

### Study population and participants

This retrospective study reviewed the genetic test analyses and clinical outcomes of young adult men who were qualified sperm donors at the Hunan Province Human Sperm Bank of China from January 1, 2018, to May 1, 2021. Their clinical information was collected and analyzed, and included the clinical outcomes after semen use. All donors sign informed consent forms during their first visit to the sperm bank, consenting to the use of their semen samples or data by the sperm bank for scientific research purposes. A total of 3231 qualified sperm donors were included in this study: all donors underwent primary screening for thalassemia and glucose-6-phosphate dehydrogenase (G6PD) deficiency. Whereafter, 278 of donors underwent genetic testing for specific genes, and 43 donors underwent whole exome sequencing. The Central South University Ethics Committee approved this study (2021-KT48).

### Criteria for sperm donor screening in China

The guidelines for screening sperm donors in China are as follows[8]: (a) donors must be between 22 and 45 years of age; (b) donors must be in good health, based on the results of a physical examination and a psychological evaluation by a qualified doctor, and have no familial history of a genetic disease; (c) fresh semen should have a liquefaction time of < 60 min, sperm concentration  $\geq 60 \times 10^6 \text{ mL}^{-1}$ , progressive sperm motility  $\geq 60\%$ , and normal morphology > 30%; (d) post-thaw semen should have motility of  $\geq 40\%$ ,  $\geq 12 \times 10^6$  motile sperm per vial, and a frozen–thaw survival rate  $\geq 60\%$ ; and (e) potential donors must undergo laboratory testing to exclude individuals at high risk for sexually transmitted infections and genetic diseases, including HIV-1 and HIV-2, hepatitis B and C, syphilis, gonorrhea, mycoplasma, chlamydia, cytomegalovirus, *Toxoplasma gondii*, rubella virus, and herpes simplex virus types 1 and 2, and undergo karyotype analysis. If the potential donor tests negative for all of the above tests and fulfills the Chinese Ministry of Health guidelines outlined above, the donation process is initiated and their semen samples are cryopreserved. The samples must be cryopreserved for a minimum 6-month quarantine period prior to rescreening for HIV.

### Criteria for management of donor sperm for external use in China

In China, sperm banks are required to maintain permanent records of the initial selection process and the subsequent tests and evaluations of each donor. The clinical outcomes (pregnant or not) for each treatment cycle (artificial insemination by donor [AID] and in vitro fertilization [IVF]) and fertility outcomes (abortion, healthy live birth, or birth defect) must also be recorded. Moreover, a sperm donor is permitted to impregnate a maximum of five women[2].

### Blood and DNA samples

The analyses were performed using blood and DNA samples collected for clinical purposes according to established protocols. 1.5ml blood samples were collected from qualified donors. DNA was extracted from the whole blood of the qualified sperm donors using commercially available DNA isolation kits (Genra Systems, Minneapolis, MN, USA) following the manufacturer's instructions.

### Premarital screening of thalassemia and G6PD

The blood sample for routine blood testing, hematology analyzers (HA) are used to measure the complete blood count which was screening sperm donors for thalassemia. G6PD deficiency Rapid Diagnostic Test (G6PD RDT, GuangZhou Jianlun biology Technology Co, Guangzhou, China) are used to screening sperm donors for G6PD deficiency. 2 $\mu$ l whole blood was placed into the sample window, immediately followed by two drops of buffer solution into the assay window according to the manufacturer's instructions. After ten minutes at the room temperature, the RDT was visually read and classified as deficient or normal.

### Genetic testing

Genetic testing was performed via whole exome sequencing or for specific genes. Samples from all participants in this study were submitted to the Baylor Miraca Genetics Laboratory for clinical exome sequencing (CES, performed at the AmCare Genomics Laboratory, Guangzhou, China), which included coding exons for about 5000 clinically relevant disease-causing genes[9]. The genomic DNA sequencing library was prepared by ultrasonic fragmentation, terminal repair, amplification, and purification, and the DNA sequence of the target region was enriched using custom-designed NimbleGen SeqCap probes (Roche NimbleGen, Madison, WI, USA). Then, the enriched DNA samples were sequenced on an Illumina HiSeq 2000 unit (Illumina, San Diego, CA, USA) with 150-bp single-end read length. All coding exons and 20 bp of the flanking intronic regions of the target genes were sequenced, with an average coverage depth of  $\sim 200\times$ . The original raw data obtained from the Illumina

