Whole exome sequencing identifies genes associated with non-obstructive azoospermia

Abstract

Background: Genetic etiology is the main cause of non-obstructive azoospermia, but little is known about the landscape of the disease causative genes.

Objective: To identify the association of non-obstructive azoospermia and the putative causative genetic factors.

Design, *setting*, *and participants*: A single-center perspective case-control study of 133 patients, with clinicopathologic non-obstructive azoospermia and 495 fertile men control was performed. Eleven trio families were available and enrolled from the 133 patients' families.

Outcome measurements and statistical analysis: Whole exome sequencing based rare variant association study between the cases and controls was performed by means of gene burden association testing. Linkage analysis on the trio family was also described to screen the causative genes.

Results and limitations: Totally 80 genes (p < 0.05) were identified associated with non-obstructive azoospermia (2 of which were previously reported), meanwhile 5 novel genes out of which were also found potentially causative through the linkage analysis on the trio families. The pathway enrichment analysis was also provided to assess the potential interaction between genes identified in this study and previously reported together. The 5 novel identified overlap genes by both above mentioned test with the rare mutations account for an overall 20% (26 /133 patients) incidence, together with the 2 known genes together would account for an overall 20% incidence for non-obstructive azoospermia in this study. The study is limited by the lack of functional biological study.

Conclusions: Five novel genes were identified associated with non-obstructive azoospermia by means of both rare variant association study and linkage analysis through trio families. They could account for about 20% clinical incidence among the patients in our study.

Patient summary: 133 infertile patients (11 of them with parents enrolled) with idiopathic non-obstructive azoospermia and 300 fertile male controls were recruited from single clinic center.

All patients underwent semen analyses at least on three different occasions.

Keywords

non-obstructive azoospermia; rare variant association study; linkage analysis; causative genes; incidence

1. Introduction

Approximately 7% of male population worldwide suffer from infertility, which is mainly caused by azoospermia, a condition with the highest frequency of known genetic factors contributing to male infertility (about 25%). Given that nearly 50% of infertility cases are associated with genetic defects, the genetic etiology for azoospermia is still very elusive, the genetic basis of azoospermia in most cases remains unknown. In terms of pathobiological mechanism, azoospermia can be categorized into two groups: non-obstructive azoospermia which is defined as no sperm in the ejaculate due to failure of spermatogenesis, and obstructive azoospermia which is defined as the absence of spermatozoa in the ejaculate despite of the normal spermatogenesis.

Non-obstructive azoospermia (NOA) is the most severe form of azoospermia and lacking of effective cures, whereas patients with obstructive azoospermia usually can be treated with assisted reproductive technologies (ART) surgery and bypass infertility in this way. To date, mutations on several genes have been identified causative for, or associated with NOA, the acknowledged genes include X chromosome linked genes TEX11,⁴⁻⁷ FAM47C,⁸ MAGEB4,⁹ and AR;^{10,11} and autosomal genes PIWIL1,¹²⁻¹⁵ BRCA2,¹⁶ BOLL and DAZL,^{17,18} TEX15,¹⁹⁻²¹ SYCP3,²² NR5A1,²³ SOHLH1,²⁴ WT1,^{25,26} TAF4B and ZMYND15,²⁷ FANCM,²⁸ SPINK2,²⁹ TEX14,³⁰ DNAH6,³⁰ MEIOB,³¹ RNF212 and STAG3,³² XRCC2,^{33,34} TDRD7,³⁵ TDRD9,³⁶ DMC1,³⁷ SYCE1,³⁸ NPAS2,^{39,40} MCM8,⁴¹ MEI1,^{42,43} and STX2.⁴⁴ A surprisingly large proportion of total genes are specifically expressed in the male testis,⁴⁵ suggesting that thousands of genes may involve in the spermatogenic process. Therefore, the physiological pathway complexity may contribute to the large number of unresolved idiopathic NOA patients.

