Identification of pathogenic genomic variants in idiopathic azoospermia by long-fragment-read genome sequencing: abridged secondary publication

Z Dong *, TY Leung, JPW Chung, Y Cao

KEY MESSAGES

1. We recruited 100 men with idiopathic infertility who had negative findings in karyotype analysis and Y-microdeletion testing for long-fragment-read genome sequencing (GS).
2. Long-fragment-read GS identified likely clinically significant variants in 32 men (19 with nonobstructive azoospermia and 13 with severe oligoasthenozoospermia); the diagnostic yield was 32.0%.
3. Genetic investigation of variants by long-fragment-read GS facilitates both understanding of infertility mechanisms and patient management, including sperm retrieval through testicular biopsy, intracytoplasmic sperm injection, and/or preimplantation genetic testing via in vitro fertilisation.
4. Long-fragment-read GS can be used as a second-tier approach for the genetic diagnosis of infertile men.

Hong Kong Med J 2026;32(Suppl 1):S17-21
HMRF project number: 09200236

Z Dong, TY Leung, JPW Chung, Y Cao

Department of Obstetrics and Gynaecology, The Chinese University of Hong Kong, Hong Kong SAR, China

* Principal applicant and corresponding author: elvisdong@cuhk.edu.hk

Introduction

Male infertility affects approximately 7% of the male population.[1] Its phenotypic spectrum is heterogeneous, ranging from the absence of sperm to alterations in sperm quality.[1] Genetic testing is recommended by the American Society for Reproductive Medicine[2] for men with nonobstructive azoospermia (NOA) or severe oligoasthenozoospermia (OAT) to clarify the aetiology prior to surgical intervention,[3] including polymerase chain reaction (PCR) and karyotype analyses for Y-microdeletions and chromosomal abnormalities, respectively. Exome sequencing provides a diagnostic yield of approximately 20%.[4,5] We developed multiple-barcoded long-fragment-read genome sequencing (GS)[6] for comprehensive detection of copy-number variants (CNVs), structural variants (SVs), single-nucleotide variants (SNVs), small insertions/deletions (indels), and regions of absence of heterozygosity (AOH). We used this long-fragment-read GS to assess 100 men with idiopathic infertility and identified potential causative variants in 32 men, revealing a diagnostic yield of 32.0%.

Methods

We recruited men with severe OAT (<5 million sperm/mL ejaculate fluid and/or sperm malformation) or azoospermia with elevated follicle-stimulating hormone, or with negative results from Y-microdeletion and karyotype analyses.

Buffy coats were retrieved for long DNA extraction and subjected to long-fragment-read GS at >40-fold read depth.[6] RNA was extracted. SNVs/indels, CNVs, SVs, and regions with AOH were verified by Sanger sequencing, breakpoint-junction-specific PCR, and quantitative PCR. Ten cases with potential causative variants were also subjected to long-read sequencing. RNA sequencing with enrichment of messenger RNA was performed for >10 Gb data per sample. Paired-end reads were aligned to the human reference genome (hg19) with STAR. Expression patterns of candidate genes were compared with our in-house fertile controls. Two published single-cell sequencing datasets,[7,8] consisting of 74 845 cells derived from 12 time points in prenatal and postnatal testicular development, were curated by our Temporal Expression during Development Database[9] for cell type–specific expression analysis.

Results

In total, 63 men with NOA and 37 men with severe OAT were recruited (Table 1). Long-fragment-read GS identified likely clinically significant variants in 32 men (19 with NOA and 13 with severe OAT); the diagnostic yield was 32.0%. There were no significant differences in the diagnostic yields of each variant type between NOA and OAT groups (Table 2).

TABLE 1. Characteristics of 100 men with idiopathic infertility.

Characteristic	All (n=100)*	Nonobstructive azoospermia (n=63)*	Severe oligoasthenozoospermia (n=37)*	P value
Age, y	36±5	36±5.7	36±4.4	0.998
Follicle-stimulating hormone, IU/mL	20.7±13.2	24.9±13.1	11.1±6.0	<0.001
Testosterone, ng/dL	12.7±6.4	12.3±12.3	13.9±4.1	0.450

* Data are presented as mean ± standard deviation.

