ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Gene Expression Profiles of Spermatozoa in Teratozoospermia.

Soudabeh Sabetian*, Mohd Shahir Shamsir, Mohammed Abu Naser, and Bashir Sajo Mienda.

Department of Biological and Health Sciences, Faculty of Bioscience & Medical Engineering, Universiti Teknologi Malaysia, 81310 Johor, Malaysia.

ABSTRACT

Teratospermia or teratozoospermia (TS) is a condition characterized by the presence of sperm with abnormal morphology that disturbs fertility in males. The exact causes of this condition are still unknown. This study aimed to analyze the differentially expressed genes (DEGs) in sperm having TS and identify the featured genes related to TS. The gene expression profile GSE6872 was downloaded from Gene Expression Omnibus, including 8 TS and 13 control samples. The differentially expressed genes were identified by testing students consisting of 165 upregulated and 377 downregulated DEGs. Then, a Reactome functional interaction network for DEGs using Cytoscape was constructed and the hub DEGs nodes were identified. These DEGs were then mapped to GO and KEGG databases and important pathways associated with the AD were isolated. The GO ontology and KEGG pathway analysis of high degree DEGs showed that cytosolic ribosome processes such as RPS4X and mitochondrial enzyme activities including NDUFB10 were significantly enriched (adjusted P value <. 05). These findings suggested that the genes having high degree in Reactome functional interaction network for DEGs may be the underlying potential therapeutic targets for TS.

Keywords: Teratospermia, differentially expressed genes, functional interaction network

Corresponding author

ISSN: 0975-8585

INTRODUCTION

Teratospermia or teratozoospermia is caused by the presence of sperm having an abnormal morphology and low motility that gives rise to infertility. For treatment of this condition, a process of intracytoplasmic sperm injection (ICSI) is frequently utilized [1]. In most cases, the causes of teratospermia cannot be clearly identified although diseases like Hodgkin's, coeliac and Crohn's may contribute in a few cases [2, 3]. (http://www.gfmer.ch/Endo/Lectures_09/dupan1.htm).

Transformation of numerous ribosomal proteins have been found recently to be a main cause of genetic diseases, many of which interfere with the ribosome biogenesis [4]. It is also known that defects in ribosome synthesis elicit cell cycle arrest and apoptosis [5, 6]. Thus, it is probable that the causes of infertility or subfertility phenotypes are due to insufficiencies in particular ribosomal proteins [7]. In humans, a quantitative reduction synthesis of ribosomal protein s4, which is linked to the X-chromosome has been observed in individuals, which include short stature and infertility [8].

It has been found that there is a clear, direct and positive correlation in the whole population between mitochondrial enzyme-specific activities and spermatozoa motility [9]. This is evident from the results of an earlier report that the spermatozoa mitochondria of the patients having primary mitochondrial disorders show the characteristic ultra-structural features of mitochondrial disorders, which reduce the mobility of sperm [10]. The fertility of males is highly affected due to alterations at the molecular level [11]. The study aimed to identify the proteins that are over-or under expressed in the sperm of males suffering from teratospermia. The distinctive occurrence of some of the proteins provides an insight into the mechanistic role played by these proteins in the infertility of males and may serve as potential biomarkers.

MATERIAL AND METHODS

Affymetrix Microarray Data

The microarray data was downloaded from the Gene Expression Omnibus (GEO) database [12], accession number GSE6872, including 8 teratozoospermia (TS) and 13 control (CT) samples [13]. All analytical source samples were taken from human sperm. The microarray expression platform was Affymetrix Human Genome U133 plus 2.0.

Data Preprocessing

From the CEL files, the data was transformed into an expression profile and background correction, and quartile data normalization was performed by robust multiarray average [14] with affy package. Average gene expression values for genes which correspond to multiple probe sets which have a plurality of expression values were calculated. Eventually, the expression profiles of 54675 genes were obtained from 21 specimens.

Differentially Expressed Genes Analysis

The identification of the genes that were significantly differentially expressed between 8 TS and 13 CT samples was carried out by student t test. The P values that were less than 0.01 were adjusted by Benjamin and Hochberg (BH) method [15] and fold change (FC) value larger than 2 was set as the threshold criteria for screening DEGs [16].

Gene Ontology Enrichment Analysis

A bioinformatics project developed by the GO Consortium that aims to introduce consistency in the description of functional information pertaining to gene products is known as Gene Ontology (GO) [17]. The process of GO consists of 3 ontologies that are used to describe the molecular functions, cellular component of a gene product and biological processes [18]. The Database for Annotation, Visualization and Integrated Discovery (DAVID) consists of an integrated biological knowledgebase and functional annotation chart or table. It provides a comprehensive set of functional annotation tools for investigators to integrate the significant genes of a specific function [19]. The EASE method [20] was used to calculate significant values for GO

categories after all DEGs were input into DAVID. The P value adjusted by BH less than .05 and EASE score of 0.01 were chosen as a cutoff criteria.

