



基于高通量转录组测序技术的无精子症患者睾丸组织比较转录组分析*

罗 斌¹⁾ 吕自力¹⁾ 单旭东¹⁾ 赵情梅¹⁾ 杜珍珍¹⁾ 张 霞¹⁾ 李俊君²⁾ 梁 鑫^{1)**}

¹⁾ 成都中医药大学医学与生命科学学院/附属生殖妇幼医院, 生殖医学中心, 成都 610041;

²⁾ 成都中医药大学医学与生命科学学院/附属生殖妇幼医院, 男科, 成都 610041)

摘要 本研究探讨不同程度无精子症患者睾丸组织的转录组差异, 了解差异表达基因 (DEG) 在功能、分类和代谢通路的不同, 揭示无精子症患者精子发生分子机制, 为促进男性不育研究的发展提供理论基础. 选取 1 份非梗阻性无精子症和 4 份梗阻性无精子症患者睾丸组织样品 (从无精子到有精子), 进行 RNA 提取和文库构建, 利用 Illumina HiSeq™ 2500 高通量测序, 构建无精子症患者睾丸组织转录组文库, 并用生物信息学方法进行分析. 结果发现, 样品比对基因组数据库的平均比率为 94.38%, 共检测 2 242 个属于预测新的蛋白质编码基因的转录本. 得到差异表达基因统计结果为: NOA vs. OA1 基因上调 8 045, 下调 1 150; OA1 vs. OA2 基因上调 1 538, 下调 420; OA2 vs. OA3 基因上调 1 275, 下调 1 690; OA3 vs. OA4 基因上调 1 834, 下调 1 853. 比较 5 例无精子症睾丸组织的差异基因 KEGG, 主要富集在 RNA 降解通路、基底细胞瘤通路、癌通路、黑色素生成通路和调节干细胞多能性等信号通路. PRM1、PRM2、TNPI、UBXN6、CXCL16、NUPR2、CCDC136 和 CRISP2 等基因的表达呈递增趋势, 并具有时序特异性. 此外, 5 例无精子症睾丸组织的表达基因有不同程度的基因融合. 综上, 基因融合可能和无精子症相关, 并且不同程度无精子症患者睾丸组织的差异表达基因数量、功能、分类和代谢通路不同. 本研究筛选出精子发生、精子运动等差异表达基因, 丰富了无精子症患者睾丸组织转录组信息, 为开展无精子症患者睾丸组织相关基因及分子调控机制的研究奠定基础.

关键词 无精子症, 睾丸组织, 转录组测序, 辅助生殖, 差异表达基因

中图分类号 Q754

DOI: 10.16476/j.pibb.2020.0092

目前, 发现无精子症主要有性腺功能低下、精索静脉曲张、Y 染色体微缺失和基因突变等病因, 但许多梗阻性无精子症 (obstructive azoospermia, OA) 和非梗阻性无精子症 (non-obstructive azoospermia, NOA) 的病因仍不清楚^[1]. 因此, 阐明人类精子发生过程中的基因表达调控机制是发育生物学的重大科学问题, 对理解、诊断和治疗男性不育相关疾病至关重要. 一部分 OA 患者可通过睾丸组织手术取精方式获得活动精子, 利用卵胞浆内单精子显微注射技术 (intracytoplasmic sperm injection, ICSI) 获得后代, 而 NOA 患者或者 OA 无活动精子患者行辅助生殖 ICSI 技术, 存在巨大困难^[2]. 近年来临床工作者进行了广泛而深入的研究, 但其机制尚不清楚, 这为男性不育症的治疗提出了新的挑战^[3]. 有研究者利用基因敲除等方法发

现, 与男性不育症密切相关的基因近 200 个, 这些基因的突变、缺失或表达异常, 可能是男性不育症发生的重要原因^[4].

与基因敲除或基因组测序等研究方法相比, 高通量转录组测序 (RNA-sequencing, RNA Seq) 技术具有显著的优势: 既可在较短时间内大规模研究组织或细胞里多个基因的差异表达, 筛选相关基因, 也可以研究细胞通路和调控网络, 为解析不孕不育调控机制提供理论基础^[5-6]. 近期, 研究人员

* 成都中医药大学基金医院专项基金 (2017-EL-25), 四川省科技厅项目 (2018YSZH0028), 四川省卫生健康委课题 (19PJ033), 四川省中医药管理局中医药科研专项 (2018JYC010), 成都中医药大学实验技术项目 (Z1605) 资助.

** 通讯联系人.

Tel: 17313164091, E-mail: 184516824@qq.com

收稿日期: 2020-04-07, 接受日期: 2020-05-12

应用精度高的单细胞 RNA-Seq 方法, 建立小鼠精子发生过程的转录组图谱, 揭示了精子发生过程中基因表达调控的动态变化, 这为人类精子转录组研究奠定了基础^[7]. 通过运用 RNA-Seq 技术与 1 000 多份高丰度精子转录本比对, 初步鉴定出与精子发生、转录调控、精子运动、受精和胚胎发生等生物学过程相关的基因, 将这些基因用于比较受精卵和雄性精子的 mRNA 产物, 发现特发性不孕症患者与正常生育个体生殖细胞转录谱之间的不同^[8]. 目前已成功利用 RNA-Seq 技术在人体及其他物种植入前胚胎发育的早期胚胎发育调控机制方面开展了一些初步研究^[9-11], 而人类精子发生方面的转录组研究较少^[12]. 本研究利用转录组测序技术, 研究人类无精子症患者的精子功能, 以期发现更多与生精功能相关的分子机制.

1 材料与方法

1.1 睾丸组织

在成都中医药大学附属生殖妇幼医院选取 5 例无精子症患者 (30±2 岁) 进行体检和内分泌检查、染色体检查. 未发现其他疾病后, 经患者知情同意, 获取其睾丸活检样本, 置入 RNAlater 保护液, 冻存于液氮中 (-196℃).

1.2 RNA 提取、文库构建及测序

用 1% 琼脂糖凝胶进行 RNA 降解和污染监测. 选用生物分析仪 Agilent2100 (Agilent technologies, CA, 美国) 检测 RNA 完整性, 采用带有 Oligo dT 的核酸序列将收集到的各个阶段 RNA 进行逆转录, 形成 1st cDNA; 再利用 Smart-seq2 技术进行 PCR 扩增富集核酸, 纯化扩增产物: 包括 DNA 片段化处理、末端修复、加碱基 A、加接头以及 PCR 扩增和文库质控等, 建成文库. 采用测序策略为 PE125 模式及 Illumina HiSeq™ 2500 平台对文库进行测序.

1.3 转录组数据分析

根据 Illumina HiSeq™ 2500 平台测序得到原始数据 (raw reads), 通过去低质量序列、去接头污染等数据过滤过程, 完成数据处理, 得到高质量的序列 (clean reads), 后续所有分析都是基于 clean reads. RPKM (reads per kilobase of exon model per million mapped reads) 值来衡量基因的表达量, RPKM 是利用 RNA-Seq 技术用来定量估计

基因表达值的一个非常有效的工具^[13].

本试验采用无生物学重复样品, 处理该类样品时对采用所得的序列用 DEG-seq (differentially expressed genes-sequencing) 进行基因差异表达分析, 比较处理组与参考组, 并选取 $|\log_2 \text{Ratio}| \geq 1$ 和 $P < 0.05$ 的基因作为差异表达基因, 差异倍数 (fold change) 在 2 倍以上, 满足这些条件限制的基因被确定为差异表达显著基因, 得到上下调基因个数. DEG-seq 是基于 count 值服从二项分布的假设, 借鉴 MAplot 理论, 估计 M 值的近似正态分布, 从而计算每个基因在两个样品中不差异表达的概率值 P value, 通过矫正获得 Q value. 计算得到的 P value 通过校正阈值之后, 满足此条件的 GO Term (gene ontology) 定义为在差异表达基因中显著富集的 GO Term 以及 KEGG (Kyoto encyclopedia of genes and genomes, KEGG) 分析^[14]. 使用 Mfuzz 及 SOAPfuse 对 5 例睾丸组织数据进行融合基因分析和时间序列聚类分析^[15].

