

# Seminal Plasma piRNA Array Analysis Identifies Possible Biomarker piRNAs for the Diagnosis of Asthenozoospermia

ling he

Department of Cardiovascular Medicine, The Second Affiliated Hospital of Nanchang

Xing wu

Assisted Reproductive Center, Jiangxi Maternal and Child Health Hospital

rong wu

Department of clinical medicine, JIANGXI HEALTH VOCATIONAL COLLEGE

ping guo

Obstetrics and Gynecology, Yichun Maternal and Child Health Hospital

Wan sun

Department of Clinical Laboratory, The Second Affiliated Hospital of Nanchang University, Nanchang

wen he

Jiangxi Key Laboratory of Molecular Medicine, The Second Affiliated Hospital of Nanchang University

he chen (✉ [181962170@qq.com](mailto:181962170@qq.com))

Jiangxi Key Laboratory of Molecular Medicine, The Second Affiliated Hospital of Nanchang University

---

## Research Article

**Keywords:** Seminal plasma, piRNA, Asthenozoospermia, analysis, biomarker

**Posted Date:** January 27th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-149605/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

## Seminal plasma piRNA array analysis identifies possible biomarker piRNAs for the diagnosis of asthenozoospermia

Ling He<sup>1#</sup>, Xingwu Wu<sup>2#</sup>, Rongye Wu<sup>3#</sup>, Ping Guo<sup>4</sup>, wenfeng He<sup>6</sup>, Wanlei Sun<sup>5</sup>, He Chen<sup>6\*</sup>

1. Department of Cardiovascular Medicine, The Second Affiliated Hospital of Nanchang University, Nanchang, 330006, China

2. Assisted Reproductive Center, Jiangxi Maternal and Child Health Hospital, Nanchang, 330006, China

3. Department of clinical medicine, JIANGXI HEALTH VOCATIONAL COLLEGE, Nanchang, 330052

4. Obstetrics and Gynecology , Yichun Maternal and Child Health Hospital, Yichun, 336000, China

5. Department of Clinical Laboratory, The Second Affiliated Hospital of Nanchang University, Nanchang, 330006, China

6. Jiangxi Key Laboratory of Molecular Medicine, The Second Affiliated Hospital of Nanchang University, Nanchang, 330006, China

Co-authors: Ling He, Xingwu Wu, and Rongye Wu.

CONTACT: He Chen, MD, Jiangxi Key Laboratory of Molecular Medicine, The Second Affiliated Hospital of Nanchang University, Nanchang, 330006, China. E-mail: chenhe2021@126.com

his work was funded by the The Provincial Natural Science Foundation of Jiangxi (20181BAB205024, 20181BAB205013), Project of Jiangxi Provincial Health Department (20181075, 20181084, 2017A100), Science and technology research project of Jiangxi Provincial Department of Education (20160165, 171353 ,170096,170014) and Graduate Innovation Fund Project of Jiangxi Province (YC2012-B011)

**Key words:** Seminal plasma, piRNA, Asthenozoospermia, analysis, biomarker,

### Abstract

Asthenozoospermia (AZS) is characterized by reduced sperm motility, and its pathogenesis remains poorly understood. Piwi-interacting RNAs (piRNAs) have been recognized to play important roles in spermatogenesis. However, little is known about the correlation of piRNAs with AZS. In this study, small RNA sequencing was performed on samples from AZS patients and fertile controls. Real-time quantitative reverse transcription polymerase chain reaction (qRT-PCR) was used to validate the small RNA-seq results. Bioinformatics analyses were performed to predict the functions of differentially expressed piRNAs. Logistic regression models were constructed, and receiver operating characteristic curve (ROC) analysis was used to evaluate their diagnostic performance. A total of 114 upregulated and 169 downregulated piRNAs were detected in AZS patients. GO and KEGG analyses showed that the differentially expressed piRNAs were mainly associated with transcription, signal transduction, cell differentiation, metal ion binding and focal adhesion. These results were verified by qRT-PCR of 8 selected piRNAs. Among the differentially expressed piRNAs tested, piR-hsa-32694, piR-hsa-26591, piR-hsa-18725, piR-hsa-18586 and piR-hsa-2912 produced qRT-PCR results consistent with the sequencing results in AZS compared with controls in the first cohort, whereas the other three genes did not show significant differences in expression. piR-hsa-32694, piR-hsa-26591, piR-hsa-18725, and piR-hsa-18586 were significantly upregulated in AZS. The diagnostic power of the four

piRNAs was further analysed using ROC analysis; piR-hsa-26591 exhibited an area under the ROC curve (AUC) of 0.913 (95% CI: 0.795-0.994). Logistic regression modelling and subsequent ROC analysis indicated that the combination of the 4 piRNAs reached good diagnostic efficacy (AUC: 0.935).

## 1. Introduction

According to a report evaluating reproductive system diseases that was published by the World Health Organization (WHO) in 2004, one in four couples in developing countries experience infertility. Infertility often causes great psychological strain, which can eventually lead to mental illness. Approximately 50% of cases of infertility are caused by male factors<sup>1</sup>. AZS is one of the major causes of human male infertility and is characterized by impaired motility of sperm, which causes infertility by preventing the sperm from reaching the egg<sup>2</sup>. However, the pathogenesis of AZS has not been absolutely elucidated. Recently, studies have shown that noncoding RNAs, especially piRNAs, play important roles in the pathogenesis of AZS.

Piwi-interacting RNAs (piRNAs) are a unique class of small noncoding RNAs that are 26–30 nt in length and bind to Piwi proteins. These short RNAs were originally discovered in mouse testicular cells and have been shown to play key roles in germ cell maintenance in mammals. Mammalian spermatogenesis is a highly regulated developmental process that can be divided into three steps, in chronological order: meiosis, mitosis and differentiation. This process produces haploid gametes. Abnormal expression of piRNAs in testicular cells may lead to the failure of spermatogenesis. Heyn et al found that five piRNAs (DQ589977, DQ591415, DQ598918, DQ601291 and DQ601609) were downregulated in cases of spermatogenic failure and that this downregulation was negatively correlated with the degree of methylation of the piwi2 and Tudor domain protein 1 (tdrd1) gene promoters. Bioinformatics analysis of these five piRNAs also revealed some other potential target genes, such as IL-16, kallikrein 1 (KLK1), G protein coupled receptor 156 (gpr156), histone cluster 1h2aa (hist1h2aa), rab24 (Ras oncogene family member), and sphingomyelin phosphodiesterase 3 (smpd3)<sup>3</sup>. Zeeba et al observed that the rs508485 mutation in the hiwi2 gene influences the stability of mRNA or changes the binding affinity of miRNA and also resulted in idiopathic nonobstructive azoospermia<sup>4</sup>. Hong and her colleagues identified 5 piRNAs as biomarkers for the diagnosis of AZS<sup>5</sup>. However, little is known about the expression patterns and functions of piRNAs in AZS.

