



湖栖鳍虾虎鱼性腺转录组比较分析

董忠典, 黎学友, 黄承勤, 张海瑞, 黄顺楷,
张 宁, 郭昱嵩, 王中铎*

(广东海洋大学水产学院, 南海水产经济动物增殖重点实验室, 广东 湛江 524088)

摘要: 为获得湖栖鳍虾虎鱼性腺基因表达谱, 筛选性别决定、配子发生和繁殖相关基因, 实验对其卵巢和精巢分别进行了 RNA-Seq 分析。测序原始数据经 de novo 拼接组装共获得 62 573 个单基因 (unigene), N50 和平均长度分别为 3 082 bp 和 1 869 bp。在 NR、NT、SwissProt、PFAM、KOG、GO 及 KEGG 数据库中对上述获得的 unigene 进行注释, 分别有 41 480、32 848、37 500、35 394、18 318、35 394 和 27 009 个 unigene 获得注释。基因表达差异分析显示, 在湖栖鳍虾虎鱼卵巢和精巢中存在 10 954 个差异基因 (different expression gene, DEG), 其中 2 383 个 DEG 表达水平在精巢中上调, 8 571 个 DEG 在卵巢中上调。对 14 个 DEG 进行实时荧光定量 PCR (qPCR) 验证, 结果显示, qPCR 结果与 RNA-Seq 分析一致。在差异基因中, 筛选到了部分参与配子发生与成熟的基因, 包括 *foxl2*、*dmrt1*、*cyp19a1a*、*inha*、*inhb*、*sycp2*、*zglp1*、*tdrp*、*zps* 和 *esra*。差异基因富集结果中包含“泛素介导蛋白水解”、“卵母细胞减数分裂”、“孕酮介导卵母细胞成熟”、“p53 信号通路”、“类固醇激素合成通路”和“PI3/AKT 信号通路”等参与性腺发育及繁殖的 KEGG 通路。最后, 从湖栖鳍虾虎鱼性腺转录组中共获得 38 550 个简单重复序列标记 (simple sequence repeats, SSR) 和 192 450 个单核苷酸多态性 (single nucleotide polymorphism, SNP) 标记。研究结果有助于对湖栖鳍虾虎鱼性别相关基因的功能研究和性别相关分子标记的开发。

关键词: 湖栖鳍虾虎鱼; RNA-Seq; 配子发生; 繁殖

中图分类号: S 917.4

文献标志码: A

鱼类的性别由遗传基因和环境因子共同决定, 目前仅有少数鱼类明确了性别决定基因, 如日本青鳉 (*Oryzias latipes*) 的 *dmy*^[1]、虹鳉 (*Onchorhynchus mykiss*) 的 *sdv*^[2]、东方红鳍鲌 (*Takifugu rubripes*) 的 *amhr2*^[3]、哈奇氏牙汉鱼 (*Odontesthes hatcheri*) 的 *amh*^[4]、吕宋青鳉 (*Oryzias luzonensis*) 的 *gsdf*^[5] 和半滑舌鳎 (*Cynoglossus semilaevis*) 的 *dmrt1*^[6-7]。此外, 一些性别相关基因 (*cyp19a1a*、*foxl2*、*fox3*、*igf3*、*sox9* 和 *sf1* 等)^[8-10] 和生物学过

程 (TGF- β 、Wnt 和雌激素信号通路等) 在鱼类的性别决定中也发挥重要作用^[11-14]。鱼类种类繁多, 性别决定多样, 为更好地了解鱼类的性别决定机制, 需要在多种鱼类中进行探索。

RNA-Seq 可快速获得待测样品在特定条件下的基因表达谱, 已成功应用于鱼类性别相关基因的筛选和调控机制的研究。同样, 已利用 RNA-Seq 在多种鱼类中确定了与性腺发育和配子发生相关的基因及其表达模式^[8, 12-13, 15-16]。

收稿日期: 2020-03-30 修回日期: 2020-06-01

资助项目: 国家自然科学基金 (41806195, 31201996); 广东省科技厅项目 (2017A030303075); 广东海洋大学创新强校项目 (230419069, 230419055); 广东海洋大学博士启动项目

第一作者: 董忠典 (照片), 从事鱼类分子育种及环境毒理方面研究, E-mail: zddong@gdou.edu.cn

通信作者: 王中铎, E-mail: aduofa@hotmail.com



湖栖鳍虾虎鱼 (*Gobiopertus lacustris*) 是一种分布于印度洋和太平洋沿岸海域的小型鱼类, 对盐度的适应能力较强, 鱼体几乎透明, 器官清晰可见, 是鱼类发育、繁殖等领域潜在的理想研究模型。本课题组前期调查发现, 湖栖鳍虾虎鱼广泛分布于中国南海沿岸红树林水域^[17-19]。目前, 湖栖鳍虾虎鱼的研究尚处在资源调查和遗传多样性分析阶段, 缺乏遗传信息, 限制了湖栖鳍虾虎鱼资源的开发和利用。本研究采用 RNA-Seq 对湖栖鳍虾虎鱼性腺进行了测序、基因功能注释和表达谱分析, 为开发其基因资源提供参考。

1 材料与方法

1.1 湖栖鳍虾虎鱼性腺发育阶段鉴定及 RNA 提取

实验用性成熟湖栖鳍虾虎鱼 [体长 (18.8 ± 0.5) mm, 体质量 (0.11 ± 0.03) g] 采集自广东湛江高桥国家级红树林保护区 (N21°36'24", E109°47'8"), 于实验室暂养 3 个月 (盐度 15) 后用于样品采集。为了确定湖栖鳍虾虎鱼性腺的发育阶段, 对卵巢和精巢进行了切片观察。由于鱼体小, 湖栖鳍虾虎鱼用 20 μ g/L MS222 麻醉后直接置于波恩氏液中固定 18 h、随后进行乙醇溶液系列脱水、二甲苯透明、石蜡包埋、切片 (厚度为 8 μ m)、苏木精—伊红 (H.E) 染色后置于光学显微镜下观察并拍照。从 5 尾雌鱼中解剖卵巢混合后于液氮中速冻, 获得卵巢样本 (FG), 用于卵巢总 RNA 提取; 利用相同方法获取精巢样本 (MG)。参照 TRIzol 试剂 (Invitrogen, 美国) 操作说明提取卵巢和精巢总 RNA, 并用 DNase II (宝日医生物技术 (北京) 有限公司) 去除基因组 DNA 污染。RNA 质量和完整性分别用 1% 的琼脂糖凝胶电泳和 RNA Nano 6000 (Agilent Technologies, 美国) 检测, 使用 Qubit® 2.0 Fluorometer (Life Technologies, 美国) 测定总 RNA 浓度。

1.2 cDNA 文库构建和测序

取卵巢和精巢总 RNA 各 1.5 μ g 用于各自 cDNA 文库构建和测序。文库构建和 RNA-Seq 测序由北京诺禾致源生物信息科技有限公司完成, 文库构建和测序策略参考 Nan 等^[20]。测序原始数据 (raw data) 已上传到国家基因库生命大数据平台 (China National GenBank database, CNGBdb: <https://db.cngb.org/>), 登录号为 CNP0000359。

<http://www.china-fishery.cn>

1.3 测序数据组装及基因功能注释

为保证 RNA-Seq 分析质量, 首先对 raw reads 进行过滤, 去除低质量、接头污染以及未知碱基 N (N 表示无法确定碱基信息) 含量过高的 reads, 得到干净数据 (clean reads) 用于后续拼接组装分析。采用 Trinity (v2.5.1 版本) 软件对 clean reads 进行 de novo 组装获得转录本^[21], 对 Trinity 拼装的转录本进行去冗余, 取每个转录本聚类中最长的转录本为单基因 (unigene), 作为后续分析的参考序列。为获得全面的基因功能信息, 使用 7 大数据库对 Trinity 拼接获得的 unigene 进行功能注释 (E-value 的阈值为 1.0×10^{-5})。通过在线 BLAST (<http://www.ncbi.nlm.nih.gov>) 对 unigene 进行 Nt (NCBI nucleotide sequences) 注释; 利用 diamond (v0.8.22 版本) 对 unigene 进行 Nr (<http://www.ncbi.nlm.nih.gov>)、KOG (euKaryotic Ortholog Groups)、Swiss-Prot 注释, 利用 KAAS 进行 KEGG (Kyoto Encyclopedia of Genes and Genomes) 注释^[22], 利用 HMMER 3.0 进行 Pfam (Protein family) 注释^[23-24], 采用 Blast 2 GO 软件 (v2.5 版本) 对 NR 和 Pfam 注释结果进行 GO (Gene Ontology) 注释^[25]。