However, currently tools for understanding the genetic etiology of NOA is very limited due to the lack of large pedigrees resulting from the infertility. In this study we performed whole exome sequencing (WES), a powerful tool for identifying the pathogenic factors for inherent diseases, on 133 unrelated patients and 300 independent fertile male controls for rare variant association

study (RVAS), a kind of modified genome wide association study (GWAS);⁴⁸⁻⁵⁵ meanwhile 11 trio families (the proband with his parents) were also analyzed in a maternal inherited or a *de novo* pattern to identify the causative genes. Meanwhile proteins interaction pathway enrichment analysis was also performed to provide potential evidence to further address the abovementioned question.

2. Materials and methods

2.1. Study design and patients

This study was approved by the Medical Ethics Review Board of The First Hospital affiliated to Jilin University (Changchun, China), and with written informed consent from the patients and their parents, as well as the controls. 133 infertile patients (11 of them with their healthy parents enrolled together as trios) with idiopathic NOA (no sperm in the ejaculate even after centrifugation) and 300 fertile male controls were recruited from the Reproductive Medicine & Prenatal Diagnosis Center of The First Hospital, Jilin University (Changchun, China). All patients underwent semen analysis at least on three different occasions, and were excluded if be with a history of orchitis, obstruction of vas deferens or endocrine disorders. All the men from control group were fertile and had fathered at least one child.

We selected those affected by idiopathic NOA; 133 subjects and 11 trios signed the informed consent for this study and were enrolled. These patients were selected according to the following criteria: (i) azoospermia due to either spermatogenic arrest or Sertoli cell-only syndrome (SCOS; type I or II); (ii) normal karyotype; (iii) without absence of Y-chromosome microdeletions and TEX11 mutations (including the copy number variation (CNV) of TEX11); and (iv) absence of a list (Supplementary Table 1) of known causes of azoospermia. Clinical characteristics of the patients are reported in Table I.

2.2. Genomic analysis

The genomic DNA was prepared from the peripheral blood samples. All patients and the trio families enrolled in the study underwent WES sequencing performed by BGI Genomics (Shenzhen, China) according to previous report. Sequencing variants were annotated using ANNOVAR sequences as loss of function (LOF) or missense variants. LOFs were defined as frameshift insertions or deletions, gain or loss of stop codon, insertions, deletions, and disruption of canonical splice site dinucleotides. CNVs were analyzed with ExomeDepth and CNVkit. Pathway enrichment

analysis were performed with Cytoscape_v3.7.2 software based on the String database (https://string-db.org/). The data deposited in the CNSA (https://db.cngb.org/cnsa/) of CNGBdb with accession number CNP0001080.

2.3. Statistical analysis

Only LOFs with minor allele frequencies (MAF) \leq 0.1% as per the Genome Aggregation Database (gnomAD) were considered for inclusion in the gene burden association testing, which was performed with SKAT. The p values reported are two-tailed. When applicable, multiple testing correction was performed using the FDR method and FDR <0.05 was considered significant. Quantile-quantile plot (QQ plot) was achieved with the RVAS data and draw in R program. All analyses were performed using R v3.6.1 (www.R-project.org) and Bioconductor v3.4.

3. Results

The test flow for this study is shown in Figure 1, a total of 1021 NOA patients were enrolled in, of which 299 cases with abnormal karyotyping results were excluded, followed by 281 cases exclusion due to the AZF region microdeletion; after this, 305 cases were reluctant to accept further test and were excluded in the following study; then a list (Supplementary Table 1) of known genes of azoospermia were developed as custom genes panel screening test, 3 cases with positive results (2 with TEX11 mutation, 1 with AR mutation) were excluded in this step; eventually, 133 NOA patients without abnormal karyotyping, AZF region microdeletion, and the known causative genetic mutation were selected to perform WES; meanwhile 300 independent fertile men were selected as the control population in this study; among the 133 NOA cases, 11 were enrolled with their parents together as the trio family, linkage analysis for these 11 Trios were also performed to screen the causative genes.

RVAS were analyzed for these cases and controls, the testis expressed genes with p value lower than 0.05 were labelled in color blue; meanwhile the genes in blue color were changed into yellow color if the genes were reported previously associated with NOA trait, or changed into red if the genes were overlapped identified in the linkage analysis for trios family. The quartile-quartile plot for the RVAS analysis was also shown in Figure 2. Two previously reported genes, FAM47C and BRAC2 (Table 2), were identified associated with NOA in this study; and five novel genes were

revealed in both the RVAS (Table S2) and trios linkage analysis (Table 3) study.