Therefore, this study aimed to investigate the expression profile and functions of piRNAs in AZS patients. Bioinformatics analysis was also performed to describe the piRNA/mRNA interaction network, biological process and signal pathways. These results may provide potential targets for the development of novel diagnostic and therapeutic strategies against AZS. We also assessed the diagnostic value of the differentially expressed piRNAs in the seminal plasma as an additional tool for AZS.

## 2. Materials and Methods

### 2.1 Ethics statement

This study was approved by the Ethics Committee of The Second Affiliated Hospital of Nanchang University, China. All donors gave informed consent for the collection and use of their samples for this study.

### 2.2 Sample collection

We selected 4 patients diagnosed with male sterility due to AZS and 3 normal healthy males at the Maternal and Child Health Hospital in Jiangxi Province. Semen samples (6-8 ml) were collected in a sterile container. We used a sperm quality analyser to evaluate the basic semen parameters, including sperm morphology, concentration, motility, viability and quantity. According to the World Health Organization (WHO, 2010, 5th Edition) diagnostic criteria for AZS, the proportions of sperm with fast progressive motility ( $A < 25\%$ ) and total progressive motility ( $A + B < 50\%$ ) in fresh ejaculates were determined. Normal sperm samples were defined as those with the following parameters: semen concentration  $\geq 15 \times 10^6/\text{ml}$ , semen volume  $\geq 2 \text{ ml}$ , pH  $\geq 7.2$ , progressive motility (PR)  $\geq 32\%$ , and PR + nonprogressive motility (NP)  $\geq 40\%$ . The age range of both the AZS patients and the normal controls was 28-40 years old; the patients were also well matched to controls in terms of drinking and smoking status. Controls with underlying disease or abnormal sperm parameters were excluded from this study. Patients with gonadal dysplasia due to congenital or inherited disease were excluded from the study. The semen samples were liquefied at 37 °C for 30 min and separated by a discontinuous 45%/85% Percoll density gradient to eliminate the contamination from round cells, including germ cells and white blood cells. The spermatozoa in the lower (85% Percoll) layer were collected and washed 3 times with phosphate-buffered saline (PBS). The quality of the purified sperm was examined by microscopy.

### 2.3 RNA isolation and qRT-PCR assays

Total RNA was extracted from semen samples using TRIzol reagent according to the manufacturer's specifications. The yield of RNA was determined using a NanoDrop 2000 spectrophotometer (Thermo Scientific, USA), and the RNA integrity was evaluated using electrophoresis on an agarose gel stained with ethidium bromide.

Real-time PCR was performed in a LightCycler® 480 II Real-time PCR Instrument (Roche, Switzerland) in a 10  $\mu\text{l}$  PCR mixture that included 1  $\mu\text{l}$  of cDNA, 5  $\mu\text{l}$  of 2 $\times$ PerfectStart™ Green qPCR SuperMix, 0.2  $\mu\text{l}$  of universal primer, 0.2  $\mu\text{l}$  of miRNA-specific primer and 3.6  $\mu\text{l}$  of nuclease-free water. Reactions were incubated in a 384-well optical plate (Roche, Switzerland) at 94°C for 30 s, followed by 45 cycles of 94°C for 5 s and 60°C for 30 s. Each sample was run in triplicate. At the end of the PCR cycles, melting curve analysis was performed to validate the specific generation of the expected PCR product. The miRNA-specific primer sequences were designed in the laboratory and synthesized by TsingKe Biotech based on the miRNA sequences obtained from the piRNABase database (Release 20.0), as shown in Table 2. The piRNA expression levels were normalized to that of (input the reference gene, e.g., RNU6B, 5S rRNA) and were calculated using the  $2^{-\Delta\Delta\text{Ct}}$  method (Livak and Schmittgen, 2001).

## 2.4 Bioinformatic analysis

Briefly, small RNA sequencing and analysis were conducted by OE Biotech Co., Ltd. (Shanghai, China). The base reads were converted into sequence data (also called raw data/reads) by base calling. Low-quality reads were filtered out, and reads with 5' primer or poly(A) contaminants were removed. The reads without 3' adapter and insert tags and reads shorter than 15 nt and longer than 41 nt from the raw data were filtered, and clean reads were obtained. The clean reads were matched to the reference human (*Homo sapiens*) genome<sup>6</sup>. Sequences corresponding to known piRNAs were identified by perfect sequence matching to the piRBase database (piRBase release v2.0)<sup>7</sup>. miRNA sequences were determined by matching to the miRNA database (miRBase 19.0), and other small RNAs (rRNA, tRNA and snoRNA) were annotated by BLAST search against the Rfam database<sup>8-10</sup>.

Differentially expressed piRNAs between the AZS group and the normal control group were identified using DEG-seq software. Genes with q-values (FDRs) lower than 0.05 and fold-changes higher than 2.0 were considered differentially expressed<sup>11,12</sup>. The targets of differentially expressed piRNAs were predicted by using the miRNA database software<sup>9,13,14</sup>

with the following parameters:  $S \geq 150$   $\Delta G$ ,  $\leq -30$  kcal/mol, and strict 5' seed pairing.

Functional annotation of predicted target genes was performed by Gene Ontology, and Kyoto Encyclopedia of Genes and Genomes pathway analysis was conducted<sup>15</sup>. Corrected P values < 0.05 were considered to indicate significantly enriched differentially expressed piRNA-target-genes.

## 3. Statistical analysis

Differences in demographic variables and expression levels of piRNA between the AZS group and the control group were evaluated with the chi-squared test. Fisher's exact test was used to assess the significance of GO terms or KEGG pathway identifiers in the differentially expressed genes.

To evaluate the specificity and sensitivity of each piRNA in the diagnostic value of AZS, we analysed ROC curves and AUC. A large Youden's index (sensitivity and specificity-1) was used to calculate the cut-off value of the optimal diagnostic point of the ROC curve. Binary logistic regression and ROC curves were used to improve diagnostic efficiency. All statistical analyses were performed in SPSS (Version 23, SPSS Inc.) and R software (version 3.5.0). A value of  $p < 0.05$  was considered statistically significant.

## 4. Results

### 4.1 Patient data analysis and seminal plasma piRNA profiling

Seven semen samples were collected from AZS and male normal healthy controls according to the WHO guidelines. The basic information and sperm parameters of AZS and healthy controls are presented in Table 1. There were no distinct differences in age, BMI, smoking or drinking status between the AZS and control groups ( $P > 0.05$ ). However, there were obvious differences in motility and sperm morphology ( $P < 0.001$ ) between the AZS and control groups. The semen samples were observed under a light microscope equipped with a 100 x oil objective and were found to be free of seminiferous epithelium cells, somatic cells and white blood cells.