1.4 基因表达和差异分析

将 Trinity 拼接获得的转录组作为参考序列, 通过 RSEM 软件的 bowtie 2 程序将每个样本的 clean reads 比对到组装好的参考序列上^[26]。根据比对结果得到每个基因的测序数目 (read count), 采用 TMM 软件对 read count 数据进行标准化处理, 再用 DEGseq 软件进行差异分析, 筛选阈值为 $|\log_2(\text{Fold Change})| > 1$ 且 $q\text{-value} < 0.005$ ^[27]。使用 Goseq 和 KOBAS 软件分别对差异基因进行富集分析 (corrected $P\text{-value} < 0.05$)。

1.5 实时荧光定量 PCR (qPCR) 验证

利用 qPCR 对 14 个差异基因在性成熟湖栖鳍虾虎鱼性腺中的表达水平进行检测 (表 1)。qPCR 反应体系 15.0 μ L, 包含 7.5 μ L $2 \times$ SYBR Premix Ex Taq II [宝日医生物技术 (北京) 有限公司], 0.6 μ L 上下游引物 (10 μ mol/L), 1.5 μ L cDNA, 4.8 μ L ddH₂O。每个样品技术重复 3 次, 反应于 Roche Light Cycler 96 定量仪 (罗氏, 瑞士) 完成。反应程序: 预变性 94 °C 30 s; 95 °C 5 s, 58 °C 30 s, 72 °C 30 s (采集荧光), 40 个循环; 熔解曲线分析 PCR 产物的特异性。eef1b 和 hprt1 作为内参基因, 采用 $2^{-\Delta\Delta C_t}$ 方法分析基因相对表达水平^[28]。使用

中国水产学会主办 sponsored by China Society of Fisheries

SPSS 17.0 软件对结果进行统计， $P<0.05$ 表示差异具有统计学意义。对其中 9 个基因的转录本在雌雄各组织 (脑、肝脏、肠、性腺、肌肉和眼睛) 中的表达情况进行定性检测。

表 1 引物序列及扩增效率
Tab. 1 Primers and PCR efficiency of candidate genes

基因ID gene ID	引物名称 primer name	序列(5'-3') sequence	扩增效率/% efficiency	相关系数(R^2) correlation coefficient
cluster-9641.24201	<i>ef1b</i>	F: AGACTTTGCGATAACAGCCG R: GACGATGATGACCTTGACCTG	94.55	0.999 0
cluster-9641.33084	<i>hprt1</i>	F: CGGTCACTGTTTCTATTACGCG R: CTGGAGCGAGTTTACATTCTCTC	98.01	0.998 6
cluster-9641.47188	<i>e3</i>	F: CCATAAAGAGCAGCAGCAGC R: AGGCTACTCCGTCCACTACAG	93.69	0.999 1
cluster-9641.38934	<i>krt2</i>	F: AGAGCATCAACCTTTGCCTC R: AATGAACTGGAGAGCGAGC	97.17	0.999 9
cluster-9641.47528	<i>dmrt1</i>	F: GCAGACTATTGTTGTCCCTCAGG R: AGGAGAAACGAACCCGTGG	92.88	0.999 7
cluster-9641.3059	<i>cas3</i>	F: TTCATCAATGCTTGTCGTGG R: AGAAAGTCTGCCTCGTTTCGG	98.23	0.998 8
cluster-9641.43183	<i>gapdh</i>	F: TTGTGATGGGTGTCAACCAC R: CTGTGATGGCGTGAACGTG	94.99	0.995 1
cluster-9641.3821	<i>test</i>	F: GCGATTACATCATCATCACACG R: AAAGAGCGGACCAAACTGAG	97.96	0.999 5
cluster-9641.6064	<i>inhb</i>	F: TGTGACCTTACGACGCAGAC R: TTCGCAGAATCAGATGAACG	97.76	0.998 5
cluster-9641.35558	<i>tdrp</i>	F: TGAAAGAGAGGAACAGCGAG R: GAGGCGTAATCCACCAACAC	94.93	0.999 8
cluster-9641.24002	<i>rpl7</i>	F: TTGGCAACTTCTACGTCCC R: ACACACCATTGAAAATCTGGC	92.14	0.999 4
cluster-9641.11433	<i>inha</i>	F: GGTGAACCAGAACAATCCAAG R: AGTGACTATCGTTTGACCTGAAGG	97.83	0.997 9
cluster-2179.0	<i>cyp19a1a</i>	F: CCAACAGACTCTTCTGGATG R: CTCAACTAACTCTCTATGGCATTTC	95.93	0.998 1
cluster-9641.16580	<i>foxl2</i>	F: ATCACAAACCACAACCAGAGTAGC R: GCGAAACCATAGAAAGTCCAAG	99.52	0.999 5
cluster-9641.24530	<i>sycp2</i>	F: TGCTCTCTTCAGGACTCACC R: CGTGCGACACAGATACATAAGG	96.20	0.991 0
cluster-9641.25293	<i>zglp1</i>	F: TTCACCAAGTGGCAGAAATCAC R: ATGACTCCAAGAGCAGTGC	97.43	0.999 3

1.6 简单重复序列标记 (simple sequence repeats, SSR) 和单核苷酸多态性标记 (single nucleotide polymorphism, SNP) 检测

采用在线软件 MISA (<http://pgrc.ipk-gatersleben.de/misa/misa.html>) 对湖栖鲮虾虎鱼性腺转录组拼接获得的 unigene 进行 SSR 检测，设置各检测单元的最少重复次数分别为 1-10、2-6、3-5、4-5、5-5 和 6-5 (1-10 表示以单核苷酸为重复单位时，其重复数至少为 10 才可被检测到；2-6 表示以双核苷酸为重复单位时，其最少重复数为 6；3-5、4-5、5-5 和 6-5 分别表示以三核苷酸、四核苷酸、五核苷酸和六核苷酸为重复单位时，其最少重

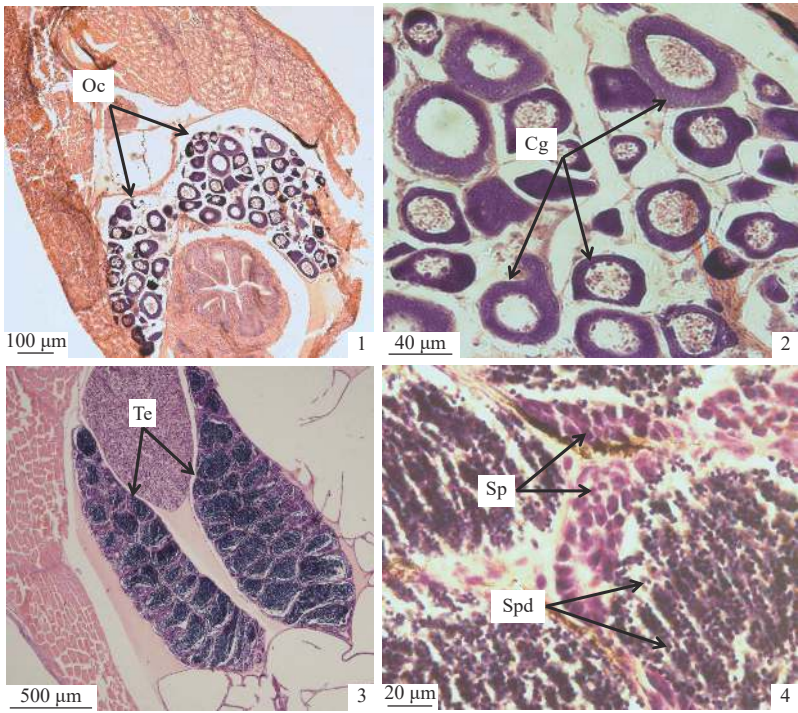
复数为 5)。以所有 unigene 序列作为参考，通过变异检测软件 samtools 1.4 进行 SNP 检测。