platform were considered raw reads, and recorded in FASTQ files (fq). The data were filtered to generate clean reads by removing adapters and reads with base quality of < Q20. The sequencing reads were aligned to the human reference genome GRCh37/hg19. Specific gene testing only analyze regions of specific genes in the original raw data, while whole exome sequencing analyses all the regions in the original raw data. Single-nucleotide variants (SNVs) were observed with NextGENE software (version 2.4.1.2) (SoftGenetics, State College, PA, USA). Gene variants were annotated using population and literature databases, including the Genome Aggregation Database (gnomAD, <https://gnomad.broadinstitute.org/>), ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>), and Online Mendelian Inheritance in Man (OMIM). Variants with minor allele frequency (MAF) > 1% in Asian populations were filtered out. Pathogenicity and evolutionary conservative analyses were performed using PolyPhen and VarCards. The pathogenicity of the gene variants was classified according to American College of Medical Genetics and Genomics (ACMG) guidelines[10]. We eliminated low-quality sequencing data (the average coverage of the coding sequence was < 3×) to analyze copy number variations (CNVs). The copy number changes of exons were analyzed with homogenization calculation; 1 indicated a normal level (×2), 0.5 indicated one missing copy (×1), and 0 indicated two missing copies (×0), that is, homozygous/haplotype deletion. The pathogenicity of the CNVs was evaluated by referring to databases and published literature such as the Database of Genomic Variants (DGV), DECIPHER, and OMIM. Sperm donors without clinical symptoms were considered positive if they carried a pathogenic or likely pathogenic variation in the gene of concern to the recipient.

## Interpretation of variants

All variants found were classified as pathogenic (pathogenic and/or likely pathogenic), common (polymorphisms), or of uncertain significance [10]. Pathogenic variants were mutations classed as pathogenic or likely pathogenic in the ClinVar database and the Human Gene Mutation Database (HGMD, <http://www.hgmd.cf.ac.uk>), or described in the literature as associated with the disease. Variants of uncertain significance were variants not previously reported as pathogenic in the ClinVar database or HGMD and that presented frequencies in the population compatible with those of the disease of interest.

## Results

### Thalassemia and G6PD deficiency testing

That at the base of criteria for sperm donor screening in China, our sperm bank has increased the screening of thalassemia and G6PD deficiency for qualified sperm donors. The results as show in Table I, 2.4% of qualified sperm donors might have thalassemia and 1.4% might have G6PD deficiency. Sperm donors with thalassemia and G6PD deficiency would be eliminated.

### Specific gene testing

Table II summarizes the results for a part of genes that received requests for specific testing from donor semen recipients. Specific gene testing identified 7 of the 278 donors (2.5%) as carriers of at least one pathogenic or likely pathogenic variant in a gene, including 1.9% of 154 donors (3/154) as carrier variants in  $\alpha$ -Like or  $\beta$ -Like globin genes, 17.6% of 17 donors (3/17) as carrier variants in *GJB2*, 12.5% of 8 donors (1/8) as carrier variants in *SMN1*. In addition, most of the requested genes were the  $\alpha$ -like or  $\beta$ -like globin genes, followed by the genes related to G6PD deficiency, GJB2-related DFNB1 nonsyndromic hearing loss and deafness, and spinal muscular atrophy (*SMN1*). These were external-use donor semen recipients for more than 10 ART cycles. The remaining requested genes were *SLC26A4*, *HLAB27*, *PKHD1*, *ATPTB*, *TSCI*, *PAH*, and *ACSF3*.