Small RNAs were extracted from semen samples of healthy controls and AZS patients and analysed by Illumina high-throughput sequencing. First, the splice sequences were

eliminated from the raw reads and small RNA sequences of different lengths were obtained. After quality control steps, small RNAs in the range of 10 to 45 nucleotides were obtained. Analysis of the length distribution showed that all three seminal plasma samples contained many small RNAs with sizes consistent with miRNAs (18–24 nucleotides) and piRNAs (26–31 nucleotides) (Figure 1). In general, piRNAs account for approximately 0.05% of the total small RNA content and 0.5% of the total small RNA reads in seminal plasma. Among the 77 million piRNAs annotated in the piRBase database, 30,348 and 53,932 piRNAs were detected in the seminal plasma of healthy controls and AZS patients, respectively.

#### 4.2 Identification of differential piRNA expression

We used DEG-seq software to analyse the differences in piRNA expression levels between the AZS and normal groups. A total of 283 DEGs were obtained with a significance threshold of  $|\log_2FC| \geq 1$  and  $P\text{-value} \leq 0.05$ , comprising 114 upregulated DEGs and 169 downregulated DEGs. We displayed volcano plots and heatmaps of DEGs in Figure 2A, B. Twenty differentially expressed piRNAs, including the top 10 most strongly upregulated and top 10 most strongly downregulated piRNAs, are presented in Figure 2C, D and Table 2.

#### 4.3 GO and KEGG pathway analysis

GO enrichment analysis was used to explore the potential functions of these DEGs. The top 30 most frequent GO terms are given in Figure 3. As shown in Figure 3A, B, and C, the top cellular component terms were nucleus, cytosol, cytoplasm and nucleoplasm; the top biological process terms were regulation of transcription-DNA templated, signal transduction and positive regulation of transcription from the RNA polymerase II promoter; and the top molecular function terms were metal ion binding, ATP binding, DNA binding and RNA binding.

KEGG pathway analysis was used to identify all pathways and the top 20 most highly enriched pathways related to the potential target genes of the differentially expressed piRNAs (Figure 3C, D ,E). The target genes enriched among the AZS DEGs were mainly involved in focal adhesion, actin cytoskeleton, axon guidance, ErbB signalling pathway and other signal transduction pathways.

#### 4.4 Validation of differentially expressed piRNAs with qRT-PCR

To validate the small RNA-seq results, qRT-PCR analysis was performed on 8 randomly selected differentially expressed piRNAs. We selected 4 upregulated piRNAs (piR-hsa-32694, piR-hsa-26591, piR-hsa-18725, piR-hsa-18586) and 4 downregulated piRNAs (piR-hsa-6148, piR-hsa-32713, piR-hsa-2912, piR-hsa-15551) and performed qRT-PCR to quantify their expression in AZS and normal group semen samples. As shown in Figure 4, the expression of piR-hsa-32694, piR-hsa-26591, piR-hsa-18725, and piR-hsa-18586 were upregulated in the AZS group compared with the normal group, and piR-hsa-2912 indicated a downregulated expression pattern, which was consistent with the small RNA-seq results. However, the other three genes did not show significant differences in expression between the two samples.

#### 4.5 Evaluate piRNAs as a diagnostic marker for AZS.

Illumina high-throughput sequencing analysis indicated that piR-hsa-32694, piR-hsa-26591, piR-hsa-18725, and piR-hsa-18586 expression levels were obviously different between AZS and controls. Therefore, the diagnostic value these four piRNAs as potential biomarkers for AZS was assessed and evaluated. We used ROC curve analysis to

calculate whether piR-hsa-32694, piR-hsa-26591, piR-hsa-18725, and piR-hsa-18586 could distinguish AZS patients from controls. As shown in Figure 5, the AUCs of piR-hsa-18586, piR-hsa-18725, piR-hsa-26591, and piR-hsa-32694 were 0.850 ( $P < 0.007$ ), 0.875 ( $P < 0.004$ ), 0.913 ( $P < 0.001$ ), and 0.883 ( $P < 0.003$ ), respectively. To improve the diagnostic power, we combined these four piRNAs in 80 patients with AZS and 100 normal controls as a training cohort to build a diagnostic model using binary logistic regression. An AUC of 0.935 was found (Figure 6) when the model was applied to an independent testing cohort (40 controls and 40 AZS patients) to estimate the diagnostic accuracy of the joint diagnosis model. These results showed that the combined diagnostic model can distinguish AZS patients from controls (specificity = 95%, sensitivity = 90%).

## 5. Discussion

In this study, we identified the comprehensive piRNA expression pattern in semen samples of AZS patients. The small RNA sequencing data revealed that a total of 283 significant DEGs (114 upregulated and 169 downregulated) were obtained at a threshold of  $|\log_2FC| \geq 1$  and  $P\text{-value} \leq 0.05$ . Further *in silico* analyses, including piRNA/mRNA interaction network, GO and KEGG pathway analyses, were performed to predict the functions of differentially expressed piRNAs manifesting the critical roles of piRNAs in regulating spermatogenesis. These results were validated by qRT-PCR assays of 8 randomly selected piRNAs, specifically, 4 upregulated piRNAs (piR-hsa-32694, piR-hsa-26591, piR-hsa-18725, piR-hsa-18586) and 4 downregulated piRNAs (piR-hsa-6148, piR-hsa-32713, piR-hsa-2912, piR-hsa-15551). Of these, piR-hsa-32694, piR-hsa-26591, piR-hsa-18725, piR-hsa-18586 and piR-hsa-2912 produced qRT-PCR results consistent with the sequencing results, whereas the other three genes were did not show significant differences in expression between AZS samples and controls. The piR-hsa-32694, piR-hsa-26591, piR-hsa-18725, and piR-hsa-18586 piRNAs were significantly upregulated in AZS. The diagnostic power of these four piRNAs was further analysed using ROC analysis, and piR-hsa-26591 was found to be the most diagnostic for AZS, with an AUC of 0.913 (95% CI: 0.795-0.994). A logistic regression model including all four piRNAs was constructed, and ROC analysis revealed that the combination model reached good diagnostic efficacy (AUC: 0.935). Therefore, these piRNAs were identified as hub genes that participate in AZS pathogenesis and could serve as biomarkers for this disease.

Human spermatogenesis is a carefully coordinated and precisely regulated biological process that includes rigorous regulation of gene expression at specific phases<sup>3</sup>. Noncoding RNAs such as circular RNAs (circRNAs), long noncoding RNAs (lncRNAs) and microRNAs (miRNAs) are critical regulators of gene expression in multiple phases of spermatogenesis<sup>16-18</sup>. piRNAs are a recently discovered class of small RNAs that are mainly expressed in germ cells. piRNAs play important roles in maintaining the integrity of germline DNA, inhibiting transposon transcription, inhibiting translation, participating in the formation of heterochromatin, epigenetic regulation and germ cell genesis<sup>19</sup>. To date, there have been few reports describing the changes in piRNA profile in human AZS. In view of the importance of piRNAs in male reproduction, we studied the expression pattern of piRNAs in AZS sperm, analysing the differential expression of piRNAs in AZS and controls to explore the potential pathological mechanism of AZS.