2 结果

2.1 性腺切片观察和转录组 de novo 拼接组装

为了明确湖栖鲮虾虎鱼性腺发育阶段，实验制备了雌、雄湖栖鲮虾虎鱼性腺切片。结果显示，湖栖鲮虾虎鱼的性腺对称分布于腹腔的两侧，在卵巢中可以观察到成熟卵母细胞，在精巢中可以观察到精子细胞，说明湖栖鲮虾虎鱼已发育至性成熟 (图版 1)。

本实验分别构建了湖栖鲮虾虎鱼卵巢和精



图版 I 湖栖鳍虾虎鱼性腺切片

1~2. 卵巢石蜡切片, 3~4. 精巢石蜡切片; Oc. 卵巢腔, Cg. 卵母细胞, Te. 精巢, Sp. 精母细胞, Spd. 精子细胞

Plate I Histological examination of *G. lacustris* gonad

1-2. The ovary of *G. lacustris*, 3-4. the testis of *G. lacustris*; Oc. ovarian cavity, Cg. oocyte, Te. testis, Sp. spermatocyte, Spd. spermatid

巢 cDNA 文库, 通过 Illumina 测序分别获得了 63 391 676 和 54 542 800 个 raw reads, 经过质控后分别获得 60 657 644 和 52 016 136 个 clean reads (表 2)。经 Trinity 组装, 共得到 62 573 个 unigene,

其中 35 081 个长度大于 1 000 bp (图 1), N50 (按照长度将拼接的转录本从大到小排序并累加转录本的长度, 到不小于总长 50% 的拼接转录本的长度) 长度为 3 082 bp。

表 2 RNA-Seq 数据统计

Tab. 2 Statistics summary of RNA-Seq data

样本 samples	原始数据/个 raw reads	干净数据/个 clean reads	干净碱基/bp clean bases	错误率% error	Q20/%	Q30/%	GC含量/% GC
卵巢 FG	63 391 676	60 657 644	9.1G	0.02	96.03	90.26	44.97
精巢 MG	54 542 800	52 016 136	7.8G	0.02	96.54	91.05	47.95

注: Q20和Q30分别代表碱基被测错的概率为1%和0.1%

Notes: Q20 and Q30 represent 1% and 0.1% probabilities of base errors, respectively

2.2 基因注释和分类

使用 NR、NT、SwissProt、PFAM、KOG 数据库对上述 62 573 个 unigene 进行注释, 分别有 41 480、32 848、37 500、35 394 和 18 318 个 unigene 获得注释。与 NR 数据库比对发现, 有 40 726 个 unigene 与 236 个物种的已知基因高度同源, 其中 64.77%(26 379 个) 的 unigene 和大弹涂鱼 (*Boleophthalmus pectinirostris*) 高度同源。为进一步了解湖栖鳍虾虎鱼性腺表达基因的功能, 将 62 573

个 unigene 在 GO 和 KEGG 数据库中进行了比对分析。35 394 个 unigene 被归类到“生物学过程”、“分子功能”和“细胞成分”三大 GO 分类的 56 个功能分类中, 其中“细胞过程”(21 652 个 unigene), “结合”(22 407 个) 和“细胞”(12 906 个) 分别在“生物学过程”、“分子功能”和“细胞成分”中包含最多的 unigene (图 2)。根据参与的 KEGG 代谢, 可将 27 009 个 unigene 分为 5 个分支 26 个二级通路, 其中“信号转导”包含最多的 unigene (表 3)。

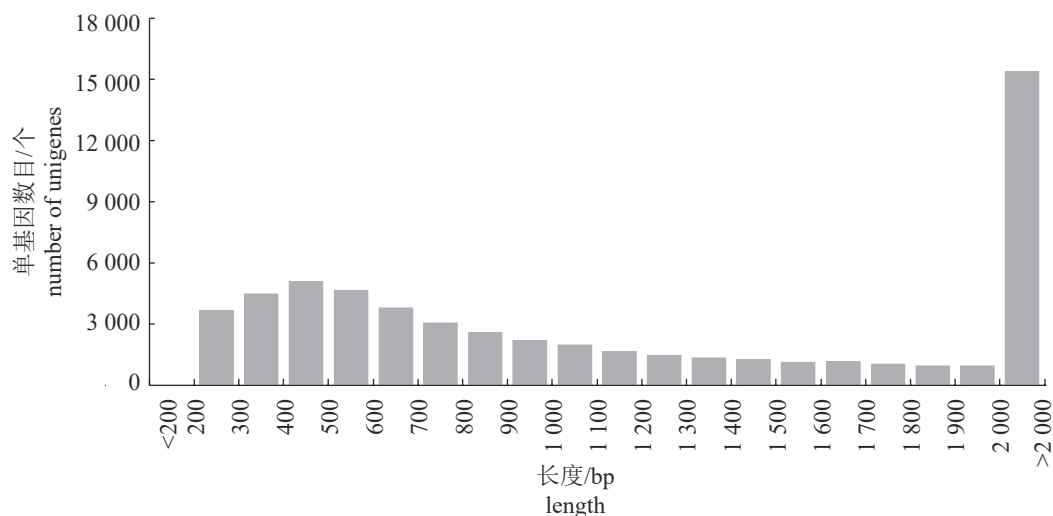


图1 unigene 长度分布

Fig. 1 Length distribution of assembled unigene

2.3 差异基因表达分析

湖栖鲮虾虎鱼卵巢和精巢 unigene 表达差异分析表明, 8 571 个 unigene 在卵巢中上调, 2 383 个 unigene 在精巢中上调 (图 3)。多个与鱼类配子发生、成熟及繁殖相关的基因 (*dmrt1*、*foxl2*、*cyp19a1a*、*tdrp*、*zglp1* 和 *inhs* 等) 被鉴定 (表 4)。差异基因 GO 富集分析结果表明, “细胞进程”在生物学过程中包含最多 unigene; 在细胞成分中, “细胞”和“细胞部分”含有最多的 unigene; 而在分子功能中, “结合”含有最多的 unigene (图 4)。同时, 有 3 618 个差异基因富集到 298 个 KEGG 通路中 (表 5), 其中“卵母细胞减数分裂”(oocyte meiosis) 和“孕酮介导卵母细胞成熟”(progesterone-mediated oocyte maturation) 等通路与生殖关系密切。

2.4 qPCR 验证和组织表达检测

利用 qPCR 对 14 个差异基因的 RNA-Seq 进行了验证, 结果显示, qPCR 结果与 RNA-Seq 分析一致 (表 6), 说明 RNA-Seq 分析结果可信。通过反转录 PCR 检测了 9 个基因在雌雄湖栖鲮虾虎鱼不同组织中的表达情况: *inha*、*foxl2*、*zglp1*、*sycp2* 和 *cyp19a1a* 主要表达于雌鱼卵巢和肠组织; *dmrt1* 主要在精巢中表达; *cas3*、*inhb* 和 *test* 在雌、雄多个组织中都有表达。

2.5 SSR 和 SNP 检测

利用 MISA 软件对湖栖鲮虾虎鱼进行 SSR

鉴定, 结果显示, 在 20 517 个 unigene 中共含有 38 550 个 SSR, 其中 8 730 个 unigene 含有一个以上的 SSR。6 种类型 SSR 出现的频率不同, 单核苷酸重复 SSR 出现的频率最高, 其次是二核苷酸重复 SSR。另外, 不同重复序列的相对丰度差异很大: A/T 在单核苷酸 SSR 中最常见; 在双核苷酸 SSR 中 AG/CT 最为常见; AGG/CCT、AAAT/ATTT、AAGGC/CCTTG 和 ACCTGG/AGGTCC 是三、四、五和六核苷酸 SSR 中最常见的元件。可依据这些 SSR 位点侧翼序列设计特异性引物, 用于湖栖鲮虾虎鱼遗传多样性分析。另外, SNP 分析结果显示, 湖栖鲮虾虎鱼雌、雄性腺转录组中共有 192 450 个 SNP, 其中 C/T、A/G 转换分别为 58 914 和 58 476 个; A/T、A/C、T/G 和 C/G 颠换分别为 24 732、18 095、18 444 和 13 789 个。

3 讨论

为了阐明湖栖鲮虾虎鱼性腺基因表达模式, 探究性腺发育及繁殖机制, 本实验首次构建了湖栖鲮虾虎鱼卵巢和精巢 cDNA 文库, 并通过 Illumina 平台进行了 RNA-Seq 测序分析。N50 长度是评价转录组数据 de novo 组装质量的重要参数, 本实验中湖栖鲮虾虎鱼性腺转录组 N50 长度为 3 082 bp, 高于多个已发表的鱼类性腺转录本组装结果, 包括牙鲈 (*Paralichthys olivaceus*) (809 bp)^[8]、多鳞鳉 (*Sillago sihama*) (2 190 bp)^[13]、金钱鱼 (*Scatophagus argus*) (3 038 bp)^[16]。有 23.47%

