Towards a Comprehensive Catalog of Human Genes Associated with Main Forms of Pathoszoopermia and its Functional Annotation

Elena V. Ignatieva , Alexander V. Osadchuk, Maxim A. Kleshev, Ludmila V. Osadchuk

Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

**e-mail: eignat@bionet.nsc.ru; osadchuk@bionet.nsc.ru; max82cll@bionet.nsc.ru; losadch@bionet.nsc.ru*

Motivation and Aim: **The genetic causes of the global decline in male fertility is among the hot spots of scientific research in reproductive genetics. The most common way to evaluate male fertility in clinical trials is to determine the quality of the ejaculate. A pathological condition characterized by lower quality indicators of the ejaculate compared to the norm and leading to impaired fertility is called pathospermia. Pathospermia is a syndrome that occurs in many diseases and can be caused by many factors, including genetic ones. To systematize information about genes associated with main forms of pathozoospermia, we created a catalog of such genes and analyzed their functional characteristics.**

Methods and Algorithms: **Data on associations between allelic variants of genes and specific forms of pathospermia were extracted from scientific publications that were selected using ANDSystem (http://www-bionet.sscc.ru/and/cell/). Search terms that refer to specific forms of pathozoospermia were obtained from the 2010 WHO guidelines. Gene ontology (GO) enrichment analysis was performed using DAVID tool (https://david.ncifcrf.gov/). The enriched GO terms (FDR<0.05) from the biological processes (GOTERM_BP_5) and cellular compartments (GOTERM_CC_5) vocabularies were considered in our study.**

Results: Male infertility genes (MIGenes) – a catalog of genes associated with pathozoospermia

1) A dictionary of keywords denoting specific forms of pathozoospermia was formed:

1) non-obstructive azoospermia (NOA);
2) cryptozoospermia;
3) oligozoospermia;
4) severe oligozoospermia;
5) asthenozoospermia;
6) teratozoospermia;
7) oligoasthenoteratozoospermia (OAT);
8) oligoasthenozoospermia;
9) oligoteratozoospermia;
10) globozoospermia

2) Using these key words as a search terms we created a catalog of genes, associated with pathozoospermia.

ANDSystem + manual verification

Data was extracted from experimental research papers

MIGenes:
88 human genes
> 150 polymorphic loci associated with pathozoospermia

Each entry of MIGenes contain data on the gene, polymorphic locus, form of pathozoospermia, population, and scientific publication.

GeneID	Official Symbol	Official Full Name	Map location	Infertility phenotype	Reference	Population	SNP	Sentences
1761	*DMRT1*	doublesex and mab-3 related transcription factor 1	9p24.3	Non-obstructive azoospermia	PMID 2019: 31479588	Brazilian	rs140506267 (Asn224Ser)	We identified…
190	*NR0B1*	nuclear receptor subfamily 0 group B member 1	Xp21.2	Oligozoospermia	PMID 2013: 23384712	French	rs1569269179 (Trp39Ter)	A man with a DAX1/NR0B1 mutation….

Results: MIGenes - information content and functional annotation

Most genes (81 out of 88) are located on autosomes and are associated with such forms of pathozoospermia as nonobstructive azoospermia, oligozoospermia and asthenozoospermia.

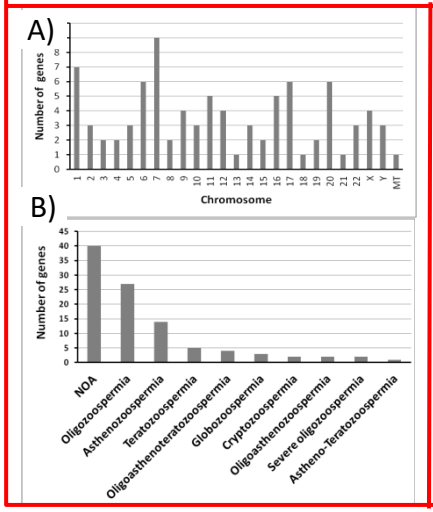

Figure 1. Information content of the MIGenes catalog. Panel A - chromosomal localization of genes; panel B - the number of genes associated with a specific form of pathozoospermia. *NOA* - non-obstructive azoospermia.

Enrichment analysis performed using the DAVID-tool revealed six significantly enriched terms denoting cellular localization. They all indicate localization on the chromosome. Among the GO terms that indicate biological processes, we identified nine key significantly enriched terms. These terms were associated with spermatogenesis, gamete generation, DNA metabolic process, cell division, cytoskeletal function.

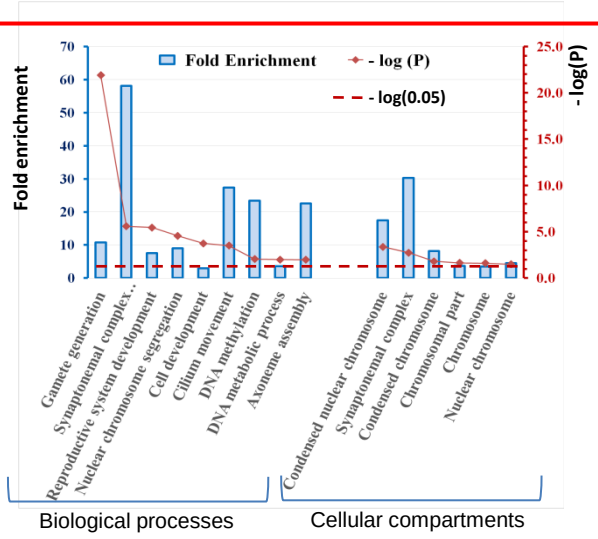

Figure 2. Association of genes from the catalog with GO terms from biological processes and cellular compartments vocabularies. GO terms with FDR < 0.05 are presented.

Conclusion: We have created a catalog of genes and their allelic variants associated with various forms of pathozoospermia. This information was obtained by using ANDSystem in combination with subsequent manual verification, which indicates a high level of reliability. The data we have accumulated and the results of functional gene annotation can be useful both for interpreting the results of genome- and exome-wide association studies, as well as for developing new approaches to the diagnosis and treatment of male infertility.

Acknowledgements: This study was supported by the grant from the Russian Science Foundation (project no. 19-15-00075)