Studies have shown that piRNAs regulate gene silencing mainly by binding with piwi

subfamily proteins to form piRNA complexes (piRCs). The human genome encodes four piwi protein family members, Hiwi, Hili, Hiwi2 and PIWIL3, which are highly expressed in testicular cells<sup>20</sup>. Liu et al. showed that a Hiwi mutation is pathogenic, causing male infertility, and established a clear relationship between the human piwi gene and male infertility<sup>21</sup>. Gou and other research groups have demonstrated that in addition to its main functions in transposon silencing, the piwi/piRNA mechanism mediates the degradation of a large number of mRNA transcripts in various germ cells through miRNA- or/and siRNA-like mechanisms<sup>22,23</sup>. In this study, we predicted the target genes of piRNAs that were differentially expressed between the control group and the AZS group to explore how the differentially expressed piRNAs might change sperm function. In addition, enrichment analysis based on these predicted target genes showed that cell macromolecule biosynthesis, spermatogenesis, male gametogenesis, cytoplasmic transport and RNA metabolism were changed in the AZS group. KEGG pathway analysis showed that the differentially expressed piRNA targets were involved in focal adhesion kinase (FAK), actin cytoskeleton, axon guidance, the ErbB signalling pathway and other signal transduction pathways. An interesting study showed that FAK coordinates the transport of spermatozoa and prealbumin spermatids through the epithelium along the functional axis of the acroplasmic thinning (ES)-blood testicular barrier basement membrane (BM)<sup>24</sup>. A study on pig sperm found that many differentially expressed mRNAs related to infertility are involved in metabolic processes<sup>25</sup>. Specifically, during the sperm deformation stage, macromolecule metabolism and transportation are well controlled, and the production of fertile sperm is ensured through 26S proteasome exchange, digestion and absorption of nuclear protein, and elimination of RNA through phagocytosis of cytoplasmic droplets, residues and Sertoli cells<sup>26</sup>. Our results also support the view that focal adhesion kinase and metabolism of fatty acids and other macromolecules may play key roles in spermatogenesis and in AZS.

Previous reports identified mRNAs that were related to sperm motility, particularly transcripts encoding flagella-associated protein 45 (CFAP45) and bromodomain-containing 2 (BRD2). The content of CFAP45 in low-motility sperm was significantly higher than that in high-motility sperm, while the content of BRD2 mRNA in AZS was significantly lower<sup>27,28</sup>. This indicates that low sperm motility could be attributed to the abnormal expression of CFAP45 and BRD2. Here, we found that the expression of four piRNAs, piR-hsa-32694, piR-hsa-26591, piR-hsa-18725 and piR-hsa-18586, was significantly upregulated in AZS patients. Furthermore, the ROC curve analysis revealed that the AUC of the combination of the 4 piRNAs was equal to 0.935 and demonstrated 93.8% sensitivity and 89.3% specificity. These results suggested that these four piRNAs were significantly and independently associated with sperm motility and male infertility. Bioinformatics analysis showed that CAFP45 and BRD2 were the potential target genes of piR-hsa-26591. Whether piR-hsa-26591 silences the expression of CAFP45 and BRD2 and leads to decreased sperm motility and infertility in AZS patients needs further study. Overall, this study may provide insights into the use of human sperm cell-specific piRNAs as biomarkers for AZS screening. Our study has some limitations. First, the small number of samples limited the statistical power of the microarray analysis. Second, we did not use a luciferase reporter assay to further verify the relationship between piR-hsa-26591 and CAFP45.

In summary, we have defined a distinctive seminal plasma piRNA signature in AZS. Moreover, our findings may provide additional information regarding the molecular mechanisms of male infertility.

Figure 1. Small RNA length distribution map (miRNAs (18-24 nucleotides), piRNAs (26-31 nucleotides))

Figure 2. Differentially expressed genes between AZS patients and healthy controls. (A-B) Volcano plot and heatmap of DEGs between AZS and normal groups. (C-D) Volcano plot and heatmap of Top 20 DEGs between AZS and normal groups. Grey dots represent genes that are not significantly differentially expressed. Green and red dots indicate significant up-regulated and down-regulated genes ( $|\log_2$  fold change $| \geq 1.0$  and  $P$  value  $\leq 0.05$ ), respectively. Row and column were used to represent DEGs and samples. case: AZS; con: normal. The color scale represented the expression levels of DEGs.

Figure 3. Enrichment analyses of differentially expressed genes. (A-C) Top 30 GO enriched terms among all differentially expressed genes and among upregulated and downregulated genes separately. (D-F) Top 20 KEGG pathways enriched among the total differentially expressed, upregulated and downregulated genes.

Figure 4. Expression levels of piRNAs in semen in AZS and healthy controls. (A) piR-hsa-32694, (B) piR-hsa-26591, (C) piR-hsa-18725, (D) piR-hsa-18586, (E) piR-hsa-6148, (F) piR-hsa-3217, (G) piR-hsa-15551, (H) piR-hsa-2912.

Figure 5 Diagnostic power of the four piRNA potential biomarkers determined using receiver operating characteristic (ROC) curve analysis(A-D). (A) piR-hsa-18586, (B) piR-hsa-18725, (C) piR-hsa-26591, (D) piR-hsa-32594. (The optimal diagnostic point was assessed at cutoff values with largest Youden's index (sensitivity and specificity - 1))

Figure 6 Diagnostic model of AZS using 4 piRNAs (using expression of piR-hsa-18586, piR-hsa-18725, piR-hsa-26591 and piR-hsa-32594 as a training set)

## 6. References

- 1 Vander Borgh, M. & Wyns, C. Fertility and infertility: Definition and epidemiology. *Clin Biochem* **62**, 2-10, doi:10.1016/j.clinbiochem.2018.03.012 (2018).
- 2 Wu, X. *et al.* The Biogenesis and Functions of piRNAs in Human Diseases. *Molecular Therapy - Nucleic Acids* **21**, 108-120, doi:10.1016/j.omtn.2020.05.023 (2020).
- 3 Marsit, C. J. *et al.* Epigenetic Disruption of the PIWI Pathway in Human Spermatogenic Disorders. *PLoS ONE* **7**, doi:10.1371/journal.pone.0047892 (2012).
- 4 Kamaliyan, Z., Pouriamanesh, S., Soosanabadi, M., Gholami, M. & Mirfakhraie, R. Investigation of piwi-interacting RNA pathway genes role in idiopathic non-obstructive azoospermia. *Scientific Reports* **8**, doi:10.1038/s41598-017-17518-4 (2018).
- 5 Hong, Y. *et al.* Systematic characterization of seminal plasma piRNAs as molecular biomarkers for male infertility. *Sci Rep* **6**, 24229, doi:10.1038/srep24229 (2016).
- 6 Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *Embnet Journal* **2011**, **17(1)** (2011).
- 7 Wang, J. *et al.* piRBase: a comprehensive database of piRNA sequences. *Nucleic Acids Res* **47**, D175-D180, doi:10.1093/nar/gky1043 (2019).
- 8 Griffiths-Jones S, S. H., van Dongen S. miRBase: tools for microRNA genomics. *Nucleic*