TABLE 2. Diagnostic yields of different causative variants in 100 men with idiopathic infertility.

Variant	All (n=100)*	Nonobstructive azoospermia (n=63)*	Severe oligoasthenozoospermia (n=37)*	P value
Single-nucleotide variants, small insertions/deletions	11	8	3	0.705
Copy-number variants, structural variants	16	9	7	0.743
Both	5	2	3	0.537

* Data are presented as No. of patients.

Long-fragment-read GS identified compound heterozygosity for pathogenic/likely pathogenic (P/LP) variants in seven men. Five of them involved one SNV and one CNV in an autosomal recessive disease-causing gene. For instance, in patient MI#33, we identified a heterozygous missense SNV (NM_015180:c.18389A>T:p.K6130I) and a heterozygous deletion (seq[GRCh37] del(14)(q23.2) NC_000014.8:g.64454090_64458143del) in the *SYNE2* gene (Fig 1). Both variants were predicted to result in loss of function. Given that *SYNE2* is part of a family of proteins involved in sperm-head formation[10] and shows significantly reduced expression in men with NOA,[11] biallelic loss of function might underlie its pathogenicity.

Among the 16 men with identified P/LP CNVs/SVs, five had these variants on the X chromosome. Patient MI#16 was diagnosed with NOA. Ultrasound examination revealed small testes (5 cc) and a small epididymal cyst. A hemizygous 1.14 Mb deletion in Xq28 was detected (seq[GRCh37] del(X)(q28) chrX:g.150838316_151978891del), involving the *MAGEA* and *CSAG* gene clusters, as well as the *FATE1* gene; however, none was expressed in the buffy coats (Fig 2). The *MAGEA* gene cluster is crucial for maintaining normal testicular size in mice, and knockout findings suggest that specific defects occur during the first wave of spermatogenesis. Single-cell sequencing data indicated *MAGEA* expression in primordial germ cells. The protein encoded by *FATE1* is known to regulate early testicular differentiation and cell proliferation, and *FATE1* is highly expressed in Sertoli cells during early testicular development. In comparison, depletion of *CSAG1* disrupts centrosomes and leads to multipolar spindles during mitosis, although no expression of CSAG1 was identified in fetal testicular samples. Both *FATE1* and *CSAG1* are highly expressed in adult germ cells. Therefore, the lack of expression of these genes likely contributed to small testes (*MAGEA* gene cluster and *FATE1*) and azoospermia (*CSAG1*).

P/LP CNVs/SVs involving autosomal chromosomes were identified in 11 men (five with NOA and six with OAT). Patient MI#48 had severe OAT. A 53-kb copy-number gain (seq[GRCh37] dup(2)(q31.1) chr2:g.171170285_171223517dup) involving the eighth exon (NM_001083615) of *MYO3B* on 2q31.1 was inserted into the fourth intron (NM_021951) of *DMRT1* (Fig 3). Flanking the insertion was a duplication (seq[GRCh37] dup(9)(p24.3) chr9:g.910281_931988dup) involving the fourth exon of *DMRT1*, which resulted in *DMRT1* truncation. *DMRT1* is highly expressed in prenatal Sertoli cells and primordial germ cells, in adulthood, in spermatogonia and early spermatids. Additionally, heterozygous deletion of *DMRT1* causes non-syndromic 46,XY disorders of sexual development.[12]

Long-fragment-read GS also detected multiple regions with AOH in four men with NOA. Two interstitial regions with AOH were identified, with an overall size of 18.7 Mb in patient MI#53, with suspected uniparental disomy of chromosome 6. However, there was an absence of causative SNVs/indels in genes related to male infertility. In comparison, the overall sizes of regions with AOH identified in patients MI#81, MI#77, and MI#73 were 68 Mb, 81.3 Mb, and 218.6 Mb, respectively, suggesting parental consanguinity. In patient MI#77, a 7.9-kb homozygous deletion involving the 5′ end