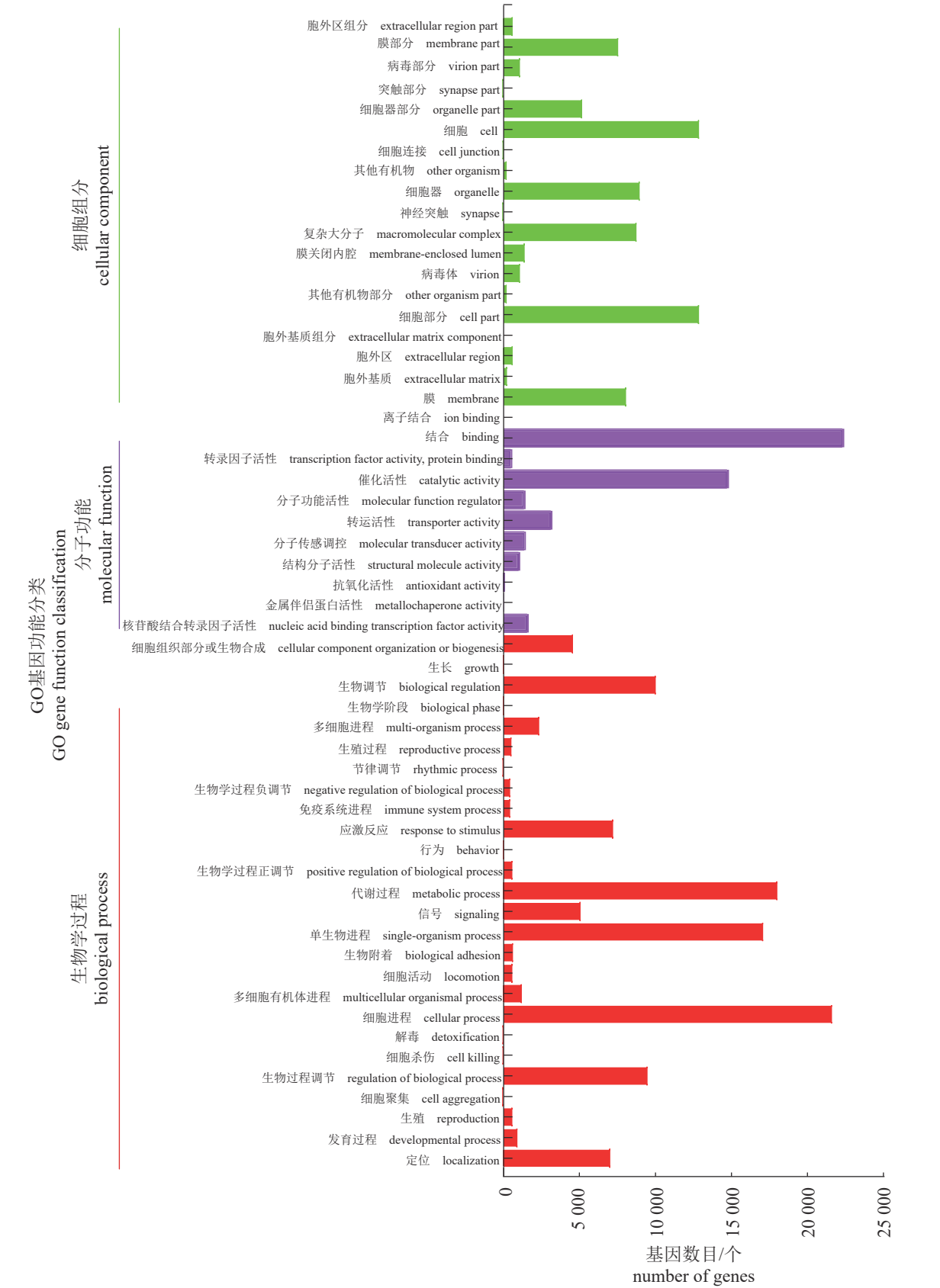


图 2 湖栖鳊虾虎鱼 unigene 的 GO 功能注释

Fig. 2 Gene ontology (GO) assignment of assembled unigenes of *G. lacustris*

表 3 湖栖鲮虾虎鱼性腺 unigene KEGG 富集分类
Tab. 3 Gonad unigene KEGG enrichment classification of *G. lacustris*

一级通路 pathway hierarchy 1		二级通路 pathway hierarchy 2	基因数目/个 gene number
细胞过程 cellular processes		细胞生长和死亡 cell growth and death	728
		细胞运动 cell motility	516
		细胞联盟 cellular community	1 536
		转运和分解代谢 transport and catabolism	1 509
环境信息处理 environmental information processing		膜运输 membrane transport	171
		信号转导 signal transduction	3 794
		信号分子与相互作用 signaling molecules and interaction	885
遗传信息处理 genetic information processing		折叠、分类和降解 folding, sorting and degradation	1 068
		复制与修复 replication and repair	295
		转录 transcription	416
		翻译 translation	1 185
代谢 metabolism		氨基酸代谢 amino acid metabolism	700
		其他次生代谢的生物合成 biosynthesis of other secondary metabolites	22
		碳水化合物代谢 carbohydrate metabolism	832
		能量代谢 energy metabolism	275
		多糖的生物合成与代谢 glycan biosynthesis and metabolism	471
		脂质代谢 lipid metabolism	826
		辅因子和维生素代谢 metabolism of cofactors and vitamins	354
		其他氨基酸代谢 metabolism of other amino acids	215
		萜类和聚酮化合物的代谢 metabolism of terpenoids and polyketides	63
		核苷酸代谢 nucleotide metabolism	388
		概述 overview	505
		外源物生物降解和代谢 exobiotics biodegradation and metabolism	153
有机系统 organismal systems		循环系统 circulatory system	790
		发育 development	667
		消化系统 digestive system	1 038
		内分泌系统 endocrine system	1 961
		环境适应 environmental adaptation	390
		排泄系统 excretory system	389
		免疫系统 immune system	1 556
		神经系统 nervous system	1 215
		感觉系统 sensory system	323

(14 682 个) 的 unigene 未获得功能注释，这可能是由于缺乏湖栖鲮虾虎鱼的基因组信息，或者

这些 unigene 包含 3'或 5'非翻译区、非编码 RNA，或属于新基因，或包含未知蛋白质结构域等^[29-30]。

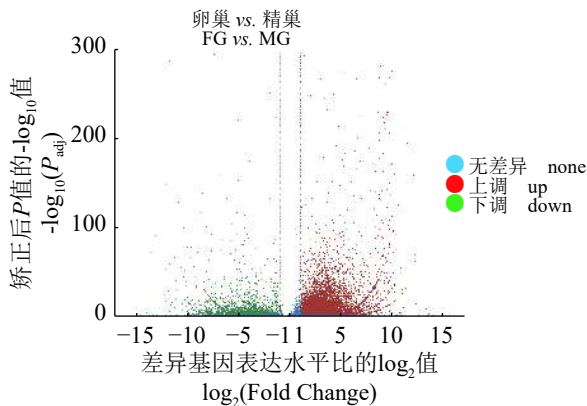


图3 湖栖鲮虾虎鱼性腺转录组差异基因火山图

Fig. 3 Volcanoplot of *G. lacustris* gonadal transcriptome differential genes

有 64.77% 的 (26 379 个) unigene 在 NR 数据库比对到与大弹涂鱼基因高度同源, 这是因为湖栖鲮虾虎鱼和大弹涂鱼在进化上亲缘关系较近, 均属于虾虎鱼科 (Gobiidae)^[31]。GO 分析显示, 湖栖鲮虾虎鱼性腺转录组序列在 GO 数据库中分别富集在“细胞过程”、“细胞”和“结合”三大功能类别中, 这与其他鱼类性腺 RNA-Seq 结果相一致, 表明这些类别的基因在功能上较为保守^[12-13, 29]。