2 结果与分析

2.1 测序质量评估及基本数据分析

5 例睾丸组织样本如表 1 所示: NOA 为非梗阻性无精子症, 活检未见精子; OA1~OA4 为梗阻性无精子症, 其中 OA1 活检发现 0~1 个精子/200 倍视野, 未见活动精子; OA2 活检发现 2~5 个精子/200 倍视野, 未见活动精子; OA3 活检发现 2~5 个精子/200 倍视野, 发现 <1 个活动精子/200 倍视野, 有活力, 非前向运动; OA4 活检发现 2~5 个精子/200 倍视野, 发现 1~2 个活动精子/200 倍视野, 有前向运动精子. OA2、OA3 和 OA4 患者在我院进行辅助生殖 ICSI 技术治疗, 其中 OA4 获得临床妊娠.

经检测, 本研究中构建的人类 NOA 和 OA 等 5 个无精子症睾丸组织样本微量文库满足转录组测序要求. 5 个样本 Q30 的百分比为 93.97%~95.13%, 说明测序质量和文库构建质量高, 测序数据准确可靠, 可满足后续分析. 测序结果中 5 个样本碱基 A-T、C-G 含量都基本对应重合, 说明碱基组成稳定平衡, 测序质量高. 从碱基质量分布可知, 5 个样本碱基质量稳定在 30%~40% 之间, 低质量碱基比例小, 说明测序质量较好. 根据饱和量分析可知, 5 个样本均得到了较高的基因表达量.

Table 1 Biopsy of 5 testicular tissue and the outcome of ICSI

Samples	Sperm count (per sperm / 200×field of vision)	Motile sperm (individual sperm /200×field of vision)	Vitality	Age	Outcome of ICSI
NOA	0	-	-		-
OA1	0~1	-	-		-
OA2	2~5	0	-	30±2	No pregnancy
OA3	2~5	<1	NP		No pregnancy
OA4	2~5	1~2	PR-NP		Clinical pregnancy

NOA is non-instructive testicular tissue, OA1~OA4 are obstructive testicular tissue; NP: no forward motility sperm; PR-NP: having forward motility sperm.

2.2 测序结果比对分析

将得到的序列数与人基因组 (*Homo_sapiens*: NCBI_GRCh37.p13) 比对, 样品比基因组数据库的平均比率为94.38%, 比对上的转录本为: NOA 为 28 640 个, OA1 为 30 411 个, OA2 为 30 941 个, OA3 为 30 522 个, OA4 为 30 753 个 (图1). 预测能编码蛋白质的新转录本共 2 242 个, 5 组睾丸组织分别为: 1 779、1 924、1 920、1 929 和 1 907 个 (表2).

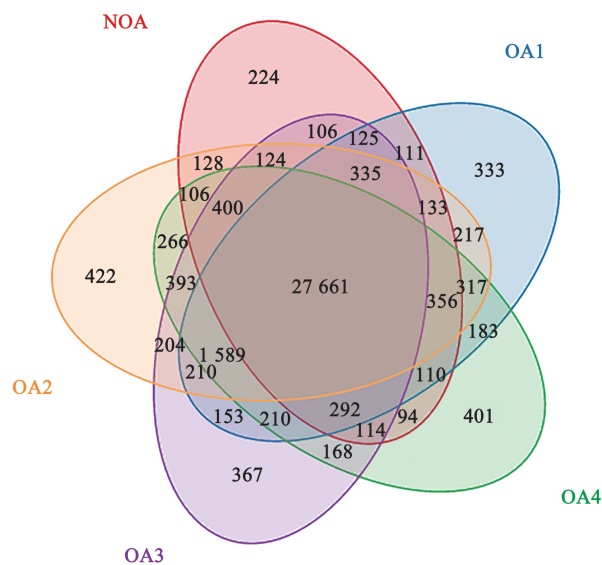


Fig. 1 Venn diagram of gene expression during *in vitro* pre-implantation development of yak crossbred embryos

Table 2 The mapped transcripts and novel transcripts across testicular tissue of azoospermia

Samples	Transcripts on the alignment	New transcripts
NOA	28 640	1 779
OA1	30 411	1 924
OA2	30 941	1 920
OA3	30 522	1 929
OA4	30 753	1 907

2.3 差异表达基因分析

得到差异表达基因统计结果为: NOA vs. OA1 总基因数是 9 195, 上调 8 045, 下调 1 150; OA1 vs. OA2 总数是 1 958, 上调 1 538, 下调 420; OA2 vs. OA3 总数是 2 965, 上调 1 275, 下调 1 690; OA3 vs. OA4 总数为 3 687, 上调 1 834, 下调 1 853 (图2). 5 组无精子症睾丸组织部分差异表达基因统计如表3~6所示.

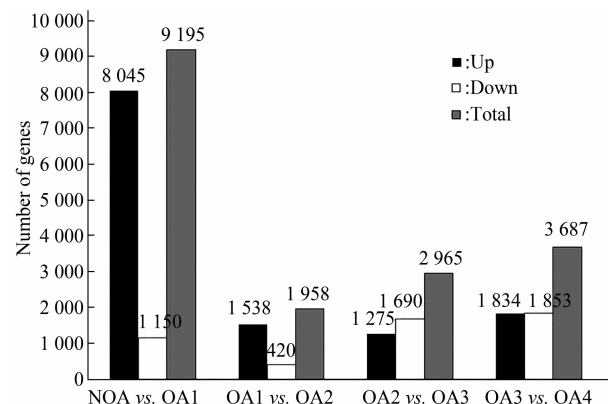


Fig. 2 Number of DE genes during testicular tissue of azoospermia

2.4 时间序列分析

图3显示的是基因聚类成各种基因簇的情况, 图4为基因簇1中表达量较高的一些代表基因.

2.5 差异表达基因的GO富集与KEGG通路分析

比较5例无精子症睾丸组织的表达谱发现, NOA 与 OA1 比对的差异表达基因 GO 分析生物过程 (biological process) 显著富集 Top5 的 GO 条目有细胞过程 (cellular process)、生物调节 (biological regulation)、代谢过程 (metabolic process)、生物调节过程 (regulation of biological process)、应激反应 (response to stimulus); 细胞

Table 3 A part of DE genes during testicular tissue of NOA vs. OA1 azoospermia

Genes	Length of genes	The magnitude of	The magnitude of	log ₂ Fold	Up/Down	P value
		NOA gene expression	OA1 gene expression			
GPX4	981	603.6	1 827.58	1.60	Up	0
TUBB4B	1 591	485.94	1 088.21	1.16	Up	0
PHF7	2 280	345.37	1 307.89	1.92	Up	0
ATP5IF1	572	294.76	656.55	1.16	Up	8.68×10 ⁻¹⁷¹
COX7A2	716	294.51	754.47	1.36	Up	0
YBX1	3 158	279.43	788.24	1.50	Up	0
UBE2S	1 209	263.47	527.96	1.00	Up	0
TSACC	668	233.24	993.73	2.09	Up	0
ALKBH7	1 159	231.01	761.69	1.72	Up	0
IZUMO4	1 016	227.28	839.2	1.88	Up	0
PTGDS	837	6 135.86	2 284.79	-1.43	Down	0
TMSB4X	657	1 700.56	421.46	-2.01	Down	0
IGFBP7	1 136	1 301.83	417.47	-1.64	Down	0
LGALS1	586	1 206.06	394.8	-1.61	Down	0
APOE	1 234	812.02	124.63	-2.70	Down	0
RPL12	674	732.41	351.98	-1.06	Down	1.12×10 ⁻²³⁷
MIF	561	716.11	349.94	-1.03	Down	3.68×10 ⁻¹⁶⁹
MIR202HG	693	637.56	161.6	-1.98	Down	0
APOA1	955	625.31	249.16	-1.33	Down	0
RPS4X	956	614.54	304.1	-1.01	Down	6.26×10 ⁻³⁰²