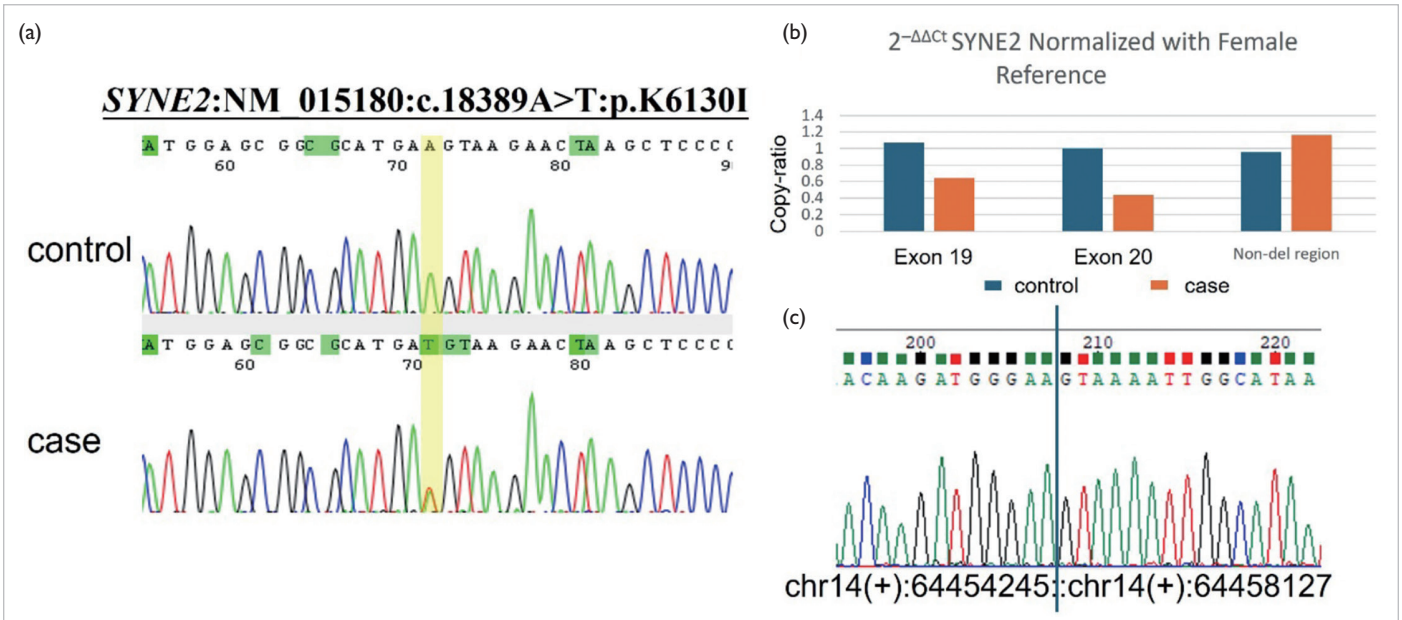

FIG 1. Compound heterozygosity of variants identified in patient MI#33: (a) Sanger sequencing confirmed the presence of the heterozygous variant. (b) Reverse transcription quantitative polymerase chain reaction confirmed a heterozygous deletion. (c) Breakpoint-specific polymerase chain reaction and Sanger sequencing revealed the junction sequence of this heterozygous deletion (3.8 kb in size).