- Acids Research* **36: D154-158**, doi:10.1093/nar/gkm952 (2008).
- 9 Griffiths-Jones, S., Grocock, R. J., van Dongen, S., Bateman, A. & Enright, A. J. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res* **34**, D140-144, doi:10.1093/nar/gkj112 (2006).
- 10 Griffiths-Jones, S., Bateman, A., Marshall, M., Khanna, A. & Eddy, S. R. Rfam: an RNA family database. *Nucleic Acids Res* **31**, 439-441, doi:10.1093/nar/gkg006 (2003).
- 11 Love, M. I. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *BMC Bioinformatics* **15(12): 550** (2009).
- 12 Tino, P. Basic properties and information theory of Audic-Claverie statistic for analyzing cDNA arrays. *BMC Bioinformatics* **10**, 310, doi:10.1186/1471-2105-10-310 (2009).
- 13 John, B. *et al.* Human MicroRNA targets. *PLoS Biol* **2**, e363, doi:10.1371/journal.pbio.0020363 (2004).
- 14 Enright AJ, J. B., Gaul U., MicroRNA targets in Drosophila. *Genome Biology* **5(1): R1** (2003).
- 15 Alexa A, R. J. topGO: enrichment analysis for gene ontology. *R package version 2.8* (2013).
- 16 Manfredola, F. *et al.* CircRNA Role and circRNA-Dependent Network (ceRNET) in Asthenozoospermia. *Front Endocrinol (Lausanne)* **11**, 395, doi:10.3389/fendo.2020.00395 (2020).
- 17 Lu, H. *et al.* RNA-sequencing and bioinformatics analysis of long noncoding RNAs and mRNAs in the asthenozoospermia. *Biosci Rep* **40**, doi:10.1042/BSR20194041 (2020).
- 18 Heidary, Z., Zaki-Dizaji, M., Saliminejad, K. & Khorram Khorshid, H. R. MicroRNA profiling in spermatozoa of men with unexplained asthenozoospermia. *Andrologia* **51**, e13284, doi:10.1111/and.13284 (2019).
- 19 Larriba, E. & Del Mazo, J. An integrative piRNA analysis of mouse gametes and zygotes reveals new potential origins and gene regulatory roles. *Sci Rep* **8**, 12832, doi:10.1038/s41598-018-31032-1 (2018).
- 20 Ozata, D. M., Gainetdinov, I., Zoch, A., O'Carroll, D. & Zamore, P. D. PIWI-interacting RNAs: small RNAs with big functions. *Nat Rev Genet* **20**, 89-108, doi:10.1038/s41576-018-0073-3 (2019).
- 21 Dai, P., Wang, X. & Liu, M. F. A dual role of the PIWI/piRNA machinery in regulating mRNAs during mouse spermiogenesis. *Sci China Life Sci* **63**, 447-449, doi:10.1007/s11427-020-1632-5 (2020).
- 22 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).
- 23 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).
- 24 Xiao, X., Mruk, D. D., Wong, C. K. & Cheng, C. Y. Germ cell transport across the seminiferous epithelium during spermatogenesis. *Physiology (Bethesda)* **29**, 286-298, doi:10.1152/physiol.00001.2014 (2014).
- 25 Yang, C. C. *et al.* Seasonal effect on sperm messenger RNA profile of domestic swine (*Sus Scrofa*). *Anim Reprod Sci* **119**, 76-84, doi:10.1016/j.anireprosci.2009.12.002 (2010).
- 26 Percipalle, P. New insights into co-transcriptional sorting of mRNA for cytoplasmic

- transport during development. *Semin Cell Dev Biol* **32**, 55-62, doi:10.1016/j.semcdb.2014.03.009 (2014).
- 27 Dougherty, G. W. *et al.* CFAP45 deficiency causes situs abnormalities and asthenospermia by disrupting an axonemal adenine nucleotide homeostasis module. *Nat Commun* **11**, 5520, doi:10.1038/s41467-020-19113-0 (2020).
- 28 Jodar, M., Kalko, S., Castillo, J., Balleca, J. L. & Oliva, R. Differential RNAs in the sperm cells of asthenozoospermic patients. *Hum Reprod* **27**, 1431-1438, doi:10.1093/humrep/des021 (2012).