Table 4 A part of DE genes during testicular tissue of OA1 vs. OA2 azoospermia

Genes	Length of genes	The magnitude of	The magnitude of	log ₂ Fold	Up/Down	P value
		OA1 gene expression	OA2 gene expression			
PRM1	426	9 875.78	2 4951.89	1.87	Up	0
TNP1	415	8 580.38	1 6477.75	1.34	Up	0
PRM2	682	8 030.86	1 8100.88	1.79	Up	0
RN7SL2	299	2 487.83	4 296.18	0.82	UP	1.17×10 ⁻⁵⁴
CST3	929	912.45	1 602.11	0.48	UP	3.87×10 ⁻¹⁶⁸
LINC01921	413	401.85	1 213.5	1.49	Up	1.35×10 ⁻²⁵³
PTGDS	837	2 284.79	4 217.17	0.95	Up	0
CLU	2 017	1 289.8	1 984.39	0.49	Up	6.95×10 ⁻⁰⁸
CST3	929	912.45	1 602.11	0.48	Up	3.87×10 ⁻¹⁶⁸
RNASE1	825	759.98	1 375.07	0.28	Up	0.06997
SPANXA1	418	992.69	106.87	-2.06	Down	0
VCY	551	594.66	152.32	-2.00	Down	0
HLA-A	1 592	411.53	182.19	-1.06	Down	0
HLA-B	1 682	378.7	233.2	-1.02	Down	0
TEX101	1 176	422.62	165.56	-1.23	Down	0
LOC105375920	523	453.5	99.71	-2.70	Down	0
CXorf51A	438	477.99	48	-2.54	Down	4.42×10 ⁻²²⁶
PRSS21	1 158	378.53	140.79	-1.44	Down	0
BUD23	1 268	346.88	175.69	-1.07	Down	1.37×10 ⁻²²²
CT83	580	290.64	88.83	-1.39	Down	2.66×10 ⁻¹⁰⁴

Table 5 A part of DE genes during testicular tissue of OA2 vs. OA3 azoospermia

Genes	Length of genes	The magnitude of OA2 gene expression	The magnitude of OA3 gene expression	log ₂ Fold	Up/Down	P value
PTGDS	837	4 217.17	15 238.32	1.85	Up	0
IGFBP7	1 136	648.36	1 744.95	1.43	Up	0
RPS28	401	565.98	1 503.69	1.41	Up	6.32×10 ⁻²⁷¹
TMSB4X	657	509.32	1 436.78	1.50	Up	0
APOE	1 234	455.13	1 116.26	1.29	Up	0
RPL12	674	355.62	767.01	1.11	Up	7.32×10 ⁻²⁵¹
IFITM3	677	331.04	725.51	1.13	Up	2.15×10 ⁻²⁴⁷
VIM	1 878	265.08	593.72	1.16	Up	0
FTL	891	237.67	905.91	1.93	Up	0
TIMP1	934	233.96	613.02	1.39	Up	0
TUBA3E	1 554	349.16	2.21	-7.30	Down	0
CLDN11	2 712	161.82	78.76	-1.04	Down	1.63×10 ⁻²⁸¹
GGTLC2	951	90.92	6.67	-3.77	Down	2.40×10 ⁻²⁴⁵
GAGE12H	564	80.31	0.01	-12.97	Down	1.79×10 ⁻¹³⁵
SLC7A5	4 349	75.48	31.03	-1.28	Down	0
CYB5R1	1 675	68.86	27.72	-1.31	Down	1.14×10 ⁻⁹⁷
GATM	2 616	66.11	28.42	-1.22	Down	7.93×10 ⁻¹³⁷
GPRC5B	2 935	63	30.18	-1.06	Down	9.63×10 ⁻¹²⁰
TMEM130	2 963	61.2	18.88	-1.70	Down	2.36×10 ⁻²⁴⁷
NPIP15	1 582	58.69	10.56	-2.47	Down	3.47×10 ⁻¹³²

Table 6 A part of DE genes during testicular tissue of OA3 vs. OA4 azoospermia

Genes	Length of genes	The magnitude of OA3 gene expression	The magnitude of OA4 gene expression	log ₂ Fold	Up/Down	P value
BANF2	585	114.52	243.17	1.08	Up	1.54×10 ⁻⁵⁰
TUBA3E	1 554	2.21	223.44	6.65	Up	0
TEX35	959	89.02	206.31	1.21	Up	5.19×10 ⁻¹³⁵
OXCT2	1 824	56.85	169.72	1.57	Up	0
PRH1-PRR4	963	65.22	163.01	1.32	Up	1.38×10 ⁻¹¹⁴
ACTL10	2 028	73.56	162.67	1.14	Up	3.79×10 ⁻²³⁶
CCDC169	978	80.49	161.92	1.00	Up	7.30×10 ⁻⁸⁵
AZIN1-AS1	833	54.93	159.46	1.53	Up	4.78×10 ⁻¹²³
PSD3	1 285	60.22	147.13	1.28	Up	5.06×10 ⁻⁶²
SLFN1	2 277	53.41	131.53	1.30	Up	1.61×10 ⁻³⁰¹
PTGDS	837	15 238.32	4 400.81	-1.79	Down	0
IGFBP7	1 136	1 744.95	482.28	-1.85	Down	0
RPS28	401	1 503.69	415.12	-1.85	Down	0
TMSB4X	657	1 436.78	348.49	-2.04	Down	0
JUNB	1 832	1 314.61	298.11	-2.14	Down	0
APOE	1 234	1 116.26	398.28	-1.48	Down	0
INSL3	834	1 088.02	292.07	-1.89	Down	0
RPL21	582	942.39	465.41	-1.01	Down	1.49×10 ⁻²¹⁴
FTL	891	905.91	252.25	-1.84	Down	0
IFITM3	677	725.51	212.47	-1.77	Down	0

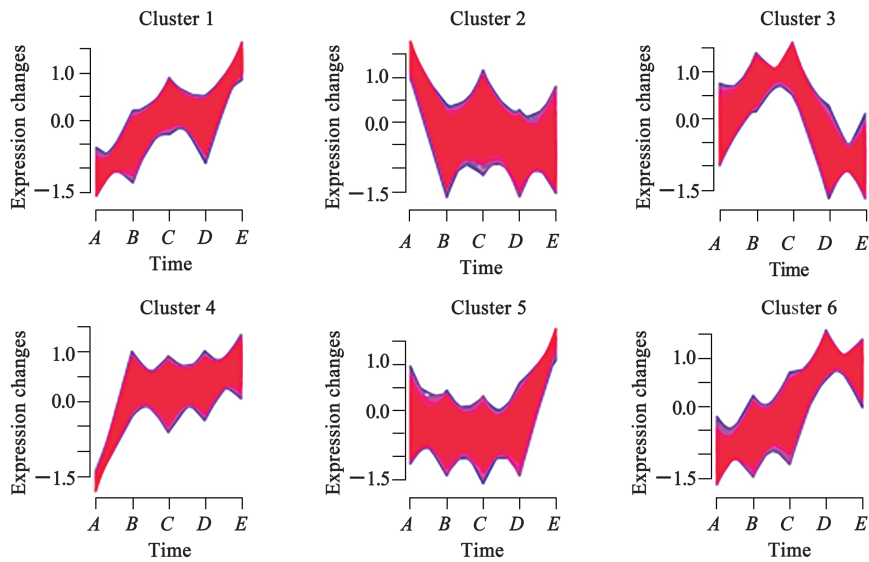


Fig. 3 Mfuzz map of gene cluster expression in testicular tissues of 5 groups

The X-axis represents five time points (A: NOA; B: OA1; C: OA2; D: OA3; E: OA4), and the Y-axis represents the value of expression after homogenization.