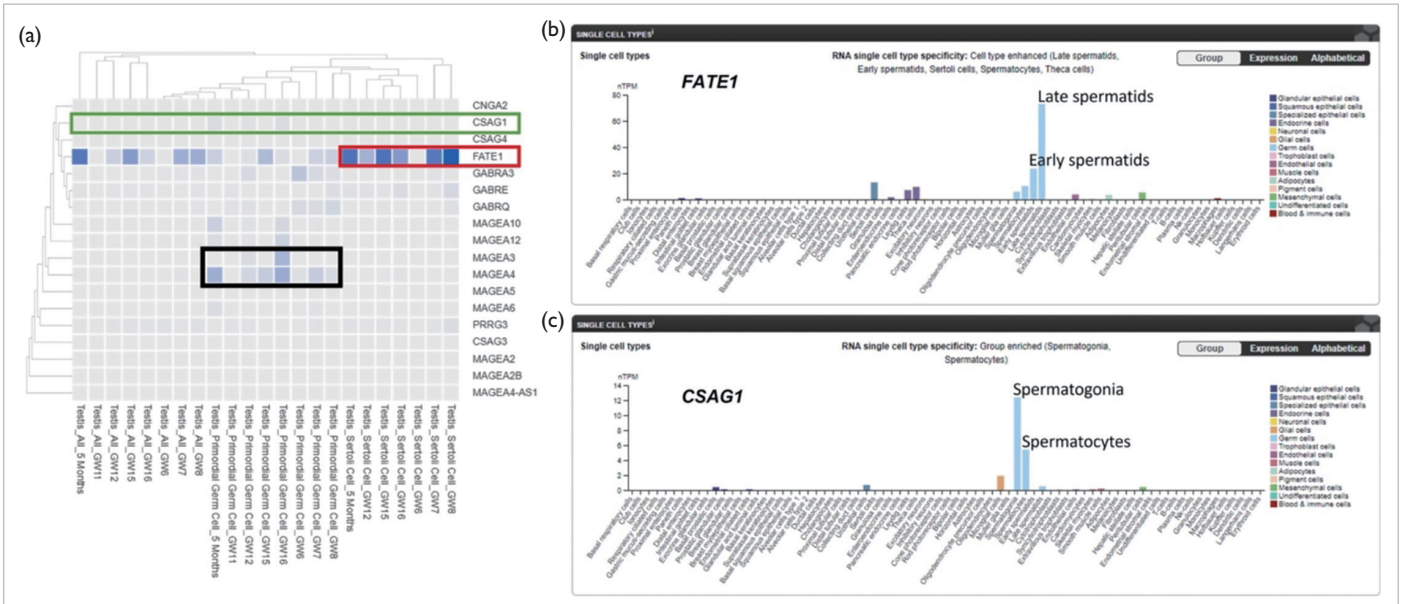

FIG 2. Gene expression involving the hemizygous deletion in patient MI#16: (a) Clustering of gene expression from single-cell sequencing during fetal testicular development: the top box shows the absence of expression in each identified cell type for *CSAG1*, the middle box shows the expression profiles of *FATE1* in Sertoli cells, and the bottom box displays the expression patterns of *MAGEA* gene cluster in primordial germ cells. (b) and (c) Expression profiles of *FATE1* and *CSAG1* in adult germ cells.

and first exon of the *NSUN7* gene was identified due to parental consanguinity.[13]

Discussion

Long-fragment-read GS identified clinically significant variants that potentially explain male infertility phenotypes in 32 of 100 men with idiopathic infertility. It was able to detect point mutations and smaller CNVs/SVs, compared with mate-pair sequencing.[14]