The RVAS identified associated genes, as well as the previously reported genes, were analyzed for the pathway enrichment in Figure 3. Previously reported genes were labelled in green color, interestingly, previously reported genes associated with NOA, except MAGEB4, FAM47C, SPINK2 and STX2, tend to accumulate in the meiosis related protein interaction pathway which is centered around PIWIL1, TEX11, TEX15, DAZL, TEX14, etc. A part of the RVAS identified genes also involved in the enriched pathway and provided functional potentials to be associated with NOA; meanwhile a small part of genes, like HLA-A, tend to be enriched in the immunity pathway, however no previous reports or linkage analysis evidence can support the association between the immunity pathway we identified through RVAS analysis and NOA trait. Since FAM47C was identified as highly associated with NOA previously, the FAM47C mutation (frameshift mutation similar with previous report) from patient 58 in our study was also identified as the genetic etiology accordingly.

Trios family WES data were analyzed and summarized in Table 3 according to a maternal inherited pattern or the *de novo* mutation mechanism, paternal inherited variations were excluded, also the probands' variations were preserved in the summary table 3 only if the variations were not found in the 300 male control population. All of the WES identified variations were validated with sanger sequencing in this Table 3.

4. Discussion

Combining the data of RVAS analysis and the trio family analysis show the overlapped genes may associated with NOA. Especially, genes also involved in the enriched pathway with previously reported genes, suggesting a high potential to be associated with and causative for NOA. The previous reported genes, FAM47C and BRCA2, were also identified in this study, the FAM47C mutation (frameshift mutation similar with previous report) from patient 58 and the BRCA2 mutation (insertion similar with previous report) from patient 102 in our study was also identified as the genetic etiology accordingly. Nevertheless, the limitation of this study lies in that prospectively more evidence like the functional cellular or physiological experiments could be needed to address this question better, and that the missense mutation which may be pathogenic was not taken into account in this study.

The variations can be classified as Likely Pathogenic according to ACMG guidelines, ⁵⁸ PS2 (strong evidence) and PS4 (strong evidence) and PM2 (moderate evidence) and PM4 (moderate evidence).

5. Conclusion

We demonstrate the association study between the 133 NOA patients and 300 fertile men control, as well trio families were also analyzed to further investigate the causative genes for NOA trait. We conclude that the overlapped identified genes and may associated with NOA. The other RVAS or trio analysis identified genes, especially the genes involved in the enriched pathway may also associated with NOA. This study provides a potential possibility for better diagnosis in the future clinical application.

References

- Lotti, F. & Maggi, M. Ultrasound of the male genital tract in relation to male reproductive health. *Hum Reprod Update* **21**, 56-83, doi:10.1093/humupd/dmu042 (2015).
- 2 Krausz, C. & Riera-Escamilla, A. Genetics of male infertility. *Nat Rev Urol* 15, 369-384, doi:10.1038/s41585-018-0003-3 (2018).
- Rodrigues, V. O., Polisseni, F., Pannain, G. D. & Carvalho, M. A. G. Genetics in human reproduction. *JBRA Assist Reprod*, doi:10.5935/1518-0557.20200007 (2020).
- Boroujeni, P. B. *et al.* Expression analysis of genes encoding TEX11, TEX12, TEX14 and TEX15 in testis tissues of men with non-obstructive azoospermia. *JBRA Assist Reprod* **22**, 185-192, doi:10.5935/1518-0557.20180030 (2018).
- 5 Yatsenko, A. N. *et al.* X-linked TEX11 mutations, meiotic arrest, and azoospermia in infertile men. *N Engl J Med* **372**, 2097-2107, doi:10.1056/NEJMoa1406192 (2015).