3.1 湖栖鲮虾虎鱼配子发生及成熟相关基因

在湖栖鲮虾虎鱼卵巢和精巢差异基因中筛选到多个与配子发生及成熟等相关的基因, 包括 *dmrt1*、*foxl2*、*gtsf1*、*sycp2*、*soxs*、*zglp1*、*tdrp*、*cyp19a1a* 和 *zps* 等 (表 4)。Dmrt1 (doublesex-and mab-3-related transcription factor 1) 对鱼类雄性性别决定至关重要, 在金钱鱼和尼罗罗非鱼中 *dmrt1* 决定了雄鱼的性别^[6, 32-33], 该基因编码的转录因子具有维持雄鱼精子的作用, 常染色体上的 *dmrt1* 突变可以导致日本青鳉雌性性逆转^[34]。在斑马鱼中, *dmrt1* 突变的雄鱼通过抑制精子的产生以及促进生殖细胞凋亡而引起精巢发育缺陷^[35]。Foxl2 (forkhead box protein L2) 在脊椎动物雌性性腺分化过程中抑制了雄性遗传途径, 同时也维持卵巢的功能。敲除尼罗罗非鱼 (*Oreochromis niloticus*) 的 *foxl2* 后, 遗传雌性性逆转为生理雄鱼^[36]。斑马鱼含有 2 个 *foxl2* (*foxl2a* 和 *foxl2b*), 其中任何 1 个基因发生突变不影响卵巢发育和卵母细胞的发生, 但若 2 个基因同时发生突变则会导致斑马鱼从雌性向雄性性逆转^[37]。本研究, *dmrt1* 转录产物在性成熟湖栖鲮虾虎鱼精巢中高

表达, 而 *foxl2* 在卵巢中高表达, 二者可能在湖栖鲮虾虎鱼性腺发育及配子发生过程中构成拮抗调节系统。在鼠 (*Mus musculus*) 中, *Foxl2* 表达缺失会导致 *Dmrt1* 基因异位表达^[38], 在尼罗罗非鱼中, *foxl2* 和 *dmrt1* 通过对 *cyp19a1a* 表达的调节在性别分化中相互拮抗^[39]。

Zglp1 (GATA-type zinc finger protein 1) 是一种核内锌指蛋白, 可调节哺乳动物性腺发育过程中体细胞与生殖细胞的相互作用^[40-41], 而在鱼类研究中其可能主要参与了卵子发生与成熟^[42]。本研究中 *zglp1* 转录本在湖栖鲮虾虎鱼成熟卵巢中表达显著高于精巢, 可能说明了 *zglp1* 在鱼类卵子发生中的功能较为保守。Zp (zona pellucida) 参与脊椎动物卵母细胞发生和受精^[43]。在鱼类中, *zp* 的 mRNA 表达水平随着卵子的发生逐渐升高, 特别是在卵黄形成时期, 其表达水平显著高于未发育阶段^[44]。有关斑马鱼的研究表明, Zp2 在卵母细胞膜的早期形成中起关键作用, 而 Zp3 则主要参与了雌性斑马鱼的生殖过程^[45]。本研究中, *zp2*、*zp3* 和 *zp4* 转录产物于卵巢高表达 (表 4), 说明 *zps* 参与了湖栖鲮虾虎鱼卵母细胞成熟及繁殖过程^[13, 46]。

类固醇激素可以调节脊椎动物的性分化和性行为, 对鱼类生殖细胞的发生和维持至关重要^[47]。芳香化酶家族中 Cyp11b 和 Cyp21 在催化皮质醇的产生中起关键作用, 并参与性类固醇的产生^[48]。在欧洲鲈 (*Dicentrarchus labrax*) 和虹鳟中, *cyp11b* 的 mRNA 在精巢中的表达水平显著高于卵巢, 主要参与了精子的发生与成熟^[49-50]。而 *cyp19a1a* 编码产物是催化孕酮转化为雌激素的关键酶, 主要在卵巢中表达^[51]。本研究中, *cyp7a1*、*cyp7b* 和 *cyp27a1* 转录产物在性成熟湖栖鲮虾虎鱼精巢中高表达, 而 *cyp19a1a*、*cyp20a* 和 *cyp27c1* 在卵巢中高表达 (表 4), 表明这些芳香化酶基因可能参与了湖栖鲮虾虎鱼配子发生并调节了类固醇激素的合成。

Inhibins 作为 TGF- β 家族的成员, 对脊椎动物配子发生及维持具有重要作用^[52]。Wu 等^[53]发现, 该基因 *inha* 以剂量依赖的方式调节斑马鱼卵母细胞的成熟。Lankford 等^[54-55]在虹鳟中发现调节卵泡成熟的 MIH 也可以调节 *inha* 的表达。本研究中, *inha* 转录产物在卵巢中高表达, 而 *inhb* 的转录产物则在精巢中高表达, 表明二者可能参与了湖栖鲮虾虎鱼的配子发生。另外, Tdrp (testis development-related protein) 是一种调节精子

表 4 湖栖鳍虾虎鱼部分性别相关基因表达模式

Tab. 4 The expression pattern of differentially expressed sex-related genes of *G. lacustris*

基因ID gene ID	log ₂ (Fold Change)	q校验收 q value	登录号 accession ID	基因名 gene name
cluster-9641.22807	9.666 2	9.10×10^{-42}	XP_020784565.1	<i>sox19</i>
cluster-9641.25293	9.606 1	1.89×10^{-40}	XP_024146001.1	<i>zglp1</i>
cluster-9641.25160	9.604 1	0	XP_020795598.1	<i>zp2</i>
cluster-9641.25068	9.443 0	0	XP_020792434.1	<i>zp3</i>
cluster-9641.24727	8.683 7	1.86×10^{-208}	XP_020781516.1	<i>zp4</i>
cluster-9641.16922	8.079 7	1.89×10^{-34}	XP_020784412.1	<i>gdf9</i>
cluster-9641.24530	7.844 2	2.22×10^{-20}	XP_023123606.1	<i>sycp2</i>
cluster-9641.25344	7.524 3	3.07×10^{-41}	XP_023250053.1	<i>sox11b</i>
cluster-9641.27486	7.150 7	6.73×10^{-19}	XP_020782655.1	<i>gdf3</i>
cluster-9641.29598	5.630 0	1.00×10^{-16}	XP_018535887.1	<i>nanos</i>
cluster-9641.16580	4.730 9	1.43×10^{-4}	XP_020773153.1	<i>foxl2</i>
cluster-9641.26847	4.690 1	6.37×10^{-66}	XP_020780621.1	<i>hsd17b12a</i>
cluster-9641.21833	4.606 5	1.06×10^{-25}	XP_020783195.1	<i>hsd17b2</i>
cluster-9641.20480	4.413 6	4.12×10^{-6}	XP_020797477.1	<i>cyp20a</i>
cluster-9641.15558	4.154 4	1.82×10^{-3}	XP_020775367.1	<i>sox11</i>
cluster-9641.35246	4.102 7	8.82×10^{-8}	XP_020774028.1	<i>cyp26a</i>
cluster-2179.0	3.749 3	1.175×10^{-2}	AAS58448.1	<i>cyp19a</i>
cluster-9641.19795	3.473 3	4.01×10^{-5}	XP_020797477.1	<i>cyp20a</i>
cluster-9641.18341	3.418 1	2.25×10^{-41}	XP_020785212.1	<i>gtsf1</i>
cluster-9641.23440	2.896 2	3.17×10^{-5}	XP_020784490.1	<i>cyp11a</i>
cluster-9641.32089	2.704 9	4.26×10^{-3}	XP_020790615.1	<i>hdac10</i>
cluster-9641.11433	2.691 3	6.95×10^{-4}	XP_020792251.1	<i>inha</i>
cluster-9641.24752	2.535 9	1.67×10^{-6}	XP_008292482.1	<i>hdac8</i>
cluster-9641.13414	2.362 9	3.77×10^{-3}	XP_020776784.1	<i>gdf10</i>
cluster-9641.35361	1.755 9	1.11×10^{-3}	XP_020794666.1	<i>cyp27c1</i>
cluster-9641.28295	1.740 4	7.93×10^{-11}	XP_020787387.1	<i>ddx42</i>
cluster-9641.29360	1.304 8	4.92×10^{-4}	XP_020782050.1	<i>stard3</i>
cluster-9641.16586	1.002 5	1.16×10^{-3}	XP_017329922.1	<i>hdac3</i>
cluster-9641.35558	-1.363 3	2.01×10^{-3}	XM_008397775.2	<i>tdrp</i>
cluster-9641.5267	-3.373 6	4.73×10^{-8}	XP_020788816.1	<i>hsd3b7</i>
cluster-9641.13207	-3.446 4	2.29×10^{-14}	XP_020796962.1	<i>esr1</i>
cluster-9641.1499	-3.594 6	2.78×10^{-3}	XP_023260192.1	<i>cyp2aa_d</i>
cluster-9641.43077	-4.398 0	3.17×10^{-3}	XP_020772714.1	<i>hdac4_5</i>
cluster-9641.47344	-4.631 5	3.54×10^{-4}	XP_023249319.1	<i>hsd17b3</i>
cluster-9641.13065	-4.639 8	3.33×10^{-3}	XP_020786901.1	<i>cyp27a1</i>
cluster-9641.6064	-5.155 4	2.71×10^{-05}	XP_020784595.1	<i>inhb</i>
cluster-9641.29535	-5.180 1	3.04×10^{-13}	XP_008302659.1	<i>cyp7a1</i>
cluster-9641.47528	-7.121 6	2.25×10^{-10}	AAP84972.1	<i>dmrt1</i>
cluster-9641.2530	-7.383 8	1.11×10^{-17}	XP_022066935.1	<i>cyp7b</i>
cluster-9641.38934	-7.477 5	5.37×10^{-295}	XP_022600688.1	<i>krt2</i>
cluster-9641.47188	-8.814 9	2.77×10^{-14}	XP_020774232.1	<i>e3</i>