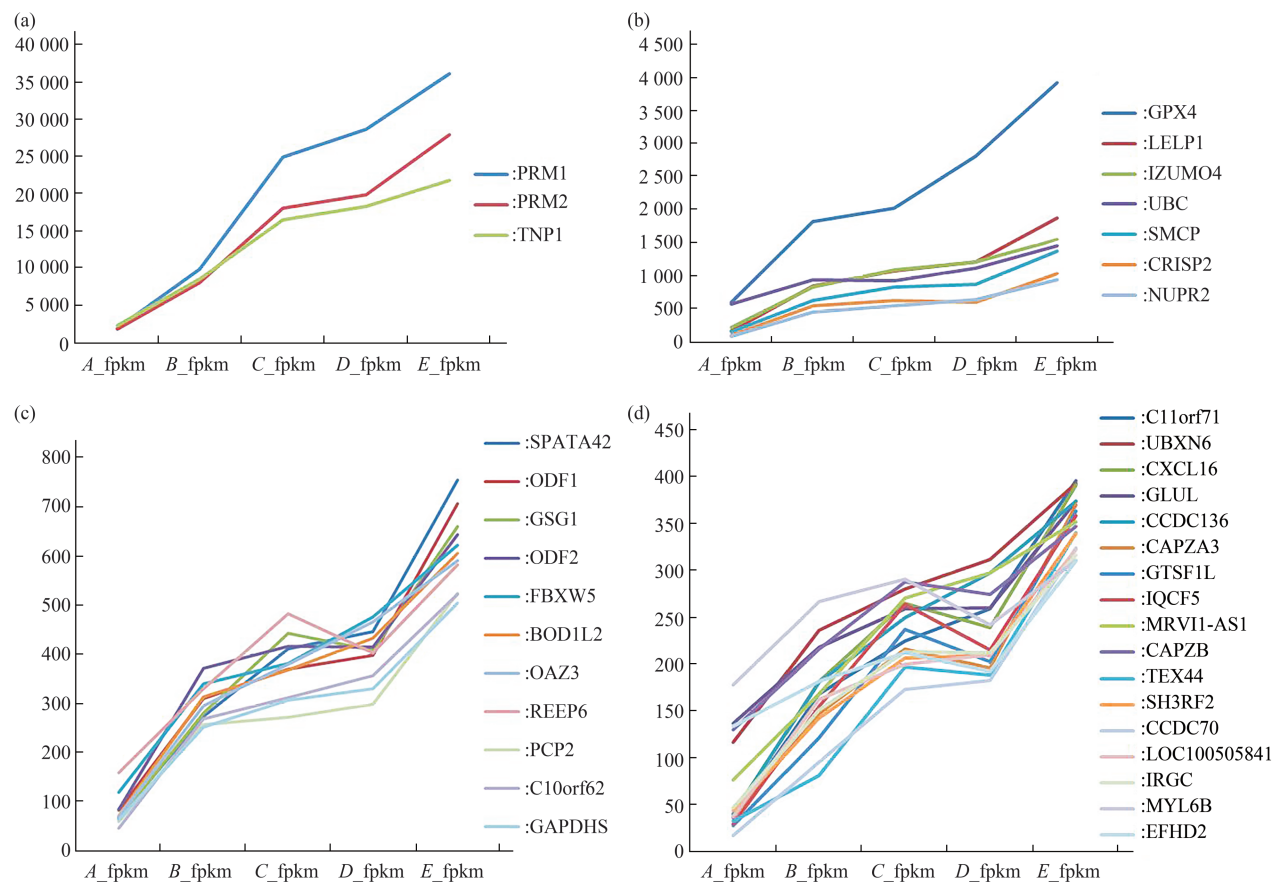


Fig. 4 Broken line diagram of gene expression in 5 groups of testicular tissues in gene cluster 1

(a)–(d) The similar value of genes expression. The X-axis represents five samples (A: NOA; B: OA1; C: OA2; D: OA3; E: OA4), and the Y-axis represents the value of genes expression (FPKM).

组分 (cellular component) 显著富集 Top5 的 GO 条目有细胞 (cell)、细胞部分 (cell part)、细胞器 (organelle)、细胞器部分 (organelle part)、胞膜 (membrane); 分子功能 (molecular function) 显著富集 Top5 的 GO 条目有绑定 (binding)、催化活性 (catalytic activity)、分子功能调节 (molecular

function regulator)、转录调节活性 (transcription regulator activity)、转运调节活性 (transcription regulator activity) (图 5)。此外, OA1 与 OA2、OA2 与 OA3、OA3 与 OA4 比对的差异表达基因富集的 GO 条目和 NOA 与 OA1 一致。