- Scovell, J. M. & Ramasamy, R. Words of Wisdom. Re: X-Linked TEX11 Mutations,

 Meiotic Arrest, and Azoospermia in Infertile Men. *Eur Urol* **68**, 1101,

 doi:10.1016/j.eururo.2015.08.050 (2015).
- Niederberger, C. Re: X-Linked TEX11 Mutations, Meiotic Arrest, and Azoospermia in Infertile Men. *J Urol* **195**, 444, doi:10.1016/j.juro.2015.10.050 (2016).
- Chertman, W., Arora, H., Griswold, A. J., Lopategui, D. M. & Ramasamy, R. Whole Exome Sequencing Identifies a Rare Nonsense Mutation in FAM47C as a Possible Cause of Severe Oligospermia in Brothers With Varicocele. *Urology* 129, 71-73, doi:10.1016/j.urology.2019.03.016 (2019).
- Okutman, O. *et al.* A no-stop mutation in MAGEB4 is a possible cause of rare X-linked azoospermia and oligozoospermia in a consanguineous Turkish family. *J Assist Reprod Genet* **34**, 683-694, doi:10.1007/s10815-017-0900-z (2017).
- Wang, Q. *et al.* Azoospermia associated with a mutation in the ligand-binding domain of an androgen receptor displaying normal ligand binding, but defective trans-activation. *J Clin Endocrinol Metab* **83**, 4303-4309, doi:10.1210/jcem.83.12.5358 (1998).
- Mou, L. & Gui, Y. A novel variant of androgen receptor is associated with idiopathic azoospermia. *Mol Med Rep* **14**, 2915-2920, doi:10.3892/mmr.2016.5587 (2016).
- 12 Kamaliyan, Z., Pouriamanesh, S., Soosanabadi, M., Gholami, M. & Mirfakhraie, R. Investigation of piwi-interacting RNA pathway genes role in idiopathic non-obstructive azoospermia. *Sci Rep* **8**, 142, doi:10.1038/s41598-017-17518-4 (2018).
- 13 Hasuwa, H., Ishino, K. & Siomi, H. Human PIWI (HIWI) is an azoospermia factor. Sci

- China Life Sci 61, 348-350, doi:10.1007/s11427-017-9149-0 (2018).
- Gou, L. T. *et al.* Ubiquitination-Deficient Mutations in Human Piwi Cause Male Infertility by Impairing Histone-to-Protamine Exchange during Spermiogenesis. *Cell* **169**, 1090-1104 e1013, doi:10.1016/j.cell.2017.04.034 (2017).
- Gu, A. *et al.* Genetic variants in Piwi-interacting RNA pathway genes confer susceptibility to spermatogenic failure in a Chinese population. *Hum Reprod* **25**, 2955-2961, doi:10.1093/humrep/deq274 (2010).
- Zhoucun, A. *et al.* The common variant N372H in BRCA2 gene may be associated with idiopathic male infertility with azoospermia or severe oligozoospermia. *Eur J Obstet Gynecol Reprod Biol* **124**, 61-64, doi:10.1016/j.ejogrb.2005.09.001 (2006).
- Zagore, L. L. *et al.* DAZL Regulates Germ Cell Survival through a Network of PolyA-Proximal mRNA Interactions. *Cell Rep* **25**, 1225-1240 e1226, doi:10.1016/j.celrep.2018.10.012 (2018).
- Zhang, C. *et al.* Highly conserved epigenetic regulation of BOULE and DAZL is associated with human fertility. *FASEB J* **30**, 3424-3440, doi:10.1096/fj.201500167R (2016).
- Colombo, R., Pontoglio, A. & Bini, M. Two Novel TEX15 Mutations in a Family with Nonobstructive Azoospermia. *Gynecol Obstet Invest* 82, 283-286, doi:10.1159/000468934 (2017).
- Wang, X. et al. Case study of a patient with cryptozoospermia associated with a recessive TEX15 nonsense mutation. Asian J Androl 20, 101-102, doi:10.4103/1008-682X.194998 (2018).