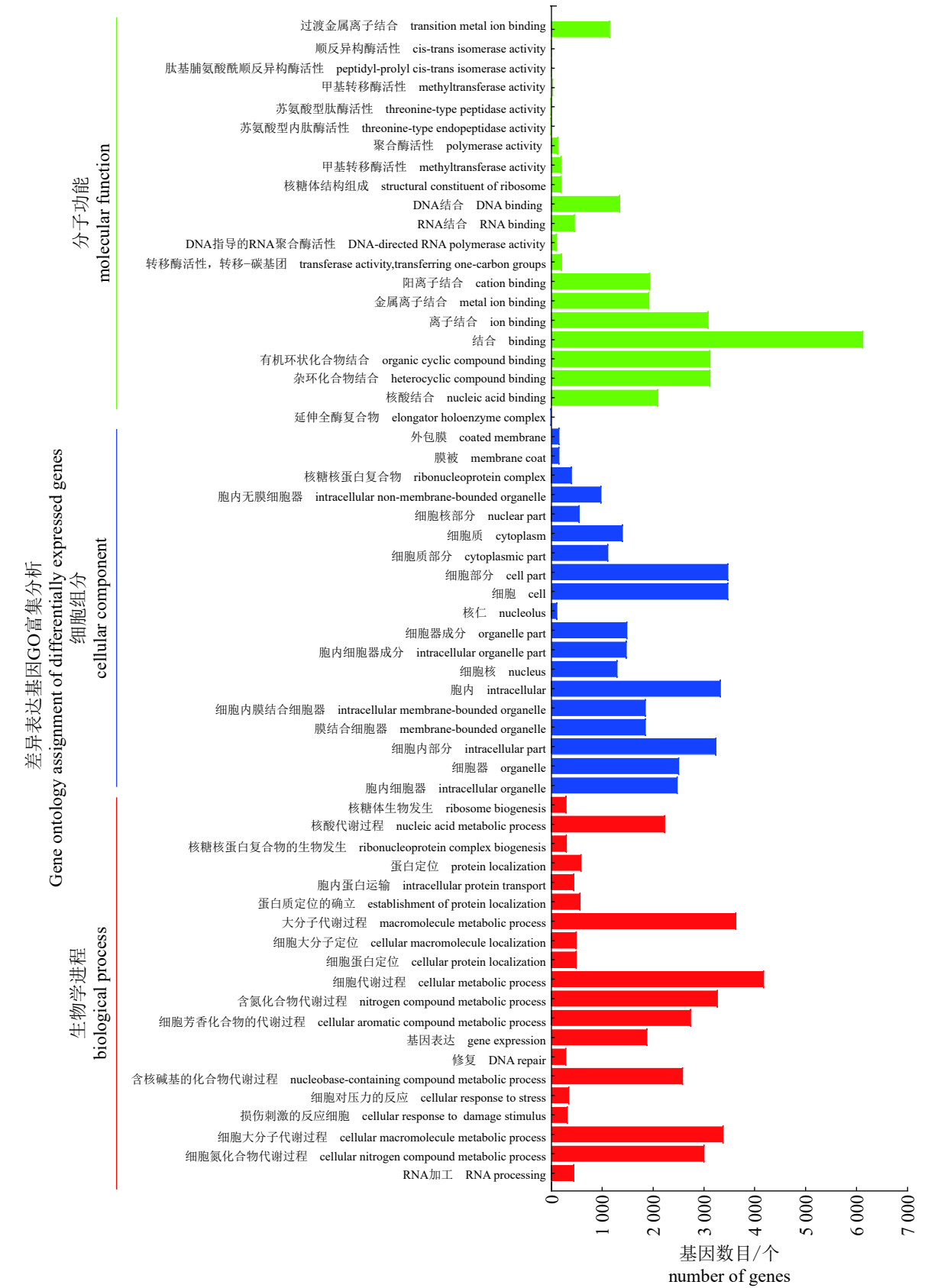


图 4 湖栖鳍虾虎鱼性腺差异表达基因 GO 富集分析

Fig. 4 Gene ontology assignment of differentially expressed genes of *G. lacustris*

表 5 卵巢 vs. 精巢差异基因 KEGG 极显著富集通路

Tab. 5 KEGG highly significant enrichment pathway of differential expression (FG vs MG) gene

通路名称 term name	通路ID term ID	基因数目/个 gene number	P值 P-value
上调通路 up-regulated pathway			
剪接体 spliceosome	ko03040	134	3.23×10^{-19}
RNA运输 RNA transport	ko03013	138	2.01×10^{-15}
真核生物核糖体生成 ribosome biogenesis in eukaryotes	ko03008	82	3.06×10^{-14}
细胞循环 cell cycle	ko04110	117	3.52×10^{-14}
嘧啶代谢 pyrimidine metabolism	ko00240	75	5.34×10^{-12}
DNA复制 DNA replication	ko03030	41	9.49×10^{-11}
泛素介导蛋白水解 ubiquitin mediated proteolysis	ko04120	104	4.73×10^{-10}
核酸切除修复 nucleotide excision repair	ko03420	44	5.27×10^{-10}
mRNA监测通路 mRNA surveillance pathway	ko03015	75	6.34×10^{-9}
RNA聚合酶 RNA polymerase	ko03020	26	1.65×10^{-7}
范可尼贫血通路 Fanconi anemia pathway	ko03460	49	1.87×10^{-7}
RNA降解 RNA degradation	ko03018	60	3.26×10^{-7}
蛋白酶体 proteasome	ko03050	34	1.00×10^{-6}
碱基切除修复 base excision repair	ko03410	31	2.27×10^{-6}
爱波斯坦巴尔病毒感染 Epstein-Barr virus infection	ko05169	123	2.77×10^{-6}
同源重组 homologous recombination	ko03440	34	3.39×10^{-6}
基底转录因子 basal transcription factors	ko03022	35	1.66×10^{-5}
嘌呤代谢 purine metabolism	ko00230	99	1.74×10^{-5}
错配修复 mismatch repair	ko03430	24	2.19×10^{-5}
卵母细胞减数分裂 oocyte meiosis	ko04114	75	1.21×10^{-4}
非同源末端连接 non-homologous end-joining	ko03450	15	1.33×10^{-3}
叶酸构成的碳池 one carbon pool by folate	ko00670	16	1.85×10^{-3}
氨酰生物合成 aminoacyl-tRNA biosynthesis	ko00970	28	1.91×10^{-3}
孕酮介导卵母细胞成熟 progesterone-mediated oocyte maturation	ko04914	56	2.93×10^{-3}
Notch信号通路 Notch signaling pathway	ko04330	37	2.93×10^{-3}
缬氨酸、亮氨酸和异亮氨酸的降解 valine, leucine and isoleucine degradation	ko00280	30	3.85×10^{-3}
硫中继系统 sulfur relay system	ko04122	6	5.93×10^{-3}
TGF- β 信号通路 TGF- β signaling pathway	ko04350	50	6.60×10^{-3}
胞质DNA-传感通路 cytosolic DNA-sensing pathway	ko04623	20	6.81×10^{-3}
p53信号通路 p53 signaling pathway	ko04115	40	7.23×10^{-3}
脂肪酸延伸 fatty acid elongation	ko00062	17	8.59×10^{-3}
糖基磷脂酰肌醇锚区生物合成 biosynthesis of GPI anchor	ko00563	14	8.75×10^{-3}
调控干细胞多能性的信号通路 signaling pathways regulating pluripotency of stem cells	ko04550	72	9.37×10^{-3}
下调通路 down-regulated pathway			
蛋白质消化吸收 protein digestion and absorption	ko04974	57	1.89×10^{-11}
细胞外基质受体相互作用 ECM-receptor interaction	ko04512	35	2.63×10^{-8}
黏着斑 focal adhesion	ko04510	72	3.34×10^{-7}
核糖体 ribosome	ko03010	65	2.28×10^{-6}
ABC转运蛋白 ABC transporters	ko02010	29	1.85×10^{-5}