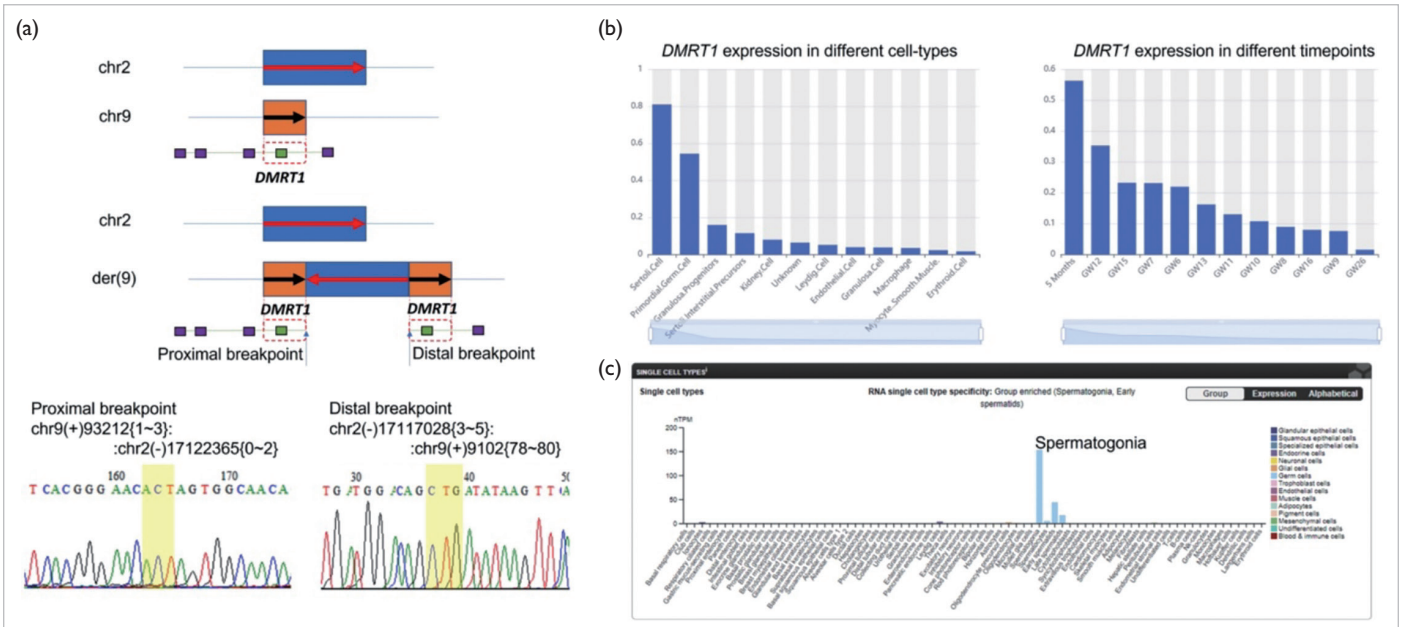

FIG 3. An interchromosomal insertion identified in patient MI#48: (a) The upper panel indicates the composition of normal chromosomes 2 and 9, whereas the lower panel indicates abnormalities at the 53.2 kb segment (chr2) and the 21.7 kb segment (chr9), respectively. Arrows indicate sequence orientation. Composition of der(9) shows insertion of a segment from 2q31.1 in antisense orientation, with subsequent duplication of a segment originating from 9p24.3. Breakpoint junction–specific polymerase chain reaction and Sanger sequencing show sequence microhomologies at both breakpoint junctions. (b) Abundance of *DMRT1* gene expression in different cell types and time points during testicular development. (c) Expression profile of *DMRT1* in adult germ cells.

Current management for patients with NOA is to perform testicular biopsy, followed by in vitro fertilisation through intracytoplasmic sperm injection (ICSI). Other factors may contribute to ICSI failure, even after sperm retrieval. Thus, information regarding the likelihood of ICSI failure is important for counselling. For severe OAT, knowledge of pathogenic variants is important to facilitate further genetic counselling if preimplantation testing is under consideration, along with conception by in vitro fertilisation and ICSI. Taken together, genetic investigation of variants using long-fragment-read GS is important for understanding the mechanisms of infertility and informing patient management.

Future research may make use of long-fragment-read GS, which achieved a diagnostic yield of 32.0%. Additionally, 21 cases were fully or partially attributed to CNVs/SVs, which could not be readily detected by exome sequencing or conventional GS. Annotation and interpretation of these intragenic changes are warranted because some might be causative in patients via dysregulation of disease-causing genes. Furthermore, establishment of the Temporal Expression during Development Database provides cell- or tissue-type- and time-point-specific expression profiles of target genes. Although RNA sequencing was conducted, the pathogenicity of genomic changes could not be confirmed by the presence of gene expression aberrations or splicing defects, given that most target genes showed no expression. Therefore, establishment of these profiles would enable better understanding of the contributions of genomic changes to infertility. Follow-up studies to investigate potential management options are warranted. Some genomic variants were correlated with clinical outcomes. Because certain variants are known to affect one or more testicular cell types, such as Leydig cells and Sertoli cells, future studies should explore potential applications of cell-type-specific treatment options or even gene therapy.