- Okutman, O. *et al.* Exome sequencing reveals a nonsense mutation in TEX15 causing spermatogenic failure in a Turkish family. *Hum Mol Genet* **24**, 5581-5588, doi:10.1093/hmg/ddv290 (2015).
- 22 Miyamoto, T. *et al.* Azoospermia in patients heterozygous for a mutation in SYCP3.
 Lancet 362, 1714-1719, doi:10.1016/S0140-6736(03)14845-3 (2003).
- Safari, S. *et al.* An Iranian family with azoospermia and premature ovarian insufficiency segregating NR5A1 mutation. *Climacteric* **17**, 301-303, doi:10.3109/13697137.2013.847079 (2014).
- 24 Choi, Y. *et al.* Mutations in SOHLH1 gene associate with nonobstructive azoospermia. *Hum Mutat* **31**, 788-793, doi:10.1002/humu.21264 (2010).
- Wang, X. N. *et al.* The Wilms tumor gene, Wt1, is critical for mouse spermatogenesis via regulation of sertoli cell polarity and is associated with non-obstructive azoospermia in humans. *PLoS Genet* **9**, e1003645, doi:10.1371/journal.pgen.1003645 (2013).
- Xu, J. *et al.* A novel functional variant in Wilms' Tumor 1 (WT1) is associated with idiopathic non-obstructive azoospermia. *Mol Reprod Dev* **84**, 222-228, doi:10.1002/mrd.22768 (2017).
- Ayhan, O. *et al.* Truncating mutations in TAF4B and ZMYND15 causing recessive azoospermia. *J Med Genet* **51**, 239-244, doi:10.1136/jmedgenet-2013-102102 (2014).
- 28 Kasak, L. *et al.* Bi-allelic Recessive Loss-of-Function Variants in FANCM Cause Non-obstructive Azoospermia. *Am J Hum Genet* **103**, 200-212, doi:10.1016/j.ajhg.2018.07.005 (2018).

- 29 Kherraf, Z. E. *et al.* SPINK2 deficiency causes infertility by inducing sperm defects in heterozygotes and azoospermia in homozygotes. *EMBO Mol Med* **9**, 1132-1149, doi:10.15252/emmm.201607461 (2017).
- Gershoni, M. *et al.* A familial study of azoospermic men identifies three novel causative mutations in three new human azoospermia genes. *Genet Med* **19**, 998-1006, doi:10.1038/gim.2016.225 (2017).
- Gershoni, M. *et al.* A new MEIOB mutation is a recurrent cause for azoospermia and testicular meiotic arrest. *Hum Reprod* **34**, 666-671, doi:10.1093/humrep/dez016 (2019).
- Riera-Escamilla, A. *et al.* Sequencing of a 'mouse azoospermia' gene panel in azoospermic men: identification of RNF212 and STAG3 mutations as novel genetic causes of meiotic arrest. *Hum Reprod* **34**, 978-988, doi:10.1093/humrep/dez042 (2019).
- Yang, Y. *et al.* XRCC2 mutation causes meiotic arrest, azoospermia and infertility. *J Med Genet* **55**, 628-636, doi:10.1136/jmedgenet-2017-105145 (2018).
- Zhang, Y. X. *et al.* XRCC2 mutation causes premature ovarian insufficiency as well as non-obstructive azoospermia in humans. *Clin Genet* **95**, 442-443, doi:10.1111/cge.13475 (2019).
- Tan, Y. Q. *et al.* Loss-of-function mutations in TDRD7 lead to a rare novel syndrome combining congenital cataract and nonobstructive azoospermia in humans. *Genet Med* 21, 1209-1217, doi:10.1038/gim.2017.130 (2019).
- 36 Arafat, M. et al. Mutation in TDRD9 causes non-obstructive azoospermia in infertile