· 续表 5 ·

通路名称 term name	通路ID term ID	基因数目/个 gene number	P值 P-value
阿米巴疾病 amoebiasis	ko05146	32	1.98×10^{-5}
近端小管重碳酸盐回收 proximal tubule bicarbonate reclamation	ko04964	18	2.28×10^{-5}
昆虫激素的生物合成 insect hormone biosynthesis	ko00981	8	3.16×10^{-5}
胆汁分泌 bile secretion	ko04976	32	3.33×10^{-4}
精氨酸和脯氨酸代谢 arginine and proline metabolism	ko00330	17	8.17×10^{-4}
糖尿病并发症的年龄-愤怒信号通路 AGE-RAGE signaling pathway in diabetic complications	ko04933	30	2.92×10^{-3}
PI3K-Akt信号通路 PI3K-Akt signaling pathway	ko04151	72	3.22×10^{-3}
肌动蛋白细胞骨架调节 regulation of actin cytoskeleton	ko04810	49	3.44×10^{-3}
视黄醇的新陈代谢 retinol metabolism	ko00830	15	3.92×10^{-3}
药物代谢-细胞色素P450 drug metabolism-cytochrome P450	ko00982	15	4.27×10^{-3}

表 6 qPCR 检测 RNA-Seq 结果
Tab. 6 Validation of the RNA-seq data by qPCR

基因ID gene ID	基因名称 genes name	RNA-Seq		qPCR	
		Log ₂ (FG/MG)	q value (FG/MG)	Log ₂ (FG/MG)	P-value (FG/MG)
cluster-9641.3059	<i>cas3</i>	-6.69	4.88×10^{-8}	-1.99	6.20×10^{-2}
cluster-2179.0	<i>cyp19a1a</i>	3.75	1.20×10^{-2}	1.75	1.60×10^{-2}
cluster-9641.47528	<i>dmrt1</i>	-7.12	2.25×10^{-10}	-7.75	3.20×10^{-2}
cluster-9641.47188	<i>e3</i>	-8.81	2.77×10^{-14}	-10.05	4.30×10^{-8}
cluster-9641.16580	<i>foxl2</i>	4.73	1.40×10^{-4}	5.15	1.80×10^{-4}
cluster-9641.43183	<i>gapdh</i>	-5.86	1.03×10^{-72}	-4.50	2.20×10^{-2}
cluster-9641.11433	<i>inha</i>	2.69	7.00×10^{-4}	1.50	2.30×10^{-2}
cluster-9641.6064	<i>inhb</i>	-5.16	2.70×10^{-5}	-4.74	5.20×10^{-4}
cluster-9641.38934	<i>krt2</i>	-7.48	5.37×10^{-295}	-11.90	5.80×10^{-3}
cluster-9641.24002	<i>rpl7</i>	-1.30	5.59×10^{-24}	-0.72	0.20
cluster-9641.24530	<i>syncp2</i>	7.84	2.22×10^{-20}	6.90	2.30×10^{-3}
cluster-9641.35558	<i>tdrp</i>	-1.36	2.01×10^{-3}	-2.22	6.70×10^{-2}
cluster-9641.3821	<i>test</i>	-5.39	6.19×10^{-08}	-8.09	2.40×10^{-2}
cluster-9641.25293	<i>zglpl</i>	9.61	1.89×10^{-40}	12.96	3.50×10^{-5}

发生的核转录因子，在本研究中其转录本在湖栖鳍虾虎鱼精巢中高表达^[56]。以上这些基因可能构成了一个网络共同调节湖栖鳍虾虎鱼配子发生和成熟。

3.2 湖栖鳍虾虎鱼性腺成熟及繁殖相关通路

卵巢的发育和成熟涉及多个生物学途径，在本研究中有 117 个差异基因富集到“细胞循环”，104 个差异基因参与“泛素介导的蛋白水解”通路，75 个差异基因参与“卵母细胞减数分裂”，56 个差异基因参与“孕酮介导的卵母细胞成熟”过程，这些基因共同调节卵巢的发育及成熟。相似的结果在三疣梭子蟹 (*Portunus trituberculatus*) 中也

有发现^[57]。p53 信号通路维持了生物过程的完整性，并对繁殖起调节作用^[12]。在斑马鱼中，p53 调节了生殖细胞的数量并间接调节了斑马鱼的性别决定^[58]。本研究中，p53 信号通路在卵巢中显著富集，表明它可能调节了湖栖鳍虾虎鱼的繁殖。

类固醇激素主要由性腺合成和分泌，并通过血液循环转运至靶组织，与核受体(如雌激素受体和雄激素受体)结合，在鱼类生殖系统的发育中发挥重要作用^[59]。p450 家族可以调节类固醇激素的生物合成，本研究中，“药物代谢—细胞色素 P450 通路”、“P450 介导异源物质代谢”和“类固醇激素生物合成”在精巢中显著富集(表 5)，

表明类固醇代谢酶可能参与湖栖鲮虾虎鱼精巢的发育并调节类固醇激素的合成。

“PI3K/AKT 信号通路”参与精巢的生长和发育, 调节雄激素的分泌和精子的产生^[60]。PI3K 是一类重要的脂质激酶, 可特异性催化磷脂酰肌醇脂质, 是 AKT 的第一个调节因子^[61]。AKT 是一种丝氨酸/苏氨酸蛋白激酶, 在精巢中, AKT 主要在生精细胞和支持细胞中表达。激活后的 PI3K 将 AKT 转运到细胞膜上以激活 AKT, 随后 AKT 促进细胞存活、增殖并调节细胞凋亡^[62]。本研究中, PI3K-Akt 信号通路在精巢中富集, 表明该通路可能调节湖栖鲮虾虎鱼雄性生殖系统发育并维持精巢的稳态。

4 结论

本研究首次构建了湖栖鲮虾虎鱼卵巢和精巢的 cDNA 文库, 并进行了转录组测序。通过 de novo 拼接组装获得了 62 573 个 unigene, 其中 47 891 个 unigene 获得了成功注释, 丰富了湖栖鲮虾虎鱼的基因资源。通过比较卵巢和精巢转录组, 鉴定到一些参与鲮虾虎鱼配子生成及繁殖相关的基因 (*foxl2*、*dmrt1*、*cyp19a1a*、*inha*、*inhib*、*sycp2*、*zglp1*、*zps*、*tdrp* 和 *esra* 等) 和通路 (“泛素介导蛋白水解”、“卵母细胞减数分裂”、“孕酮介导卵母细胞成熟”、“p53 信号通路”、“类固醇激素合成通路”及“PI3/AKT 信号通路”等)。此外, 还从转录组中筛选到 38 550 个 SSR 及 192 450 个 SNP。实验结果为进一步开展湖栖鲮虾虎鱼性别相关基因功能研究和性别相关的分子标记开发奠定基础。

参考文献 (References):