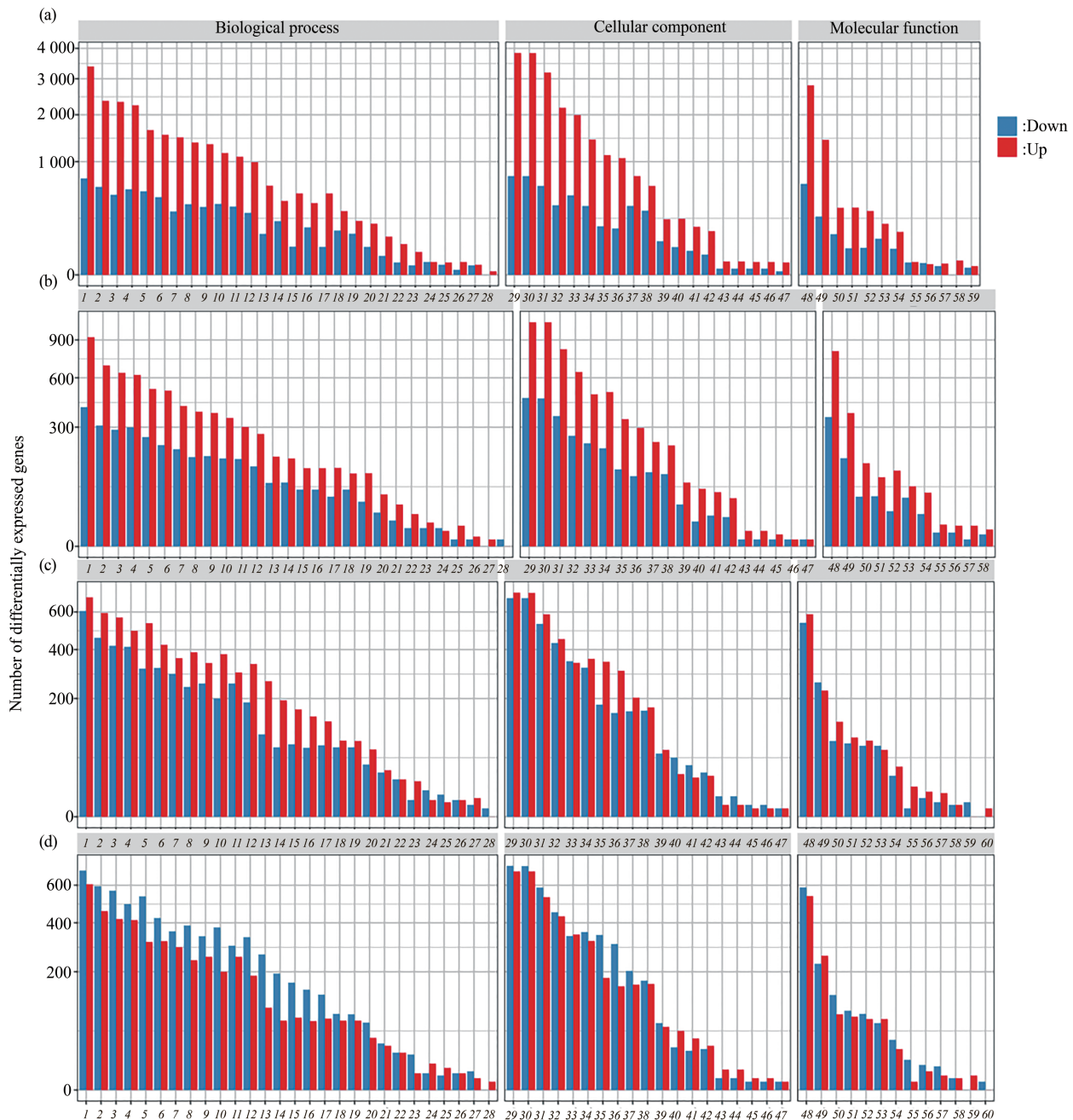


Fig. 5 GO analysis of differentially expressed genes

(a) NOA vs. OA1, (b) OA1 vs. OA2, (c) OA2 vs. OA3, (d) OA3 vs. OA4. Biological processes (1: cellular process; 2: bioregulation; 3: metabolic process; 4: bioregulation process; 5–28: stress response ~ biological stage); Cellular components (29: cells; 30: cell portion; 31: organelle; 32: organelle portion; 33–47: cell membrane ~ nucleoid); Molecular function (48: binding; 49: catalytic activity; 50: molecular function regulation; 51: transcriptional regulation activity; 52–60: transport regulation activity protein labeling) .

比较 5 例无精子症睾丸组织的差异基因 KEGG, NOA/OA1 组 Top10 通路主要富集在 RNA 降解通路 (RNA degradation)、基底细胞瘤通路 (basal cell carcinoma)、癌通路 (pathways in cancer)、黑色素生成通路 (melanogenesis)、调节干细胞多能性的信号通路 (signaling pathways regulating pluripotency of stem cells)、乳腺癌通路 (breast cancer)、河马信号通路 (hippo signaling

Table 7 KEGG pathway of 5 testicular tissues

Samples	Top10	Pathways	DE genes	Q value	ID
NOA vs. OA1	1	RNA degradation	553 (13.2%)	9.25×10^{-170}	ko03018
	2	Basal cell carcinoma	463 (11.06%)	7.06×10^{-152}	ko05217
	3	Pathways in cancer	693 (16.55%)	6.66×10^{-141}	ko05200
	4	Melanogenesis	480 (11.46%)	1.43×10^{-138}	ko04916
	5	Signaling pathways regulating pluripotency of stem cells	479 (11.44%)	8.41×10^{-136}	ko04550
	6	Breast cancer	482 (11.51%)	1.52×10^{-131}	ko05224
	7	Hippo signaling pathway	490 (11.7%)	4.93×10^{-127}	ko04390
	8	mTOR signaling pathway	483 (11.53%)	5.23×10^{-125}	ko04150
	9	Wnt signaling pathway	468 (11.17%)	1.41×10^{-115}	ko04310
	10	HTLV-I infection	522 (12.46%)	2.22×10^{-114}	ko05166
OA1 vs. OA2	1	Basal cell carcinoma	216 (12.86%)	1.77×10^{-80}	ko05217
	2	Pathways in cancer	315 (18.75%)	1.41×10^{-75}	ko05200
	3	RNA degradation	234 (13.93%)	2.20×10^{-74}	ko03018
	4	Signaling pathways regulating pluripotency of stem cells	223 (13.27%)	2.05×10^{-73}	ko04550
	5	Melanogenesis	221 (13.15%)	3.95×10^{-73}	ko04916
	6	Breast cancer	225 (13.39%)	4.93×10^{-72}	ko05224
	7	Hippo signaling pathway	229 (13.63%)	2.76×10^{-70}	ko04390
	8	mTOR signaling pathway	224 (13.33%)	4.87×10^{-68}	ko04150
	9	Proteoglycans in cancer	244 (14.52%)	6.51×10^{-66}	ko05205
	10	HTLV-I infection	243 (14.46%)	2.18×10^{-64}	ko05166
OA2 vs. OA3	1	Pathways in cancer	224 (18.14%)	1.57×10^{-50}	ko05200
	2	Basal cell carcinoma	137 (11.09%)	3.13×10^{-43}	ko05217
	3	Toxoplasmosis	104 (8.42%)	2.21×10^{-42}	ko05145
	4	Melanogenesis	146 (11.82%)	2.21×10^{-42}	ko04916
	5	Signaling pathways regulating pluripotency of stem cells	147 (11.9%)	2.21×10^{-42}	ko04550
	6	RNA degradation	152 (12.31%)	1.59×10^{-41}	ko03018
	7	HTLV-I infection	167 (13.52%)	4.10×10^{-40}	ko05166
	8	Hippo signaling pathway	147 (11.9%)	6.07×10^{-38}	ko04390
	9	Breast cancer	142 (11.5%)	1.08×10^{-37}	ko05224
	10	mTOR signaling pathway	145 (11.74%)	1.76×10^{-37}	ko04150
OA3 vs. OA4	1	Basal cell carcinoma	160 (10.82%)	6.41×10^{-49}	ko05217
	2	Pathways in cancer	246 (16.63%)	8.27×10^{-49}	ko05200
	3	HTLV-I infection	192 (12.98%)	2.66×10^{-43}	ko05166
	4	Melanogenesis	163 (11.02%)	2.66×10^{-43}	ko04916
	5	Signaling pathways regulating pluripotency of stem cells	164 (11.09%)	2.66×10^{-43}	ko04550
	6	Breast cancer	167 (11.29%)	3.69×10^{-43}	ko05224
	7	Hippo signaling pathway	172 (11.63%)	4.99×10^{-43}	ko04390
	8	Proteoglycans in cancer	187 (12.64%)	1.59×10^{-41}	ko05205
	9	mTOR signaling pathway	165 (11.16%)	8.77×10^{-40}	ko04150
	10	RNA degradation	163 (11.02%)	1.17×10^{-38}	ko03018

pathway)、mTOR 信号通路 (mTOR signaling pathway)、Wnt 信号通路 (Wnt signaling pathway)、HTLV-I 感染通路 (HTLV-I infection)。另外,和NOA/OA1组不同的是OA2/OA3组Top10通路富集弓形体病通路 (toxoplasmosis)、OA3/OA4组Top10通路富集蛋白多糖癌通路 (proteoglycans in cancer) (表7)。

2.6 融合基因

图6是显示五组样本转录本中部分基因在24条染色体和线粒体中的融合图,其中X和Y染色体未检测到融合基因。一些融合基因与无精程度呈一定的线性关系:PAICS-SRP72融合基因 (嘌呤生物合

成基因-信号识别颗粒蛋白72基因)在OA1~OA4中皆出现,NRM-MDC1 (诱变剂基因-DNA损伤检测点蛋白1基因)在NOA~OA3、冷诱导RNA结合蛋白基因CIRBP在5组中都有融合,多发性骨髓瘤基因MUM1出现在OA1~OA3中,核膜跨膜蛋白TMEM120A出现在OA2~OA4中。一些融合基因呈非线性关系,如Y-box结合蛋白1基因YBX1、细胞骨架蛋白4基因SEPT4、聚集素基因CLU等基因融合。另外,只出现在NOA中的融合基因有谷胱甘肽硫转移酶亚基2基因GSTM2和胸腺肽 β -10基因TMSB10等与线粒体基因融合 (表8)。