- men. J Med Genet 54, 633-639, doi:10.1136/jmedgenet-2017-104514 (2017).
- 37 He, W. B. *et al.* DMC1 mutation that causes human non-obstructive azoospermia and premature ovarian insufficiency identified by whole-exome sequencing. *J Med Genet* **55**, 198-204, doi:10.1136/jmedgenet-2017-104992 (2018).
- Maor-Sagie, E. *et al.* Deleterious mutation in SYCE1 is associated with non-obstructive azoospermia. *J Assist Reprod Genet* **32**, 887-891, doi:10.1007/s10815-015-0445-y (2015).
- Wang, Y. Y. *et al.* Loss of SLC9A3 decreases CFTR protein and causes obstructed azoospermia in mice. *PLoS Genet* **13**, e1006715, doi:10.1371/journal.pgen.1006715 (2017).
- Ramasamy, R. *et al.* Whole-exome sequencing identifies novel homozygous mutation in NPAS2 in family with nonobstructive azoospermia. *Fertil Steril* **104**, 286-291, doi:10.1016/j.fertnstert.2015.04.001 (2015).
- Tenenbaum-Rakover, Y. *et al.* Minichromosome maintenance complex component 8 (MCM8) gene mutations result in primary gonadal failure. *J Med Genet* **52**, 391-399, doi:10.1136/jmedgenet-2014-102921 (2015).
- 42 Ghieh, F., Mitchell, V., Mandon-Pepin, B. & Vialard, F. Genetic defects in human azoospermia. *Basic Clin Androl* **29**, 4, doi:10.1186/s12610-019-0086-6 (2019).
- Ben Khelifa, M. *et al.* A MEI1 homozygous missense mutation associated with meiotic arrest in a consanguineous family. *Hum Reprod* **33**, 1034-1037, doi:10.1093/humrep/dey073 (2018).
- 44 Nakamura, S. et al. STX2 is a causative gene for nonobstructive azoospermia. Hum

- Mutat 39, 830-833, doi:10.1002/humu.23423 (2018).
- Schultz, N., Hamra, F. K. & Garbers, D. L. A multitude of genes expressed solely in meiotic or postmeiotic spermatogenic cells offers a myriad of contraceptive targets.

 *Proc Natl Acad Sci U S A 100, 12201-12206, doi:10.1073/pnas.1635054100 (2003).
- Locke, A. E. *et al.* Exome sequencing of Finnish isolates enhances rare-variant association power. *Nature* **572**, 323-328, doi:10.1038/s41586-019-1457-z (2019).
- Sanders, S. J. *et al.* De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **485**, 237-241, doi:10.1038/nature10945 (2012).
- Ahluwalia, T. S. *et al.* A novel rare CUBN variant and three additional genes identified in Europeans with and without diabetes: results from an exome-wide association study of albuminuria. *Diabetologia* **62**, 292-305, doi:10.1007/s00125-018-4783-z (2019).
- He, Z. *et al.* Rare-variant association tests in longitudinal studies, with an application to the Multi-Ethnic Study of Atherosclerosis (MESA). *Genet Epidemiol* **41**, 801-810, doi:10.1002/gepi.22081 (2017).
- Artomov, M. *et al.* Rare Variant, Gene-Based Association Study of Hereditary Melanoma Using Whole-Exome Sequencing. *J Natl Cancer Inst* **109**, doi:10.1093/jnci/djx083 (2017).
- Aung, T. *et al.* Genetic association study of exfoliation syndrome identifies a protective rare variant at LOXL1 and five new susceptibility loci. *Nat Genet* **49**, 993-1004, doi:10.1038/ng.3875 (2017).
- 52 Tada, H., Kawashiri, M. A., Konno, T., Yamagishi, M. & Hayashi, K. Common and

- Rare Variant Association Study for Plasma Lipids and Coronary Artery Disease. *J Atheroscler Thromb* **23**, 241-256, doi:10.5551/jat.31393 (2016).
- Lin, W. Y. Beyond Rare-Variant Association Testing: Pinpointing Rare Causal Variants in Case-Control Sequencing Study. Sci Rep 6, 21824, doi:10.1038/srep21824 (2016).
- Lee, S., Abecasis, G. R., Boehnke, M. & Lin, X. Rare-variant association analysis: study designs and statistical tests. *Am J Hum Genet* **95**, 5-23, doi:10.1016/j.ajhg.2014.06.009 (2014).
- He, L. & Pitkaniemi, J. M. Family-based Bayesian collapsing method for rare-variant association study. *BMC Proc* **8**, S37, doi:10.1186/1753-6561-8-S1-S37 (2014).
- Liu, G. *et al.* Novel missense variant in TTN cosegregating with familial atrioventricular block. *Eur J Med Genet* **63**, 103752, doi:10.1016/j.ejmg.2019.103752 (2020).
- Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164, doi:10.1093/nar/gkq603 (2010).
- Richards, S. *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* 17, 405-424,
 doi:10.1038/gim.2015.30 (2015).