### Whole exome sequencing

Whole exome sequencing in the clinical context of ART preconception screening identified 90.7% of the 43 donors (39/43) as carriers of at least one pathogenic or likely pathogenic variant in a gene responsible for a Mendelian genetic disease. Most of the donors were heterozygous for multiple variants (Fig. 1); 92.3% (36/39) were carriers of  $\geq 2$  genetic variants. Only four donor was a carrier of only one variant (2.6%, 1/39). Most were heterozygous for only two or three variants (71.8%, 28/39) (Fig. 1). The whole exome sequencing detected a total of six pathogenic variants (0% AD, 83.3% AR, 16.7% AR/AD, 0% Digenic Recessive (DR)), 105 likely pathogenic variants (3.8% AD, 73.3% AR, 20% AR/AD, 2.9% DR), and 184 variants of uncertain significance (21.2% AD, 61.4% AR, 16.8% AR/AD, 0.5% DR) in the 43 sperm donors (Table S1 and III). Among the 43 sperm donors carrying the 111 pathogenic/likely pathogenic variants, eight (18.6%) were carriers of pathogenic variants of the *GJB2* gene. The frequency, therefore, was approximately 1 in 5. Two of the 43 donors (4.65%) were carriers of pathogenic/likely pathogenic variants in the *SLC25A13*, *KDM5B*, *NEM8*, *DUOX2*, *CYP4V2*, *COQ8A*, *ROM1*, *DNAH1*, and *SERPINB7* genes; the frequency was approximately 1 in 20. Variants of pathogenic or likely pathogenic carriers in 43 sperm donors as shown in Supplementary Table S2, most of variants are common in Chinese.

### Clinical outcomes of Carrier screening

Specific gene testing was performed on a total of 278 donors if requested by donor semen recipients. During the study period, 793 treatment ART cycles were completed, 433 clinical pregnancies were achieved, and 309 offspring were delivered. Meanwhile, whole exome sequencing

was performed in a total of 43 donors, and we identified 45 cycles of preassigned donor–recipient matches with high reproductive risk for transmitting a severe AR genetic condition to the offspring and that required changing the initially selected donor for another without mutations in the same genes. We found two donor–recipient matches in which both were carriers for the same mutation, so the two recipients were matched to donors a second time. Twenty-five clinical pregnancies were achieved, and 18 offspring were delivered (Supplementary Table S3).

## Discussion

Here, the data from our human sperm bank showed that, other than three-generation family history analysis, specific gene testing and whole exome sequencing can be used for the genetic testing of sperm donors.

In China, sperm donor eligibility at sperm banks is determined in part by infectious disease screening and family history risk assessment. Routine testing for genetic disease is becoming increasingly important as part of the donor selection process due to the rising number and expanding availability of genetic tests. Meanwhile, recipients' demands for sperm donor genetic testing are also increasing, which has influenced donor screening and selection at human sperm banks. Unlike sperm banks in the United States, reasons for presenting for the use of sperm donor in China can only be that current partner has fertility issues. Therefore, lots of recipients who do not unacceptable sperm donors may also have genetic diseases. However, the sperm donor screening guidelines do not describe how genetic tests published by the Chinese Ministry of Health in 2003 [3], which are currently in effect [2], should be performed.

That at the base of criteria for sperm donor screening in China, sperm donors whose semen parameters up to standard would undergo further physical and laboratory examinations, including three-generation family history analysis, karyotype analysis and so on. In addition, our sperm bank has increased the screening of thalassemia and G6PD for qualified sperm donors. Thalassemia is considered one of the most common genetic disorders resulting from globin chain synthesis impairment because of the mutation or deletion of globin gene, such as  $\alpha$ - and  $\beta$ -thalassemia, with a high frequency in Southeast Asia [11]. In recent years, large-scale surveys for thalassemia have been conducted in different parts of China, and its prevalence remains high [12, 13]. A prevalence map based on a geographic information system (GIS) showed that the geographic distribution of thalassemia was highest in the south of China, including Guangdong, Guangxi, Guizhou, Yunnan and so on, and decreased from south to north [14]. In the present study, we used blood routine testing to screening of thalassemia related gene carriers in qualified sperm donors, if blood routine results showed a decrease in mean corpuscular volume (MCV) and mean corpuscular hemoglobin (MCH), accompanied by an increase in red cell count (RBC), we would think that the sperm donors were suspected thalassemia related gene carriers, whom would be eliminated. Because of less globin production, the thalassemia RBC showed microcytic and hypochromic, MCV, MCH and RBC were valuable indicators for screening thalassemia related gene carriers [15–17]. However, MCV, MCH and RBC cannot discriminate between  $\alpha$ - and  $\beta$ -thalassemia, and cannot discover heterozygous thalassemia related gene carriers [18]. Even so, 2.4% of sperm donors (80/3231) were suspected thalassemia related gene carriers, which was significantly lower than another study in Hunan province that analyzed by using Next-Generation Sequencing [19]. G6PD deficiency is one of the most common X-linked enzymopathies caused by *G6PD* gene variants, and G6PD deficiency is a common genetic disease in China. The incidence of G6PD deficiency in China is also characterized by a high to low gradient distribution from the south to north regions [20]. In the present study, we used RDT to diagnosis G6PD deficiency in qualified sperm donors, which can accurately identify hemizygous males and homozygous females. The data showed that 1.4% of sperm donors (56/3231) were G6PD deficiency whom would be eliminated. The overall prevalence of G6PD deficiency in China was 2.10% at the national level, and in Hunan province was 1.13% [21], which was similar to our results. Therefore, through the interpretation of blood routine and RDT results of sperm donors, it can make a preliminary screening of thalassemia and G6PD deficiency, and improve the laboratory examination of sperm donor genetic testing.