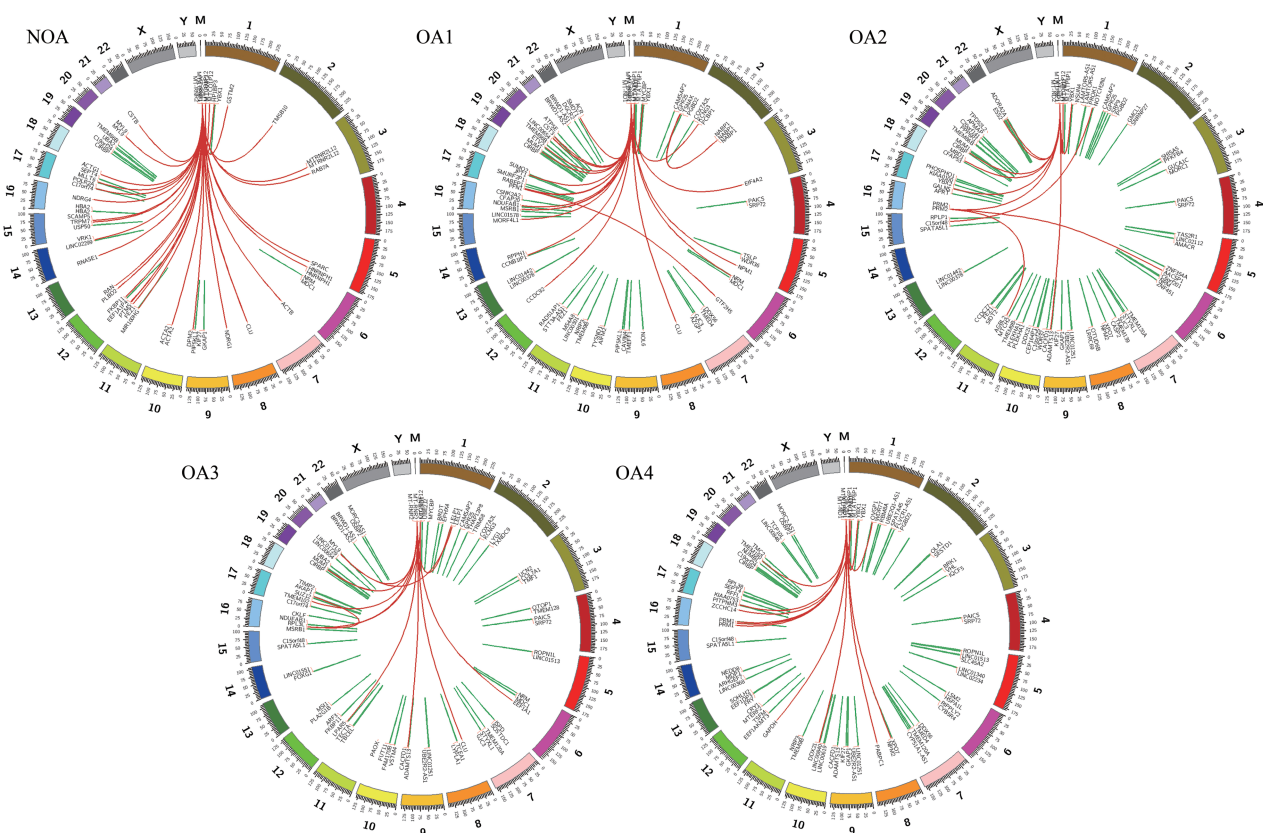


Fig. 6 The fusion gene of testicular tissue of azoospermia

The outer circle represents the chromosome information, the middle line indicates the fusion gene, the red line is the fusion gene between chromosomes, the green line is the fusion gene in the chromosome, the position of the fusion gene on the chromosome, the M is the mitochondrial transcriptome, and the 1 to Y is the 24 chromosome transcriptome.

Table 8 A part of fusion genes from NOA to OA4

Groups	NOA	OA1	OA2	OA3	OA4	Chromosome
1	-	PAICS-SRP72	PAICS-SRP72	PAICS-SRP72	PAICS-SRP72	4-4
2	NRM-MDC1	NRM-MDC1	NRM-MDC1	NRM-MDC1	-	6-6-M
3	CIRBP	CIRBP	CIRBP	CIRBP	CIRBP	19-19
4	-	MUM1	MUM1	MUM1	-	19-19-M
5	-	-	TMEM120A	TMEM120A	TMEM120A	7-7
6	YBX1	YBX1	YBX1	-	YBX1	1-1
7	SEPT4	-	-	-	SEPT4	17-17
8	CLU	CLU	-	CLU	-	8-M
9	GSTM2	-	-	-	-	1-M
10	TMSB10	-	-	-	-	2-M

the M is the mitochondrial transcriptome, and the 1 to Y is the 24 chromosome transcriptome.

3 讨 论

3.1 睾丸组织转录组特点

近年来,基因组范围的微阵列和RNA测序研究生精细胞或模型动物的睾丸样本,为人类精子发生的分子控制提供了基础^[16].研究者通过对卡尔曼综合征(Kallmann syndrome, KS)患者单个精子细胞测序与正常精子基因组序列比较,发现有数百个差异表达基因,这为KS患者临床治疗指明了方向^[17].样品在不同时间点一些基因会有相似的表达模式,根据基因的表达量信息,可以聚类成与时间相关联的基因簇,表达模式一致的基因会被聚到同一个簇,从而显示样品间的组织特异性^[18].本研究结果得到一些有意义的基因簇,随着无精症程度的加剧,一些基因呈递增或递减的表达,这说明与无精症相关的基因不是单独的,而是多个基因之间相互作用的结果.

研究者分析了1例正常人睾丸组织细胞转录组测序和1例NOA患者的174个睾丸细胞类型和基因表达特征,基于该无精子症患者睾丸中的各种细胞类型与正常人睾丸中对应的细胞类型相比发生的基因表达异常,建立基于单细胞转录组分析的无精子症分子诊断平台,为研究不育症致病机制和临床诊断提供了新的思路^[19].我们发现,与OA相比,NOA高达8 045个基因表达量降低,其中4例OA基因随发现的精子数量及活力呈时间序列变化.本研究通过检测精子发生过程中基因表达的动态变化,找出调控配子发生的关键基因,可以促进精子发生失败所致男性不育的诊断和治疗.

与生精功能相关的基因有很多,如外周致密纤维蛋白基因ODF1、ODF2和节律基因CLOCK

等^[20-22].本实验中ODF1和ODF2基因在5例睾丸组织中的表达量呈递增的趋势,表明该基因不仅在精液中与精子的运动有关,在睾丸组织中与精子的发生和运动也有关.Zhu等^[23]对27例OA患者的睾丸组织进行了活检手术,用转录组测序法测定精子发生过程中生殖细胞转录水平的变化,对参与精子发生调控的同源异型盒基因HOXs、特异性蛋白1基因SP1和转录因子3基因TCF3等潜在关键基因进行了预测,其潜在价值可作为临床应用的分子工具.本研究筛选出数千个差异表达基因,如过渡蛋白1基因TNP1、谷胱甘肽过氧化物酶4基因GPX4、精子膜蛋白4基因IZUMO4、泛素结合酶4基因UBC4、半胱氨酸的分泌蛋白2基因CRISP2和核蛋白R2基因NUPR2等,这些基因可能与精子发生和运动相关.有学者将人睾丸生精相关新基因SPAG4L进行原核表达,并对重组蛋白质进行纯化,为研究SPAG4L的生物学功能奠定基础^[24].我们预测能编码蛋白质的新转录本共2 242个,这些基因的分离纯化有利于无精症的研究.另外,本实验中,3例患者在本中心进行辅助生殖技术治疗,其中1例临床妊娠.睾丸精子活力与妊娠的关系,以及妊娠是否与睾丸组织差异基因表达相关,还需要进一步研究.