Kyoto Encyclopedia of Genes and Genomes Pathway Analysis

The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database (http://www.kegg.jp/) consists of graphical diagrams of biochemical pathways which include most of the known metabolic pathways and some known regulatory pathways [21]. These DEGs were then mapped to this database and important pathways associated with the AD were isolated. The P value adjusted by BH less than 0.05 and EASE score of 0.01 were chosen as a cutoff criteria.

Construction of Reactome functional interaction network

The DFGs were then loaded into a Cytoscape 2.8.3 [22] using the Reactome Cytoscape Plugin. The Reactome pathways stored in the database are accessed by this plugin, which ease the process of pathway enrichment analysis for a set of genes, investigates the functional relationships among genes in hit pathways and visualizes hit pathways using manually laid-out pathway diagrams directly in a Cytoscape [23]. The degree of the DFGs was evaluated and defined, one of which is the hub (high degree) node in this network.

RESULTS

Screening of DEGs

A publicly available microarray data set GSE6872 was obtained from the GEO database. Student t test was carried out to analyze the DEGs between the 8 TS and 13 CT samples. According to the threshold criteria (adjusted P value < .01 and FC value > 2) for DEGs, in total, 542 DEGs were identified including 165 up regulated and 377 down regulated DEGs.

Gene Ontology Enrichment Analysis of DEGs

To investigate a TS gene expression on a functional level, DEGs (adjusted P value < 0.01 and FC value > 2) between TS and CT samples were significantly enriched in the 10 GO terms (P value < .01) that are shown in Table 1. There were 6 reproductive process including spermatogenesis, male gamete generation, sexual reproduction, gamete generation, multicellular organism reproduction and reproductive process in a multicellular organism. The contributed genes in each biological process are shown in supplementary table 1.

Table 1: Gene Ontology Terms with Adjusted P Value < .05.a

Terms	Biological Process	P Value
GO:0006414	translational elongation	4.50E-41
GO:0006412	translation	5.11E-27
GO:0019953	sexual reproduction	7.91E-7
GO:0007283	spermatogenesis	2.62E-6
GO:0048232	male gamete generation	2.62E-6
GO:0007276	gamete generation	4.26E-5
GO:0042273	ribosomal large subunit biogenesis	5.16E-5
GO:0048609	reproductive process in a multicellular organism	1.54E-4
GO:0032504	multicellular organism reproduction	1.546E-4
GO:0007017	microtubule-based process	7.72E-4

Abbreviations: GO, Gene Ontology; ATP, adenosine triphosphate; BH, Benjamin and Hochberg.
a Adjusted P value means P value adjusted by BH.

Kyoto Encyclopedia of Genes and Genomes Pathway Analysis and Network Construction

The 542 DEGs were mapped into KEGG pathway database to screen the enrichment pathways and 5 significant pathways (adjusted P value < 0.01, Table 2) were obtained. Most of the DEGs were involved in a ribosome pathway. The genes included in the five pathways are represented in supplementary table 2.

ISSN: 0975-8585

Table 2: Kyoto Encyclopedia of Genes and Genomes Pathways of Differentially Expressed Genes.a,b

Terms	Description	Count	Adjusted P Value
hsa03010	Ribosome	39	2.52E-37
hsa05016	Huntington's disease	10	0.07
hsa03022	Basal transcription factors	4	0.08
hsa05130	Pathogenic Escherichia coli infection	5	0.08
hsa05010	Alzheimer's disease	9	0.09

Abbreviation: BH, Benjamin and Hochberg.
a Count means the number of candidate genes with that annotation.
b Adjusted P value means P value adjusted by BH.

Construction of Reactome functional interaction network

We have constructed a functional network to understand the significance and the role of the DEGs in different pathways of the human interaction network. The degrees of DEGs nodes were calculated and the high degree DEGs have been defined as hub nodes (supplementary table 3).

The GO terms of the higher degree DEGs are depicted in table 3. The Reactome functional network consists of 769 nodes and 26168 interaction edges (figure 1).