# Figures

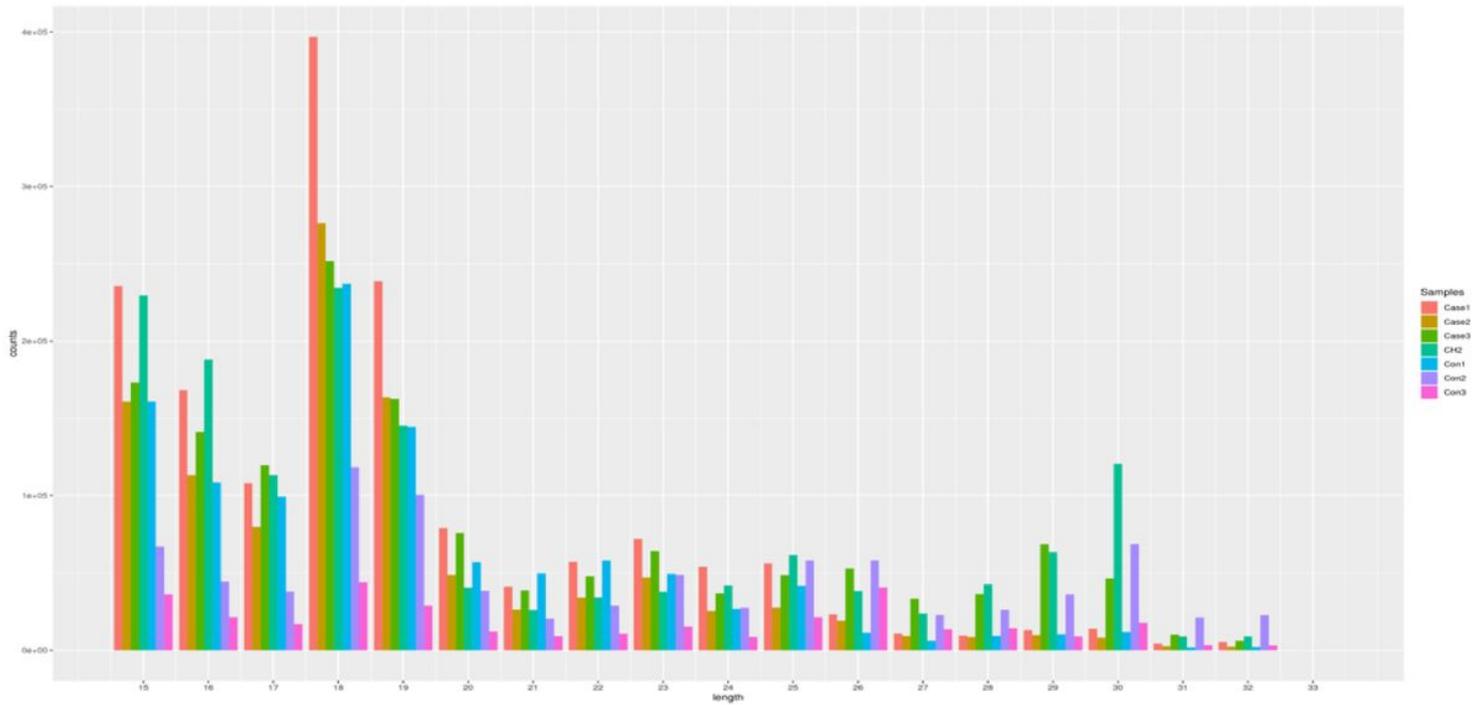


Figure 1

Small RNA length distribution map (miRNAs (18-24 nucleotides), piRNAs (26-31 nucleotides))

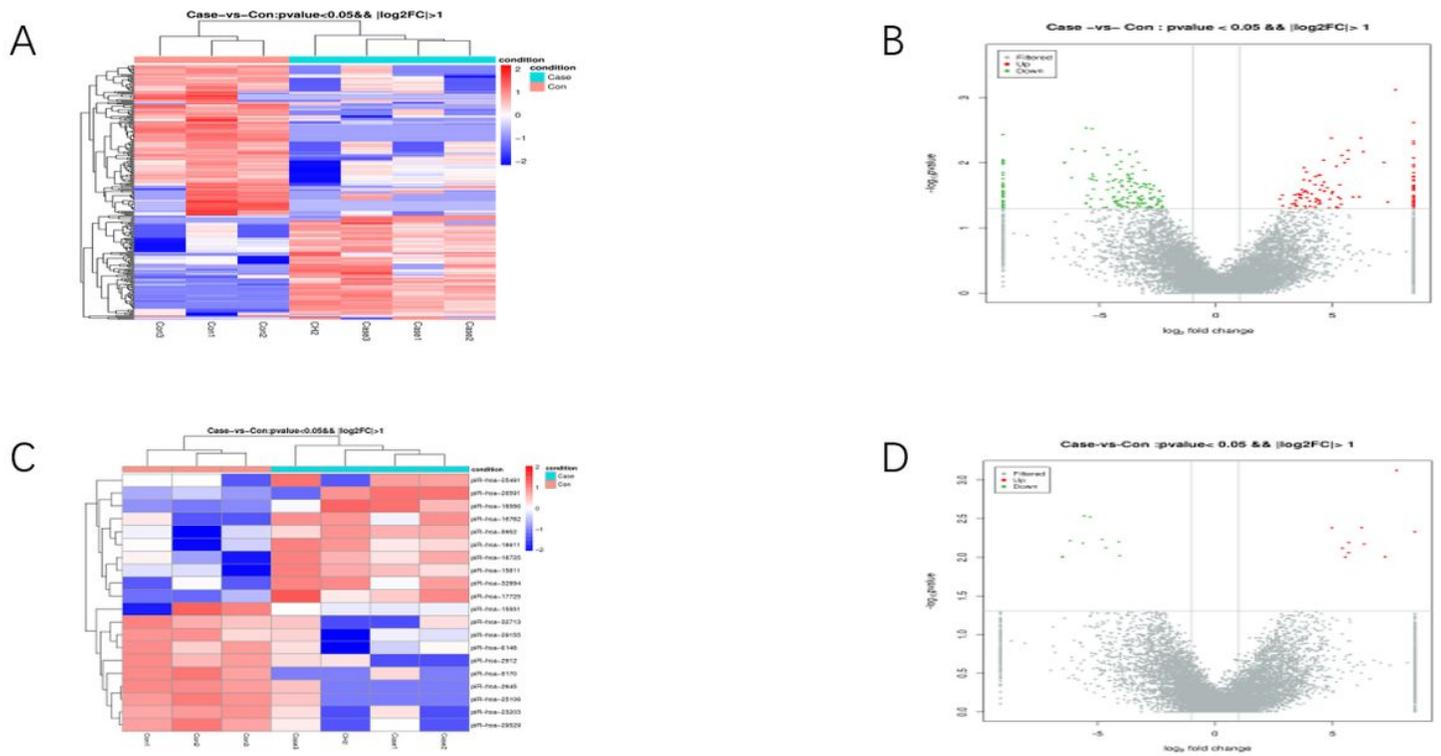
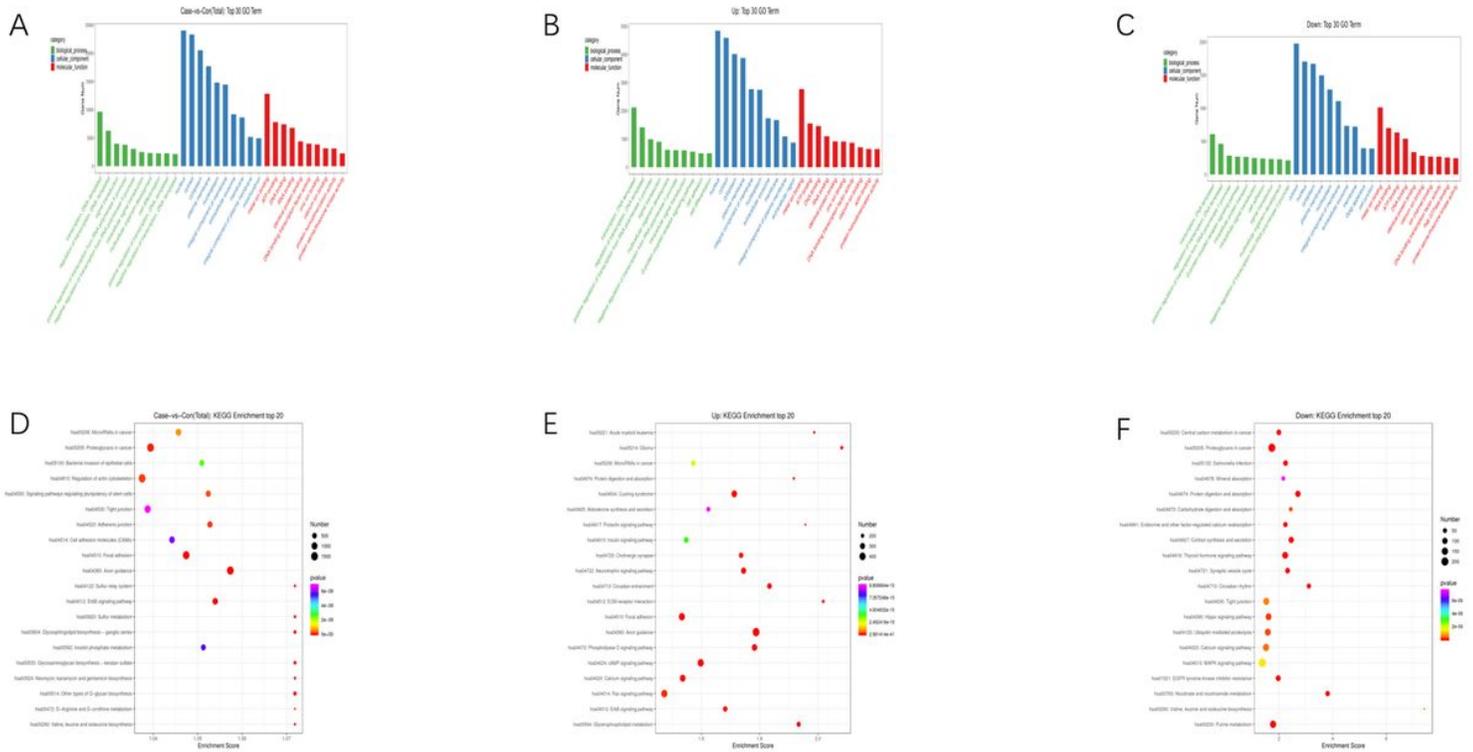