As early as 2010, all sperm banks in the United States have conducted some degree of genetic testing on their sperm donor applicants as part of the screening process. Furthermore, cystic fibrosis (CF) carrier screening, hemoglobin evaluation, and chromosome analysis are performed routinely for the majority of donor applicants [5]. CF is one of the most common severe AR diseases in the white population, with an approximate incidence of 1 in 2500 live births [22]. However, it is least common in Africans and Asians [23]. Therefore, CF carrier screening is not available at human sperm banks in China. Our data show that sperm donor carrier screening for the thalassemia-related genes was the most requested by recipients. Meanwhile, recipient requests for sperm donor carrier screening for *G6PD* genes were only second to that for thalassemia. 1.9% of 154 donors (3/154) as carrier variants in  $\alpha$ -Like or  $\beta$ -Like globin genes, although thalassemia and G6PD deficiency have been preliminary screened. This also shows while MCV, MCH and RBC were valuable indicators for screening thalassemia related gene carriers, genetic testing is the gold standard for diagnosing thalassemia related gene carriers. The incidence of genetic disease varies in different regions in China. Hence, genetic testing of sperm donors should be performed with targeted genetic disease-related genes of the different regions. In particular, donors of semen for external use in the Guangxi and Guangdong provinces in south China, which have a high incidence of thalassemia and G6PD deficiency, should be tested for the  $\alpha$ -like,  $\beta$ -like globin genes and *G6PD* gene.

In the present study, carrier genetic testing was performed using NGS, which was first used in infertile couples wishing to conceive through ART in 2015. In the clinical dataset, 2161 samples (84%) tested positive, with an average carrier burden of 2.3 variants per sample. Five percent of

the couples wishing to conceive through ART were carriers for the same mutation; genetic screening prevented the birth of 1.25% of genetically affected babies born after ART [24]. Carrier genetic testing can be applied to couples wishing to conceive through ART, as well as ART with sperm or oocyte donors, to avoid serious monogenic genetic diseases [7]. Comparison with the Exome Aggregation Consortium (ExAC) East Asian and European populations showed that the carrier frequency of disease-related genes identified at the Translational Medicine Center of the Children's Hospital of Fudan University is much more similar to that of the East Asian population than the European population. The difference in carrier frequency among populations should not be ignored and makes it necessary to establish a Chinese-specific panel for genetic testing [25]. In the present study, we performed whole exome sequencing carrier screening, which included more than 5000 genes associated with more than 7000 disorders in OMIM. We tested 43 sperm donors, with an average carrier pathogenic/likely pathogenic variant of 2.58 per sample. Forty-five cycles of preassigned donor–recipient matches were identified, 25 clinical pregnancies were achieved, and 18 offspring was delivered. Although the data are based on a limited number of participants, the results provide the carrier frequencies for many recessive disorders. Here, we found that, for many of the studied genes, the carrier frequencies for some of the most common recessive disorders, such as *GJB2*-related DFNB1 nonsyndromic hearing loss and deafness, citrullinemia argininosuccinate synthase deficiency, late-onset Alzheimer disease, and AR mental retardation, are higher than that previously reported [26–29]. This is not unexpected, as previous studies on carrier frequencies were based on known mutations, whereas here we explored the entire coding region of most genes. In addition, eight of 43 donors (18.6%) were carriers of pathogenic or likely pathogenic variants in *GJB2*; the frequency therefore was approximately 1 in 5. Meanwhile, in special gene testing, three of 17 donors (17.6%) also as carrier variants in *GJB2*. However, as different *GJB2* pathogenic variants are associated with various phenotypes of hearing loss. For instance, both c.109G > A homozygotes and compound heterozygous c.109G > A variants of *GJB2* indicate a significantly higher risk of developing hearing loss. Conversely, heterozygous c.109G > A variants alone do not increase the risk of hearing loss [30]. In the present study, five of 8 variants were heterozygous c.109G > A variants, that do not increase the risk of hearing loss. Although the carrier frequencies of *GJB2* should not be over-interpreted, the frequency of *GJB2* variants should be a concern in sperm donors in China.