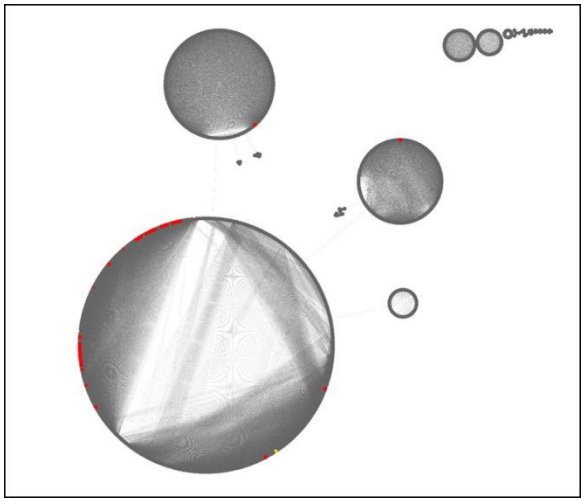

Figure 1: Reactome functional network: this network consists of 769 nodes and 26168 interaction edges. The red nodes indicated annotated downregulated differentially expressed genes (DEGs) and the yellow nodes indicated annotated upregulated DEGs.

Table 3: Gene Ontology Terms with Adjusted P Value < .05.a

Term	Biological Process	PValue
GO:0005840	ribosome	3.06E-59
GO:0022626	cytosolic ribosome	1.45E-56
GO:0033279	ribosomal subunit	7.97E-50
GO:0044445	cytosolic part	4.28E-45
GO:0030529	ribonucleoprotein complex	7.67E-45
GO:0005829	cytosol	2.00E-37
GO:0022627	cytosolic small ribosomal subunit	3.27E-26
GO:0022625	cytosolic large ribosomal subunit	3.36E-24
GO:0015935	small ribosomal subunit	5.01E-23
GO:0015934	large ribosomal subunit	1.30E-22
GO:0043228	non-membrane-bounded organelle	5.48E-21
GO:0043232	intracellular non-membrane-bounded organelle	5.48E-21

Abbreviations: GO, Gene Ontology; ATP, adenosine triphosphate; BH, Benjamin and Hochberg.
a Adjusted P value means P value adjusted by BH.

DISCUSSION

Teratozoospermia or teratospermia is an increase of abnormal human sperm. It is diagnosed when sperm with abnormal morphology exceed 85 percent and it causes infertility because most of the anomalous sperm are unable to move towards the egg and penetrate it. The exact causes of teratozoospermia are unknown. In this study we analyzed the differentially expressed genes (DEGs) in sperm having TS and identified the featured genes related to TS using GEO, KEGG, GO databases and protein interaction network approach. We have constructed a functional network to understand the significance and the role of the DEGs in different pathways of the human interaction network. The degrees of DEGs nodes were calculated and the high degree DEGs were defined as hub nodes. The hub protein nodes serve as a significant factor because they act as a potential drug targets. The function of the network can be shut down by inhibiting the functions of hubs by small molecules [24]. The highest percentage of the high degree DEGs was seen in cytosolic large and small ribosomal subunit biogenesis. It has been observed from a previous study that a protein translation in sperm involves mitochondrial rather than cytoplasmic ribosomes. These ribosomes play an active role in the translation of protein in spermatozoa. The sperm capacitation, motility, and in vitro fertilization rate is reduced significantly due to the inhibition of protein translation. Therefore, nuclear genes are expressed as proteins in sperm during their residence in the female reproductive tract until fertilization which is contrary to accepted dogma [25]. According to the latest research, chloramphenicol and the mitochondrial ribosome-specific inhibitor affect the acrosome reaction and capacitation, rather than penetration. In contrast, cyclohexide and cytoplasmic ribosome-specific inhibitor decreased the number of oocytes that reach metaphase II stage and are penetrated [26].

In humans, a quantitative reduction in the synthesis of ribosomal protein s4, which is linked to the X-chromosome, is observed in individuals, often resulting in short stature and infertility [8]. One of the high degree DEGs in our result (see supplementary table 3) is ribosomal protein s4 (RPS4X) that has been presented as a significant protein and is a down gene in the observed TS samples. This result shows that the expression level of this protein is related to fertility and the down expression may be associated with teratospermia. This outcome is in agreement with previous research.

The findings of this research suggest that the ribosomal compartment may play an essential role in disturbing the fertility outcome, which also tallies with our observations on the RNA level.

A previous report noted a close and positive relationship between sperm motility and mitochondrial enzyme-specific activities, suggesting that more specific mitochondrial dysfunctions could be the underlying cause of idiopathic asthenozoospermia [27].

NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 10 (NDUFB10) (see supplementary table 3) is a high degree DEGs in our functional network. It is thought that the protein expression level of this mitochondrial enzyme is related to fertility. The inner mitochondrial membrane (IMM) includes several complexes that make up the electron transfer chain (ETC). NDUFB10 belongs to the mitochondrial membrane respiratory chain NADH dehydrogenase (Complex I) [9]. Complex I functions in the transfer of electrons from NADH to the respiratory chain [28]. The expression of several mitochondrial proteins seems to be altered in sperm having low motility [29, 30] and the correlation between the activity of ETC enzymes and sperm parameters has been clearly shown [31]. The expression assessment of mitochondrial enzymes in the sperm samples of teratozoospermia individuals could help to identify the cases of mitochondrial-based teratozoospermia. Based on our analysis, the low expression of mitochondrial enzyme including NDUFB10 may cause teratozospermia.