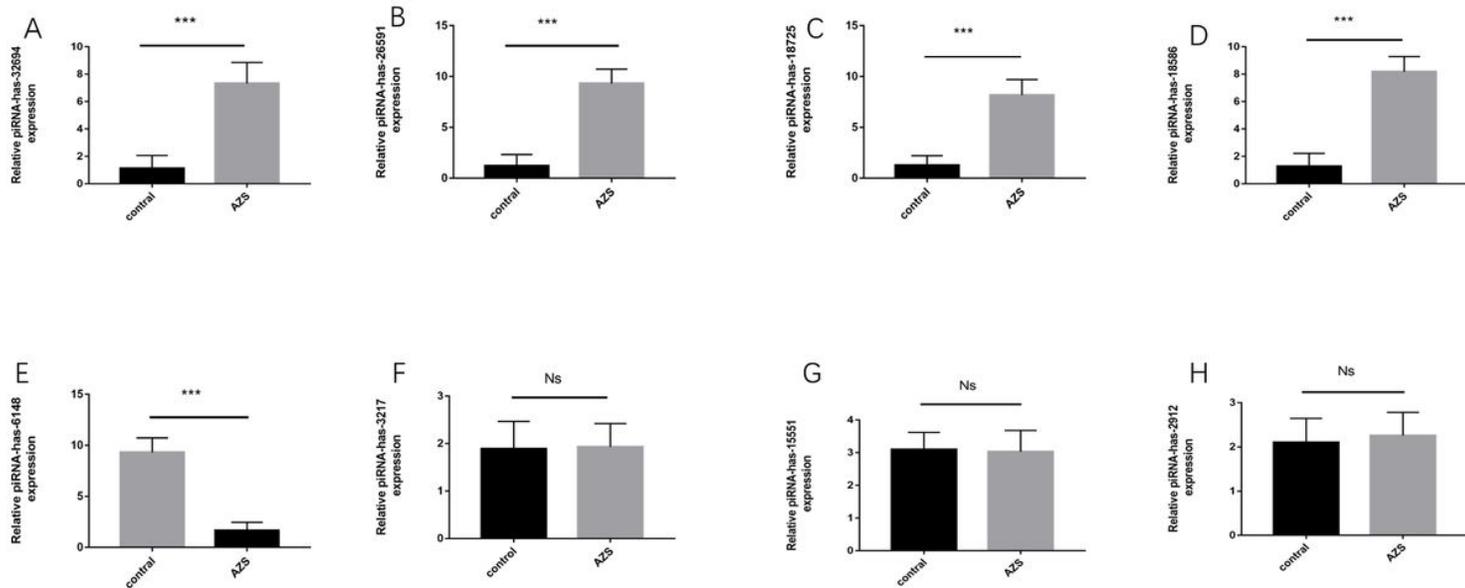
Figure 2

Differentially expressed genes between AZS patients and healthy controls. (A-B) Volcano plot and heatmap of DEGs between AZS and normal groups. (C-D) Volcano plot and heatmap of Top 20 DEGs between AZS and normal groups. Grey dots represent genes that are not significantly differentially expressed. Green and red dots indicate significant up-regulated and down-regulated genes ( $|\log_2$  fold change $\geq 1.0$  and  $P$  value  $\leq 0.05$ ) respectively. Row and column were used to represent DEGs and samples. case: AZS; con: normal. The color scale represented the expression levels of DEGs.