Whole exome sequencing and specific gene carrier screening are powerful tools, especially for certain high-risk ethnic populations. However, there is considerable debate within the sperm banking community as to which tests should be included in the donor application process, in part because genetic testing is expensive and adds considerably to the cost of qualifying as a sperm donor. Nonetheless, there is no guideline in China on the proper use of sperm samples after genetic testing, which also limits the application of genetic screening in sperm banks in China. However, many inherited risks cannot be detected before donor qualification even when a thorough genetic family history evaluation has been performed. Hence, genetic testing for sperm donors can improve the quality of donor screening effectively and reduce the genetic risks for the offspring. Meanwhile, genetic testing was required not only for sperm donors but also for recipients, aid a successful and healthy pregnancy. In additions, it is important that human sperm banks engage the services of genetics professionals so that their clients have access to counseling about their family's medical histories and the value and limitations of genetic testing and its role in reducing the risk of birth defects in future offspring.

Our study has some limitations: we did not perform genetic testing of the donors' offspring, as congenital abnormalities were not reported therein. Furthermore, our study is not based on the all general population wishing to conceive, as the study group consisted of sperm donors who were young men. Although 321 qualified sperm donors underwent genetic testing, a larger cohort would provide more information. More studies of this type are needed.

This is the first study to investigate genetic testing of donor specimens from a sperm bank in China. Our results are significant for sperm donors, recipients, and future offspring. The data clearly illustrate that the significant risks presented by donors include inherited susceptibility for AR and undiagnosed AD disorders, and suggest that used blood routine and RDT can make a preliminary screening of sperm donors, and special gene testing be performed for sperm donors according to the regional incidence of specific genetic diseases. Meanwhile, whole exome sequencing can be used as a supplementary application in sperm donor genetic testing, and aid a successful and healthy pregnancy. However, industry guidelines must be modified to incorporate its use.

## Declarations

### Ethics approval and consent to participate

The study protocol was approved by the Ethics Committee of Central South University approved of this study (approval no. 2021-KT48).

### Acknowledgements

Not applicable.

### Availability of data and materials

The datasets used or analyzed during the current study are available from the corresponding author on reasonable request.

## Funding

This work was supported by grants from the National Natural Science Foundation of China (grant number 82001634) and the China Postdoctoral Science Foundation (grant number 2019M661521).

## Authors' contributions

Study conceptualization and design, patient recruitment, and data collection: H.C., F.L.Q., and Z.W.B. Data analysis and manuscript drafting: H.C. and N.H.C. Patient recruitment and data collection: W.H.L. and L.Q. Draft revision: Z.W.J., H.Z.H., L.X.F., and T.Y.L. All authors approved the final version of the manuscript.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## References