Conclusion

Long-fragment-read GS can be used as a second-tier approach for the genetic diagnosis of infertile men.

Funding

This study was supported by the Health and Medical Research Fund, Health Bureau, Hong Kong SAR Government (#09200236). The full report is available from the Health and Medical Research Fund website (https://rfs1.healthbureau.gov.hk).

Disclosure

The results of this research have been previously published in:

1. Zhou Z, Tan C, Chau MHK, et al. TEDD: a database of temporal gene expression patterns during multiple developmental periods in human and model organisms. Nucleic Acids Res 2023;51:D1168-78.
2. Dong Z, Qian J, Law TSM, et al. Mate-pair genome sequencing reveals structural variants for idiopathic male infertility. Hum Genet 2023;142:363-77.

Acknowledgements

We thank Prof Kwong Wai Choy, Dr Matthew Hoi Kin Chau, and Dr Jicheng Qian for their contributions to the laboratory work.

References

1. Krausz C, Riera-Escamilla A. Genetics of male infertility. Nat Rev Urol 2018;15:369-84.
2. Schlegel PN, Sigman M, Collura B, et al. Diagnosis and treatment of infertility in men: AUA/ASRM guideline part I. Fertil Steril 2021;115:54-61.
3. Wosnitzer MS. Genetic evaluation of male infertility. Transl Androl Urol 2014;3:17-26.
4. Krausz C, Riera-Escamilla A, Moreno-Mendoza D, et al. Genetic dissection of spermatogenic arrest through exome analysis: clinical implications for the management of azoospermic men. Genet Med 2020;22:1956-66.
5. Alhathal N, Maddirevula S, Coskun S, et al. A genomics approach to male infertility. Genet Med 2020;22:1967-75.
6. Wang O, Cheng X, Drmanac R, Peters BA. A simple cost-effective method for whole-genome sequencing, haplotyping, and assembly. Methods Mol Biol 2023;2590:101-25.
7. Guo J, Sosa E, Chitiashvili T, et al. Single-cell analysis of the developing human testis reveals somatic niche cell specification and fetal germline stem cell establishment. Cell Stem Cell 2021;28:764-78.e4.
8. Han X, Zhou Z, Fei L, et al. Construction of a human cell landscape at single-cell level. Nature 2020;581:303-9.
9. Zhou Z, Tan C, Chau MHK, et al. TEDD: a database of temporal gene expression patterns during multiple developmental periods in human and model organisms. Nucleic Acids Res 2023;51:D1168-78.
10. Castaneda JM, Shimada K, Satouh Y, et al. FAM209 associates with DPY19L2, and is required for sperm acrosome biogenesis and fertility in mice. J Cell Sci 2021;134:jcs259206.
11. Chen S, An G, Wang H, et al. Human obstructive (postvasectomy) and nonobstructive azoospermia: insights from scRNA-Seq and transcriptome analysis. Genes Dis 2022;9:766-76.
12. Ledig S, Hiort O, Wunsch L, Wieacker P. Partial deletion of DMRT1 causes 46,XY ovotesticular disorder of sexual development. Eur J Endocrinol 2012;167:119-24.
13. Khosronezhad N, Hosseinzadeh Colagar A, Mortazavi SM. The Nsun7 (A11337)-deletion mutation, causes reduction of its protein rate and associated with sperm motility defect in infertile men. J Assist Reprod Genet 2015;32:807-15.
14. Dong Z, Qian J, Law TSM, et al. Mate-pair genome sequencing reveals structural variants for idiopathic male infertility. Hum Genet 2023;142:363-77.