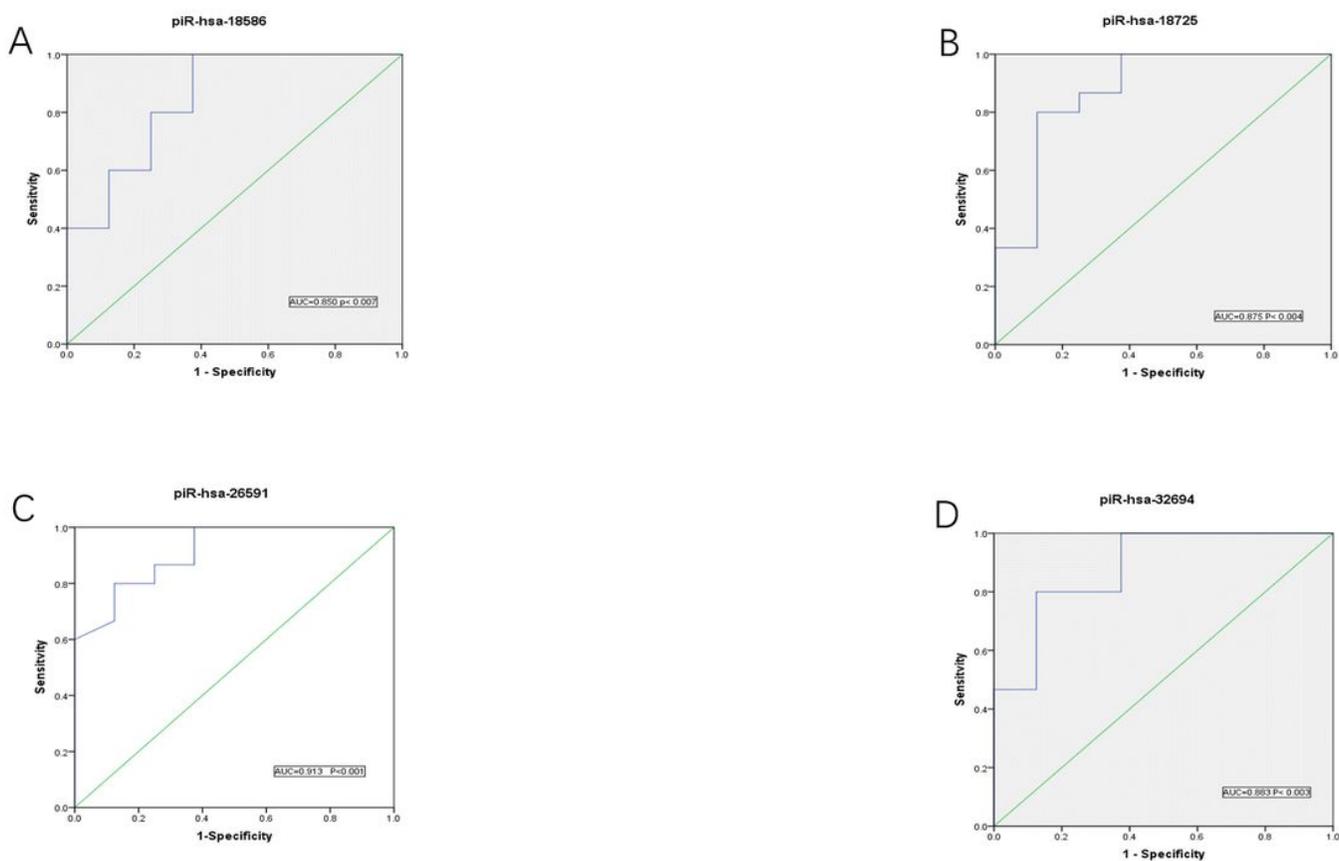
**Figure 3**

Enrichment analyses of differentially expressed genes. (A-C) Top 30 GO enriched terms among all differentially expressed genes and among upregulated and downregulated genes separately. (D-F) Top 20 KEGG pathways enriched among the total differentially expressed, upregulated and downregulated genes.



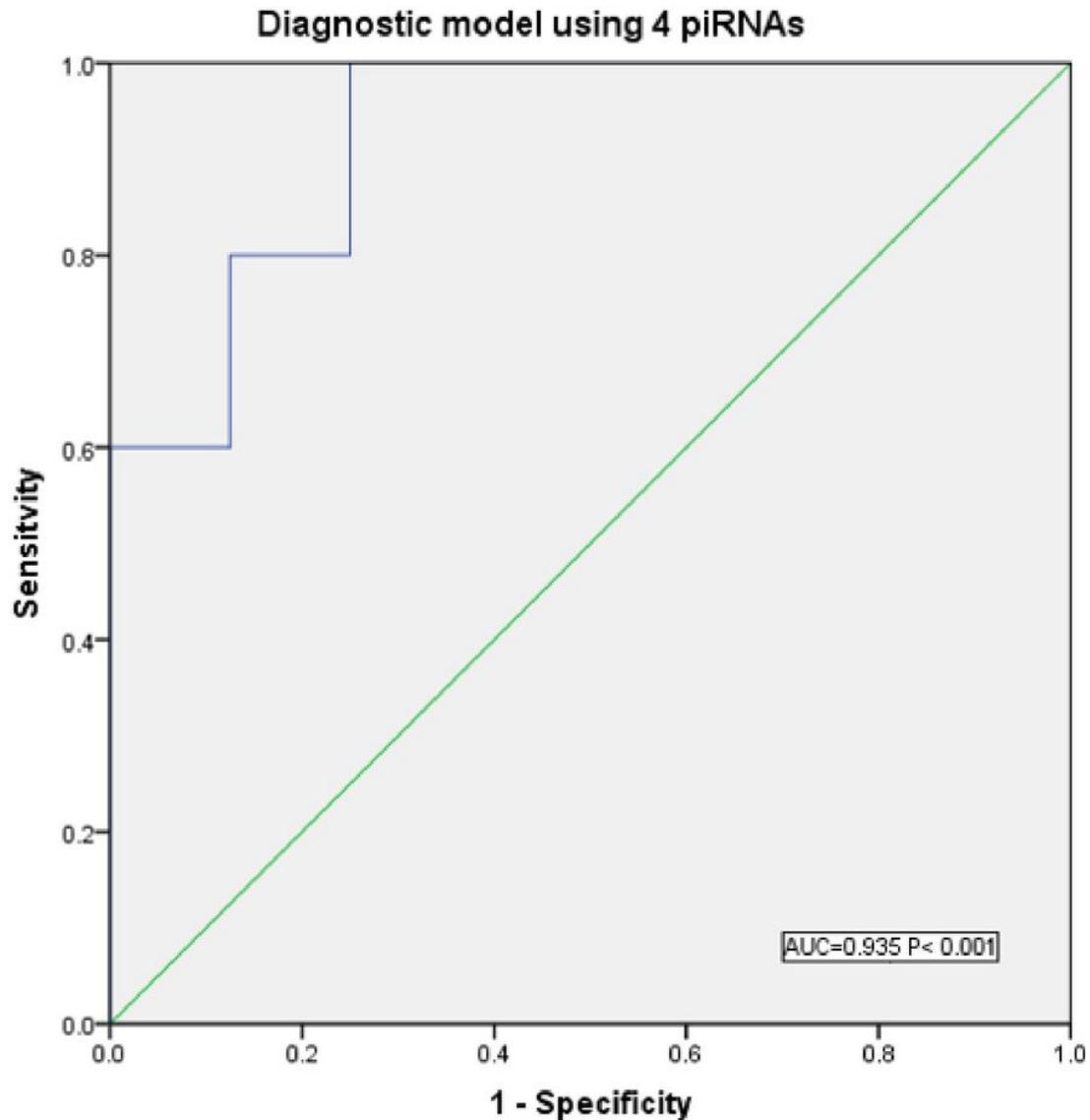
**Figure 4**

Expression levels of piRNAs in semen in AZS and healthy controls. (A) piR-hsa-32694, (B) piR-hsa-26591, (C) piR-hsa-18725, (D) piR-hsa-18586, (E) piR-hsa-6148, (F) piR-hsa-3217, (G) piR-hsa-15551, (H) piR-hsa-2912.



**Figure 5**

Diagnostic power of the four piRNA potential biomarkers determined using receiver operating characteristic (ROC) curve analysis(A-D). (A) piR-hsa-18586, (B) piR-hsa-18725, (C) piR-hsa-26591, (D) piR-hsa-32594. (The optimal diagnostic point was assessed at cutoff values with largest Youden's index (sensitivity and specificity - 1))



**Figure 6**

Diagnostic model of AZS using 4 piRNAs (using expression of piR-hsa-18586, piR-hsa-18725, piR-hsa-26591 and piR-hsa-32594 as a training set)

## Supplementary Files

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

- [table2.docx](#)

- [table.docx](#)