1. Isley L, Falk RE, Shamonki J, Sims CA, Callum P: **Management of the risks for inherited disease in donor-conceived offspring.** *Fertil Steril* 2016, **106**:1479–1484.
2. Huang C, Lei L, Wu HL, Gan RX, Yuan XB, Fan LQ, Zhu WB: **Long-term cryostorage of semen in a human sperm bank does not affect clinical outcomes.** *Fertil Steril* 2019, **112**:663–669 e661.
3. Health CMO: **Basic Standards and Technical Specifications for Human Sperm Bank.** *Chin J Reprod Heal* 2004, **15**:481.
4. Practice Committee of the American Society for Reproductive Medicine tPCftSfART: **Guidance regarding gamete and embryo donation.** *Fertil Steril* 2021, **115**:1395–1410.
5. Sims CA, Callum P, Ray M, Iger J, Falk RE: **Genetic testing of sperm donors: survey of current practices.** *Fertil Steril* 2010, **94**:126–129.
6. Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML: **Genome-wide genetic marker discovery and genotyping using next-generation sequencing.** *Nat Rev Genet* 2011, **12**:499–510.
7. Abuli A, Boada M, Rodriguez-Santiago B, Coroleu B, Veiga A, Armengol L, Barri PN, Perez-Jurado LA, Estivill X: **NGS-Based Assay for the Identification of Individuals Carrying Recessive Genetic Mutations in Reproductive Medicine.** *Hum Mutat* 2016, **37**:516–523.
8. Huang C, Li B, Xu K, Liu D, Hu J, Yang Y, Nie H, Fan L, Zhu W: **Decline in semen quality among 30,636 young Chinese men from 2001 to 2015.** *Fertil Steril* 2017, **107**:83–88 e82.
9. Meng L, Pammi M, Saronwala A, Magoulas P, Ghazi AR, Vetrini F, Zhang J, He W, Dharmadhikari AV, Qu C, et al: **Use of Exome Sequencing for Infants in Intensive Care Units: Ascertainment of Severe Single-Gene Disorders and Effect on Medical Management.** *JAMA Pediatr* 2017, **171**:e173438.
10. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, et al: **Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology.** *Genet Med* 2015, **17**:405–424.
11. Weatherall DJ, Clegg JB: **Inherited haemoglobin disorders: an increasing global health problem.** *Bull World Health Organ* 2001, **79**:704–712.
12. Tang W, Zhang C, Lu F, Tang J, Lu Y, Cui X, Qin X, Li S: **Spectrum of alpha-thalassemia and beta-thalassemia mutations in the Guilin Region of southern China.** *Clin Biochem* 2015, **48**:1068–1072.
13. Lin M, Zhong TY, Chen YG, Wang JZ, Wu JR, Lin F, Tong X, Yang HT, Hu XM, Hu R, et al: **Molecular epidemiological characterization and health burden of thalassemia in Jiangxi Province, P. R. China.** *PLoS One* 2014, **9**:e101505.
14. Lai K, Huang G, Su L, He Y: **The prevalence of thalassemia in mainland China: evidence from epidemiological surveys.** *Sci Rep* 2017, **7**:920.
15. Hartevelde CL, Higgs DR: **Alpha-thalassaemia.** *Orphanet J Rare Dis* 2010, **5**:13.
16. Brancaleoni V, Di Pierro E, Motta I, Cappellini MD: **Laboratory diagnosis of thalassemia.** *Int J Lab Hematol* 2016, **38 Suppl 1**:32–40.
17. Risoluti R, Materazzi S, Sorrentino F, Bozzi C, Caprari P: **Update on thalassemia diagnosis: New insights and methods.** *Talanta* 2018, **183**:216–222.

18. England JM, Fraser PM: **Differentiation of iron deficiency from thalassaemia trait by routine blood-count.** Lancet 1973, 1:449–452.
19. Yang M, Li CY, Lei DZ, Zhang HQ: **[Application of Next-Generation Sequencing in Screening of Thalassaemia Gene in 11212 Pregnant Women in Suxian and Beihu Districts of Chenzhou City, Hunan Province].** Zhongguo Shi Yan Xue Ye Xue Za Zhi 2021, 29:188–192.
20. Liu Z, Yu C, Li Q, Cai R, Qu Y, Wang W, Wang J, Feng J, Zhu W, Ou M, et al: **Chinese newborn screening for the incidence of G6PD deficiency and variant of G6PD gene from 2013 to 2017.** Hum Mutat 2020, 41:212–221.
21. He Y, Zhang Y, Chen X, Wang Q, Ling L, Xu Y: **Glucose-6-phosphate dehydrogenase deficiency in the Han Chinese population: molecular characterization and genotype-phenotype association throughout an activity distribution.** Sci Rep 2020, 10:17106.
22. Rowntree RK, Harris A: **The phenotypic consequences of CFTR mutations.** Ann Hum Genet 2003, 67:471–485.
23. O'Sullivan BP, Freedman SD: **Cystic fibrosis.** Lancet 2009, 373:1891–1904.
24. Martin J, Asan, Yi Y, Alberola T, Rodriguez-Iglesias B, Jimenez-Almazan J, Li Q, Du H, Alama P, Ruiz A, et al: **Comprehensive carrier genetic test using next-generation deoxyribonucleic acid sequencing in infertile couples wishing to conceive through assisted reproductive technology.** Fertil Steril 2015, 104:1286–1293.
25. Ling Y, Wang H, Wu B, Wang L, Dong C, Lu Y, Zhou W: **Carrier analysis of known disease-causing genes based on WES data from single Shanghai pediatric center.** Chin J Evid Based Pediatr 2016, 11:204–209.
26. Morton CC, Nance WE: **Newborn hearing screening—a silent revolution.** N Engl J Med 2006, 354:2151–2164.
27. Liu SL, Wang XC, Tan MS, Wang HF, Zhang W, Wang ZX, Yu JT, Tan L: **NME8 rs2718058 polymorphism with Alzheimer's disease risk: a replication and meta-analysis.** Oncotarget 2016, 7:36014–36020.
28. Saheki T, Moriyama M, Funahashi A, Kuroda E: **AGC2 (Citrin) Deficiency-From Recognition of the Disease till Construction of Therapeutic Procedures.** Biomolecules 2020, 10.
29. Xue S, Lam YM, He Z, Zheng Y, Li L, Zhang Y, Li C, Mbadhi MN, Zheng L, Cheng Z, et al: **Histone lysine demethylase KDM5B maintains chronic myeloid leukemia via multiple epigenetic actions.** Exp Hematol 2020, 82:53–65.
30. Shen N, Peng J, Wang X, Zhu Y, Liu W, Liu A, Lu Y: **Association between the p.V37I variant of GJB2 and hearing loss: a pedigree and meta-analysis.** Oncotarget 2017, 8:46681–46690.