3.2 差异表达基因的GO注释及KEGG分析

研究者对正常人睾丸组织细胞转录组测序,发现在精原干细胞中BMP(bone morphogenetic protein, BMP)与FGF信号通路(fibroblast growth factor, FGF)从配体、受体到效应分子均呈现高度富集的特点.这一结果表明,BMP和FGF信号通路很可能在人类精原干细胞的自我更新过程中发挥了关键作用^[19].利用免疫组化等方法对

PIWI蛋白相互作用RNA (PIWI/piRNA) 通路的研究,发现该通路与人类无精子症相关,可能在精子发育的过程中有着重要的保护作用^[25-26].我们发现RNA降解通路显著地富集在无精子症患者睾丸组织表达谱中,说明该通路调控与精子发生密切相关.一个复杂的“mRNA降解密码”控制着动物发育过程中基因的表达.mRNA在细胞质中的降解和稳定性的调控已经得到了很好的表征,这一调控是改变控制各种生物途径的某些转录本丰度的关键步骤^[27].Wang等^[19]发现NOA患者的体细胞中高度富集细胞凋亡、氧化应激等生物学过程.我们对差异基因进行GO分析发现,除了NOA外,OA也高度富集了生物调节、代谢和应激反应等过程.另外,我们发现癌通路、基底细胞瘤通路等通路在无精子症患者中显著富集,表明无精子症可能与机体病变有很大的关系.

3.3 融合基因分析

基因融合是指两个基因的全部或一部分序列相互融合为一个新基因的过程,由2个以上基因的编码区相连,置于相同序列的调控下构成的嵌合基因.研究发现融合基因的产生机制主要包括染色体结构畸变和转录异常,形成融合基因的机制较为复杂,包括染色体易位、中间缺失、染色体倒置以及反式剪切等多种机制^[28].基因融合主要发生在癌症等病变中,我们发现癌通路在无精子症睾丸组织转录组中显著富集,说明融合基因可能与无精子症有一定的关联,深入的论证还需要进一步研究.研究表明融合基因与癌症的发生发展有很大关联,如CIRBP蛋白与细胞凋亡有关^[29],我们也发现19号染色体的CIRBP基因在5组无精子症转录组中都有染色体融合.另外,PAICS-SRP72、NRM-MDC1、MUM1和TMEM120A等融合基因在5组转录本中都呈一定的线性关系,1号染色体的GSTM2和2号染色体的TMSB10基因与线粒体基因融合并只出现在NOA转录本中,说明这些基因的融合也可能与精子发生或运动相关.我们研究发现,随着无精程度的增加,融合基因的数量是增加的,这提示我们除了染色体多态性、非编码RNA等对男性生精能力有一定的影响以外^[30-31],转录异常导致的融合基因可能也会影响精子的发生,其具体机制还需要更进一步的研究.

本研究利用高通量测序技术对无精子症患者睾丸组织的转录组进行测序和分析,揭示了不同程度无精子症患者睾丸组织差异表达基因的数量、功

能、分类和代谢通路,筛选出精子发生、精子运动等差异表达基因,丰富了无精子症患者睾丸组织转录组信息,为开展无精子症患者睾丸组织相关基因的研究及分子调控机制研究奠定基础.综上所述,无精子症可能与基因表达、信号通路以及基因融合等综合因素相关,其中可能存在决定性的因素,如决定精子发生和运动的基因及信号开关,这些还需要利用单分子测序^[6]及RNA甲基化分析^[32]等技术进行深入研究.

参 考 文 献

- [1] 江小华,余昌萍,郭通航,等.人类无精子症的遗传基础.中国科学技术大学学报,2018,48(10):814-824
Jiang X H, Yu C P, Guo T H, *et al.* J Univ Sci Tech Chin, 2018, 48(10):814-824
- [2] 李君,李彩华,郭培培,等.严重少弱精子症ICSI失败后改行供精人工授精的妊娠结局分析.安徽医科大学学报,2019,54(5):809-812
Li J, Li C H, Guo P P, *et al.* Acta Univ Med Anhui, 2019, 54(5):809-812
- [3] 李瑞,王爱玲,董建军,等.343例非梗阻性无精子症或严重少精子症患者的遗传学分析.国际遗传学杂志,2019,42(4):267-271
Li R, Wang A L, Dong J J, *et al.* Intern J Genetics, 2019, 42(4):267-271
- [4] Matzuk M M, Lamb D J. The biology of infertility: research advances and clinical challenges. Nat Med, 2008, 14(11):1197-1213
- [5] 李彤,徐家伟,孙莹璞.单细胞转录组测序在生殖发育领域应用进展.国际生殖健康/计划生育杂志,2019,38(3):217-221
Li T, Xu J W, Sun Y P. J Intern Rep Health/Family Planning, 2019, 38(3):217-221
- [6] 俞晓玲,姜文倩,郑玲,等.单分子测序技术及应用研究进展.生物化学与生物物理进展,2020,47(1):5-16
Yu X L, Jiang W Q, Zheng L, *et al.* Prog Biochem Biophys, 2020, 47(1):5-16
- [7] Chen Y, Zheng Y, Gao Y, *et al.* Single-cell RNA-seq uncovers dynamic processes and critical regulators in mouse spermatogenesis. Cell Res, 2018, 28(9):879-896
- [8] El Fekih S, Nguyen MH, Perrin A, *et al.* Sperm RNA preparation for transcriptomic analysis: review of the techniques and personal experience. Andrologia, 2017, 49(10):e12767
- [9] Zhou F, Wang R, Yuan P, *et al.* Reconstituting the transcriptome and DNA methylome landscapes of human implantation. Nature, 2019, 572(7771):660-664
- [10] 罗斌,吕自力,单旭东,等.高通量转录组测序在人类配子发生及植入前胚胎发育研究中的应用进展,生殖医学杂志,2019,28(5):575-579
Luo B, Lü Z L, Shan X D, *et al.* J Reprod Med, 2019, 28(5):575-579