CONCLUSION

The differentially expressed genes (DEGs) between teratospermia and control samples including 165 up regulated and 377 down regulated have been identified. The Reactome functional interaction network was constructed for DEGs and their degrees were calculated. The hub nodes of DEGs were presented as the significant nodes that are related to teratospermia. The GO ontology and KEGG pathway analysis of high degree DEGs showed that cytosolic ribosomes such as RPS4X, and mitochondrial enzymes including NDUFB10 are significantly enriched, which clarifies that any deficiency or mutation on these genes may cause teratospermia. The application of this tool may be an alternative to finding important genes and drug targets

in teratospermia and the distinct presence of some of the proteins may serve as potential biomarkers and provide insight into the mechanistic role played by these proteins in male infertility.

Authors' Contribution

The work was carried out in collaboration between all the authors. Mohd Shahir Shamsir defined the research theme. Soudabeh Sabetian and Mohammed Abu Naser designed methods and analyzed the data. Mohd Shahir Shamsir, Soudabeh Sabetian and Mohammed Abu Naser interpreted the results and co-wrote the paper.

ACKNOWLEDGMENT

The authors would like to acknowledge the Universiti Teknologi Malaysia Institutional Research Grant for the funding.

REFERENCES

[1] French DB, et al. Fertility Sterility 2010;93(4):1097-1103.
[2] O'Flaherty C. Iatrogenic Genetic Damage of Spermatozoa, in Genetic Damage in Human Spermatozoa. 2014, Springer. p. 117-135.
[3] Winters BR and TJ Walsh. Urologic Clinics of North America 2014;41(1):195-204.
[4] Liu JM and SR Ellis. Blood 2006;107(12):4583-4588.
[5] Dai MS, et al. Mol Cell Biol 2004;24(17):7654-7668.
[6] Lohrum MA, et al. Cancer Cell 2003;3(6): p. 577-587.
[7] Lopes AM, et al. BMC Mol Biol 2010;11(1):33.
[8] Miyamoto H and K Kajihara. The ribosomal protein S2 from the pacific oyster, Crassostrea gigas. Memoirs of Institute of Advanced Technology, Kinki University, 2005(10): p. 37-41.
[9] Amaral A, et al. Reprod 2013;146(5):R163-R174.
[10] Pelliccione F, et al. Fertil Steril 2011;95(2): 641-6.
[11] Sharma R, et al. Reprod. Biol Endocrinol 2013;11:38.
[12] Edgar R, M Domrachev and AE Lash. Nucleic Acids Res 2002;30(1):207-210.
[13] Platts AE, et al. Hum Mol Genet 2007;16(7):763-73.
[14] Irizarry RA, et al. Biostatistics 2003;4(2):249-264.
[15] Benjamini Y and Y Hochberg. J Royal Statistical Society. Series B (Methodological), 1995:289-300.
[16] Mutch DM, et al. BMC Bioinform 2002;3(1):17.
[17] Ashburner M, et al. Nat Gen 2000;25(1):25-29.
[18] Mutowo-Meullenet P, et al. Database: the Journal Of Biological Databases And Curation 2013.
[19] Dennis Jr G, et al. Genome Biol 2003;4(5):P3.
[20] Hosack DA, et al. Genome Biol 2003;4(10):R70.
[21] Altermann E and TR Klaenhammer. BMC Genom 2005;6(1):60.
[22] Shannon P, et al. Gen Res 2003;13(11):2498-2504.
[23] Croft, D, et al. Nucleic Acids Res 2011;39(suppl 1):D691-D697.
[24] Kim KK and HB Kim. World J Gastroenterol 2009;15(36):4518.
[25] Gur Y and H Breitbart. Gen Develop 2006;20(4): 411-416.
[26] Okudaira Y and H Funahashi. Reprod Fertility Develop 2014;26(1):186-187.
[27] Ruiz-Pesini E, et al. Clin Chem 1998;44(8):1616-1620.
[28] Murray J, et al. T. J Biol Chem 2003;278(16): 13619-13622.
[29] Siva AB, et al. Mol Human Reprod 2010;16(7): 452-462.
[30] Parte PP, et al. J Proteomics 2012;75(18): 5861-71.
[31] Ruiz-Pesini E, et al. Clinica Chimica Acta 2000; 300(1):97-105.