## Tables

**Table I Primary screening for thalassaemia and G6PD with the whole blood of 3231 sperm donors**

| Genetic disorder                             | Donors, n | Abnormal, n | Population proportion |
|--|-----------|-------------|-----------------------|
| Thalassaemia $\alpha\beta^0$                 | 3231      | 80          | 2.4%                  |
| Glucose-6-phosphate dehydrogenase deficiency | 3231      | 56          | 1.4%                  |

**Table II Genetic testing for specific disorder-related genes for 278 sperm donors in sperm donor applicants.**

| Gene                | Genetic disorder  | Inheritance | Donors, n | Abnormal, n | Abnormal proportion, n | Cycles, n | Pregnancies, n | Miscarriages, n | Deliveries, n |
|---------------------|---|-------------|-----------|-------------|------------------------|-----------|----------------|-----------------|---------------|
| α-Like globin genes | Thalassemia   | AR          | 154       | 2           | 1.3%                   | 580       | 325            | 32              | 228           |
| β-Like globin genes |   |             |           |             |                        |           |                |                 |               |
| G6PD                | Glucose-6-phosphate dehydrogenase deficiency              | XD          | 23        | 1           | 4.3%                   | 59        | 28             | 3               | 22            |
| GJB2                | GJB2-related DFNB1 nonsyndromic hearing loss and deafness | AR/AD       | 17        | 3           | 17.6%                  | 23        | 7              | 1               | 5             |
| SMN1                | Spinal muscular atrophy                                   | AR          | 8         | 1           | 12.5%                  | 12        | 5              | 2               | 2             |
| SLC26A4             | Epicophosis   | AR          | 4         | 0           | 0%                     | 5         | 2              | 0               | 1             |
| HLAB27              | Ankylosing spondylitis                                    | AD          | 3         | 0           | 0%                     | 4         | 2              | 0               | 2             |
| PKHD1               | Polycystic kidney   | AR          | 3         | 0           | 0%                     | 3         | 0              | 0               | 0             |
| ATP7B               | Hepatolenticular degeneration                             | AR          | 2         | 0           | 0%                     | 2         | 1              | 0               | 1             |
| TSC1                | Tuberous sclerosis complex (Bourneville-Pringle disease)  | AD          | 2         | 0           | 0%                     | 6         | 5              | 0               | 5             |
| PAH                 | Phenylketonuria   | AR          | 2         | 0           | 0%                     | 6         | 5              | 1               | 4             |
| ACSF3               | Combined malonic and methylmalonic aciduria               | AR          | 2         | 0           | 0%                     | 5         | 4              | 2               | 2             |

Note: Genes tested only once are not listed.