- [11] 字向东, 罗斌, 夏威, 等. 基于RNA-Seq技术的牦牛体外受精胚胎发育转录组分析. 中国农业科学, 2018, **51**(8): 1577-1589
Zi X D, Luo B, Xia W, *et al.* Scientia Agricultura Sinica, 2018, **51**(8): 1577-1589
- [12] Zhu Z J, Li C, Shi Y, *et al.* Dynamics of the transcriptome during human spermatogenesis: predicting the potential key genes regulating male gametes generation. Sci Rep, 2016, **6**: 19069
- [13] Mortazavi A, Williams B A, McCue K, *et al.* Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Med, 2008, **5**(7): 621-628
- [14] Ye J, Fang L, Zheng H, *et al.* WEGO: a web tool for plotting GO annotations. Nucleic Acids Res, 2006, **34**(2): 293-297
- [15] Jia W, Qiu K, He M, *et al.* SOAPfuse: an algorithm for identifying fusion transcripts from paired-end RNA-Seq data. Genome Biol, 2013, **14**(2): R12
- [16] 刘海仙, 谭薇, 任桥, 等. Bcl-2-caspase-9 凋亡通路在哮喘小鼠睾丸组织中的相关表达. 中国免疫学杂志, 2018, **34**(7): 1064-1069
Liu H X, Tan W, Ren Q, *et al.* Chin J Immun, 2018, **34**(7): 1064-1069
- [17] D'Aurora M, Ferlin A, Di Nicola M, *et al.* Deregulation of sertoli and leydig cells function in patients with klinefelter syndrome as evidenced by testis transcriptome analysis. BMC Genomics, 2015, **16**(1): 156-164
- [18] 丁涛, 高洁. 基于基因簇判别的人类 miRNA 功能预测研究. 中国细胞生物学学报, 2016, **38**(12): 1467-1472
Ding T, Gao J. Chin J Cell Biol, 2016, **38**(12): 1467-1472
- [19] Wang M, Liu X X, Chang G, *et al.* Single-cell RNA sequencing analysis reveals sequential cell fate transition during human spermatogenesis. Cell Stem Cell, 2018, **23**(4): 599-614
- [20] 涂超峰, 袁诗敏, 蒙岚岚, 等. 非梗阻性无精子症的遗传学基础及研究进展. 生物化学与生物物理进展, 2017, **44**(6): 466-476
Tu C F, Yuan S M, Meng L L, *et al.* Prog Biochem Biophys, 2017, **44**(6): 466-476
- [21] Liang X, Cheng S, Jiang X, *et al.* The noncircadian function of the circadian clock gene in the regulation of male fertility. J Biol Rhythm, 2013, **28**(3): 208-217
- [22] 罗斌, 何维, 王世恒, 等. 弱精子症患者精子外周致密纤维 2 基因及蛋白的表达. 中华男科学杂志, 2017, **23**(11): 1002-1006
Luo B, He W, Wang S H, *et al.* National J Andr, 2017, **23**(11): 1002-1006
- [23] Zhu J, Chen G, Zhu S, *et al.* Identification of tissue-specific protein-coding and noncoding transcripts across 14 human tissues using RNA-seq. Sci Rep, 2016, **6**: 28400
- [24] 蒋先镇, 杨明刚, 邢晓为. 人睾丸新基因 SPAG4L 的原核表达与纯化. 南方医科大学学报, 2010, **30**(9): 2047-2050
Jiang X Z, Yang M G, Xing X W. J South Med Univ, 2010, **30**(9): 2047-2050
- [25] 王鑫, 李智彤, 刘默芳. Piwi/piRNA 调控异常与男性不育症. 生命科学, 2018, **30**(2): 169-177
Wang X, Li Z T, Liu M F. Chin Bull Lif Sci, 2018, **30**(2): 169-177
- [26] Gou L T, Kang J Y, Dai P, *et al.* Ubiquitination-deficient mutations in human Piwi cause male infertility by impairing histone-to-protamine exchange during spermiogenesis. Cell, 2017, **169**(6): 1090-1104
- [27] Alonso C R. A complex 'mRNA degradation code' controls gene expression during animal development. Trends Genet, 2012, **28**(2): 78-88
- [28] Wu H, Li X R, Li H. Gene fusions and chimeric RNAs, and their implications in cancer. Genes and Diseases, 2019, **6**(4): 385-390
- [29] Saito K, Fukuda N, Matsumoto T, *et al.* Moderate low temperature preserves the stemness of neural stem cells and suppresses apoptosis of the cells *via* activation of the cold inducible RNA binding protein. Brain Res, 2010, 1358: 20-29
- [30] 罗招云, 李绪杰, 林芬, 等. 非梗阻性无精子症患者 X 染色体连锁的 TEX11 基因多态性研究. 分子诊断与治疗杂志, 2020, **12**(1): 59-63
Luo Z Y, Li X J, Lin F, *et al.* J Mol Diagn Ther, 2020, **12**(1): 59-63
- [31] 刘和宇, 夏伟. 冷诱导 RNA 结合蛋白的生物学功能与信号通路. 中国生物化学与分子生物学报, 2018, **34**(7): 697-705
Liu H Y, Xia W. Chin J Biochem Mol Biol, 2018, **34**(7): 697-705
- [32] 刘恋, 张绍武, 孟佳, 等. 高通量 RNA 甲基化测序数据处理与分析研究进展. 生物化学与生物物理进展, 2015, **42**(10): 891-899
Liu L, Zhang S W, Meng J, *et al.* Prog Biochem Biophys, 2015, **42**(10): 891-899

Comparative Transcriptome Regulation Analysis of Testicular Tissue in Azoospermia Patients *Via* RNA-Seq*

LUO Bin¹⁾, LÜ Zi-li¹⁾, SHAN Xu-Dong¹⁾, ZHAO Qing-Mei¹⁾,
DU Zhen-Zhen¹⁾, ZHANG Xia¹⁾, LI Jun-Jun²⁾, LIANG Xin^{1)**}

¹⁾Center of Reproductive Medicine, School of medical and life sciences/Reproduction & Women-Children Hospital Chengdu University of Traditional Chinese Medicine, Chengdu 610041, China;

²⁾Andrology Department, School of medical and life sciences/Reproduction & Women-Children Hospital Chengdu University of Traditional Chinese Medicine, Chengdu 610041, China)

Abstract This study focuses on investigating the transcriptome differences of testicular tissues in varying degrees of azoospermia patients through analyzing the differences of differentially expressed genes (DEGs) on levels of function, classification and metabolic pathways. One testicular tissue of non-obstructive azoospermia (NOA) and four testicular tissues of obstructive azoospermia(OA) (without sperm to having sperms by biopsies) were selected for RNA extractions, library constructions and sequencings by Illumina HiSeq™ 2500, following with analyses including differential expression gene GO annotation and KEGG analysis. As a result, about 94.38% genome database were covered in the human reference genome and 2 242 transcripts were detected as predicted new protein-coding genes. Compared with NOA, 8 045 DEGs were up-regulated and 1 150 were down-regulated in OA1. For OA1 and OA2, there were 1 538 and 420 DEGs for down-regulation and up-regulation. For OA2 and OA3, there were 1 275 for up-regulation and 1 690 for down-regulation. The numbers of up-regulated (1 834) and down-regulated (1 853) DEGs between OA3 and OA4 were similar. KEGG enrichment analysis indicated that DEGs were mainly enriched in RNA degradation pathway, basal cell carcinoma pathway, cancer pathway, melanogenesis pathway and signaling pathways regulating pluripotency of stem cells. The expressions of PRM1, PRM2, TNF1, UBXN6, CXCL16, NUPR2, CCDC136 and CRISP2 genes showed an increasing trend with time-sequence specificity. In addition, different extents of gene fusions were found in 5 cases of azoospermia testicular tissue expressed genes. In summary, the number of DEGs and their function, classification and metabolic pathways were obtained with different degrees of azoospermia. The DEGs such as spermatogenesis and sperm motility were screened out, which enriched transcriptome information for azoospermia and laid a foundation for the research of testicular tissue related genes. It may provide theoretical basis for revealing the molecular mechanisms of spermatogenesis in azoospermia patients and promoting the development of male infertility research. Notably, gene fusion may be associated with azoospermia.

Key words azoospermia, testicular tissues, transcriptome sequencing, assisted reproduction, differentially expressed gene

DOI: 10.16476/j.pibb.2020.0092

* This work was supported by grants from Chengdu University of Traditional Chinese Medicine Foundation Hospital special fund (2017-EL-25), Sichuan Science and Technology Department Project (2018YSZ0028), Sichuan Health Committee Project (19PJ033), Sichuan Administration of traditional Chinese Medicine Research Project (2018JC010) and Chengdu University of Traditional Chinese Medicine Experimental Technology Project (Z1605).

** Corresponding author.

Tel:17313164091, E-mail: 184516824@qq.com

Received: April 7, 2020 Accepted: May 12, 2020