Autosomal Dominant (AD); Autosomal Recessive (AR); X-linked dominant (XD)

**Table III Whole-exome sequencing detection of identical carrying genes in 43 sperm donors**

| Gene     | Inheritance | Donors, n | Population proportion | Genetic disorder  |
|----------|-------------|-----------|-----------------------|---|
| GJB2     | AR/AD       | 8         | 18.6%                 | GJB2-related DFNB1 nonsyndromic hearing loss and deafness |
| SLC25A13 | AR          | 2         | 4.65%                 | Citrullinemia argininosuccinate synthase deficiency       |
| KDM5B    | AR          | 2         | 4.65%                 | AR mental retardation                                     |
| NME8     | AR          | 2         | 4.65%                 | Late-onset Alzheimer disease                              |
| DUOX2    | AR          | 2         | 4.65%                 | Thyroid dyshormonogenesis (dyshormonogenetic goiter)      |
| CYP4V2   | AR          | 2         | 4.65%                 | Bietti crystalline corneoretinal dystrophy                |
| COQ8A    | AR          | 2         | 4.65%                 | Coenzyme Q10 deficiency                                   |
| ROM1     | AR/AD       | 2         | 4.65%                 | Retinitis pigmentosa                                      |
| DNAH1    | AR          | 2         | 4.65%                 | Primary ciliary dyskinesia                                |
| SERPINB7 | AR          | 2         | 4.65%                 | Palmoplantar keratoderma, Nagashima type                  |

## Figures

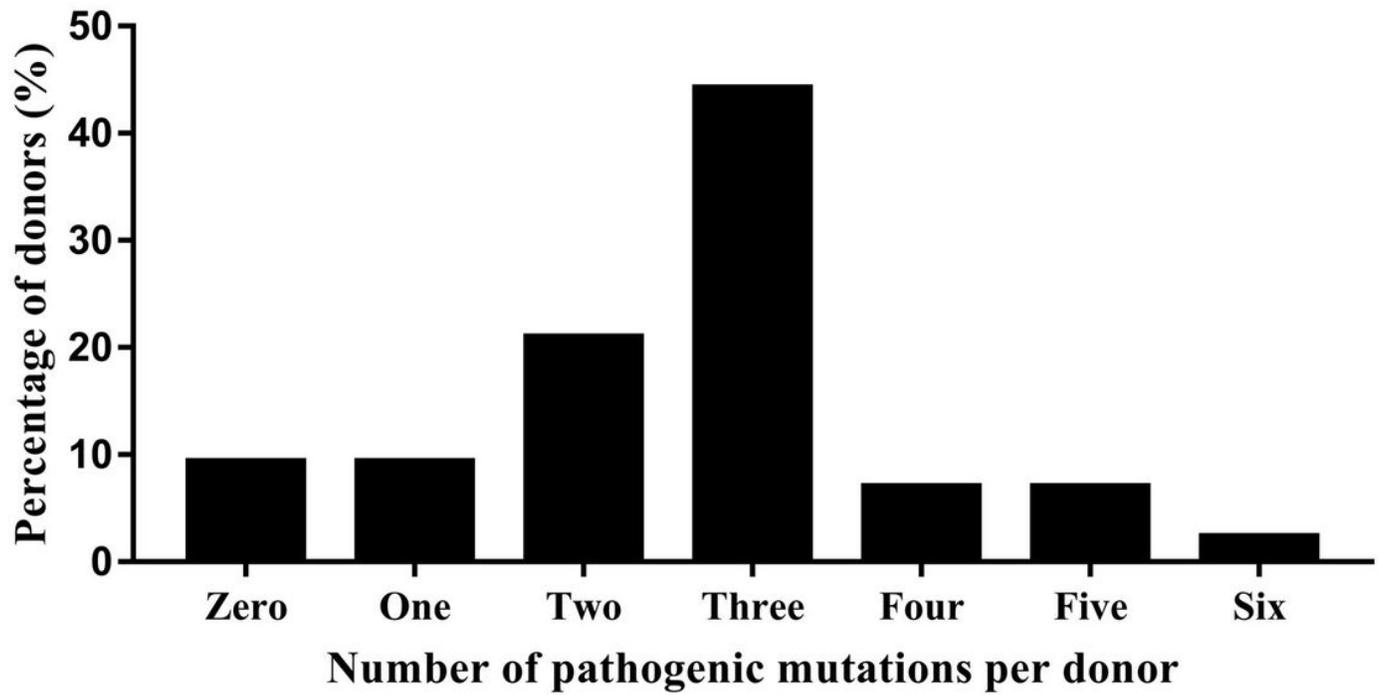


Figure 1

Whole exome sequencing detection of the number and percentage of pathogenic variants in 43 donors.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS.doc](#)