- [1] Matsuda M, Nagahama Y, Shinomiya A, *et al.* DMY is a Y-specific DM-domain gene required for male development in the medaka fish[J]. *Nature*, 2002, 417(6888): 559-563.
- [2] Yano A, Guyomard R, Nicol B, *et al.* An immune-related gene evolved into the master sex-determining gene in rainbow trout, *Oncorhynchus mykiss*[J]. *Current Biology*, 2012, 22(15): 1423-1428.
- [3] Kamiya T, Kai W, Tasumi S, *et al.* A trans-species mis-sense SNP in Amhr2 is associated with sex determination in the tiger pufferfish, *Takifugu rubripes* (fugu)[J]. *PLoS Genetics*, 2012, 8(7): e1002798.
- [4] Hattori R S, Murai Y, Oura M, *et al.* A Y-linked anti-Müllerian hormone duplication takes over a critical role in sex determination[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2012, 109(8): 2955-2959.
- [5] Myosho T, Otake H, Masuyama H, *et al.* Tracing the emergence of a novel sex-determining gene in medaka, *Oryzias luzonensis*[J]. *Genetics*, 2012, 191(1): 163-170.
- [6] Cui Z K, Liu Y, Wang W W, *et al.* Genome editing reveals dmrt1 as an essential male sex-determining gene in Chinese tongue sole (*Cynoglossus semilaevis*)[J]. *Scientific Reports*, 2017, 7: 42213.
- [7] Chen S L, Zhang G J, Shao C W, *et al.* Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic life-style[J]. *Nature Genetics*, 2014, 46(3): 253-260.
- [8] Fan Z F, You F, Wang L J, *et al.* Gonadal transcriptome analysis of male and female olive flounder (*Paralichthys olivaceus*)[J]. *BioMed Research International*, 2014, 2014: 291067.
- [9] Kikuchi K, Hamaguchi S. Novel sex-determining genes in fish and sex chromosome evolution[J]. *Developmental Dynamics*, 2013, 242(4): 339-353.
- [10] Song F B, Wang L M, Zhu W B, *et al.* A novel *igf3* gene in common carp (*Cyprinus carpio*): evidence for its role in regulating gonadal development[J]. *PLoS One*, 2016, 11(12): e0168874.
- [11] Devlin R H, Nagahama Y. Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences[J]. *Aquaculture*, 2002, 208(3-4): 191-364.
- [12] Du X X, Wang B, Liu X M, *et al.* Comparative transcriptome analysis of ovary and testis reveals potential sex-related genes and pathways in spotted knifejaw *Oplegnathus punctatus*[J]. *Gene*, 2017, 637: 203-210.
- [13] Tian C X, Li Z Y, Dong Z D, *et al.* Transcriptome analysis of male and female mature gonads of silver sillago (*Sillago sihama*)[J]. *Genes*, 2019, 10(2): 129.
- [14] 龙娟, 郑树清, 王晓双, 等. TGF- β 信号通路在鱼类性别决定与分化中的作用[J]. *水产学报*, 2020, 44(1): 166-177.
- [15] Long J, Zheng S Q, Wang X S, *et al.* Role of TGF- β signaling pathway in sex determination and differentiation in fish[J]. *Journal of Fisheries of China*, 2020, 44(1): 166-177(in Chinese).
- [16] Sun F Y, Liu S K, Gao X Y, *et al.* Male-biased genes in catfish as revealed by RNA-Seq analysis of the testis transcriptome[J]. *PLoS One*, 2013, 8(7): e68452.
- [17] He F X, Jiang D N, Huang Y Q, *et al.* Comparative transcriptome analysis of male and female gonads reveals sex-biased genes in spotted scat (*Scatophagus argus*)[J].

- Fish Physiology and Biochemistry*, 2019, 45(6): 1963-1980.
- [17] Wang Z D, Liao J, Huang C Q, *et al.* Significant genetic differentiation of *Gobiopertus lacustris*, a newly recorded transparent goby in China[J]. *Mitochondrial DNA Part A: DNA Mapping, Sequencing, and Analysis*, 2018, 29(5): 785-791.
- [18] 廖健, 张顺, 龙水生, 等. 5种虾虎鱼类线粒体COI基因序列变异及系统进化[J]. *广东海洋大学学报*, 2016, 36(1): 7-12.
- Liao J, Zhang S, Long S S, *et al.* Sequence variation and molecular phylogeny of mitochondrial COI gene segments from five species of Gobiidae family[J]. *Journal of Guangdong Ocean University*, 2016, 36(1): 7-12(in Chinese).
- [19] 黄承勤, 廖健, 张顺, 等. 中国鳍虾虎鱼属(鲈形目: 虾虎鱼科)一新纪录种[J]. *广东海洋大学学报*, 2018, 38(2): 1-6.
- Huang C Q, Liao J, Zhang S, *et al.* A new record of gobiopertus (Perciformes: Gobiidae) in China[J]. *Journal of Guangdong Ocean University*, 2018, 38(2): 1-6(in Chinese).
- [20] Nan F R, Feng J, Lv J P, *et al.* Transcriptome analysis of the typical freshwater rhodophytes *Sheathia arcuata* grown under different light intensities[J]. *PLoS One*, 2018, 13(5): e0197729.
- [21] Grabherr M G, Haas B J, Yassour M, *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome[J]. *Nature Biotechnology*, 2011, 29(7): 644-652.
- [22] Moriya Y, Itoh M, Okuda S, *et al.* KAAS: an automatic genome annotation and pathway reconstruction server[J]. *Nucleic Acids Research*, 2007, 35(S2): W182-W185.
- [23] Eddy S R. Accelerated profile HMM searches[J]. *PLoS Computational Biology*, 2011, 7(10): e1002195.
- [24] Finn R D, Mistry J, Tate J, *et al.* The Pfam protein families database[J]. *Nucleic Acids Research*, 2010, 38(S1): D211-D222.
- [25] Götz S, García-Gómez J M, Terol J, *et al.* High-throughput functional annotation and data mining with the Blast2GO suite[J]. *Nucleic Acids Research*, 2008, 36(10): 3420-3435.
- [26] Li B, Dewey C N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome[J]. *BMC Bioinformatics*, 2011, 12: 323.
- [27] Wang L K, Feng Z X, Wang X, *et al.* DEGseq: an R package for identifying differentially expressed genes from RNA-seq data[J]. *Bioinformatics*, 2010, 26(1): 136-138.
- [28] Livak K J, Schmittgen T D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method[J]. *Methods*, 2001, 25(4): 402-408.
- [29] Hu Q M, Tian H F, Li W, *et al.* Identification of critical sex-biased genes in *Andrias davidianus* by de novo transcriptome[J]. *Molecular Genetics and Genomics*, 2019, 294(2): 287-299.
- [30] Li Y P, Zhang L L, Sun Y, *et al.* Transcriptome sequencing and comparative analysis of ovary and testis identifies potential key sex-related genes and pathways in scallop *Patinopecten yessoensis*[J]. *Marine Biotechnology*, 2016, 18(4): 453-465.
- [31] You X X, Bian C, Zan Q J, *et al.* Mudskipper genomes provide insights into the terrestrial adaptation of amphibious fishes[J]. *Nature Communications*, 2014, 5: 5594.
- [32] Webster K A, Schach U, Ordaz A, *et al.* Dmrt1 is necessary for male sexual development in Zebrafish[J]. *Developmental Biology*, 2017, 422(1): 33-46.
- [33] 江东能, 彭友幸, 黄远青, 等. 基于天然XY雌鱼培育YY超雄尼罗罗非鱼的新方法[J]. *水产学报*, 2020, 44(11): 1862-1872.
- Jiang D N, Peng Y X, Huang Y Q, *et al.* Comparative transcriptome analysis of male and female gonads reveals sex-biased genes in spotted scat (*Scatophagus argus*)[J]. *Journal of Fisheries of China*, 2020, 44(11): 1862-1872(in Chinese).
- [34] Masuyama H, Yamada M, Kamei Y, *et al.* Dmrt1 mutation causes a male-to-female sex reversal after the sex determination by Dmy in the medaka[J]. *Chromosome Research*, 2012, 20(1): 163-176.
- [35] Lin Q H, Mei J, Li Z, *et al.* Distinct and cooperative roles of *amh* and *dmrt1* in self-renewal and differentiation of male germ cells in Zebrafish[J]. *Genetics*, 2017, 207(3): 1007-1022.
- [36] Zhang X B, Li M R, Ma H, *et al.* Mutation of *foxl2* or *cyp19a1a* results in female to male sex reversal in *X Nile tilapia*[J]. *Endocrinology*, 2017, 158(8): 2634-2647.
- [37] Yang Y J, Wang Y, Li Z, *et al.* Sequential, divergent, and cooperative requirements of *foxl2a* and *foxl2b* in ovary development and maintenance of Zebrafish[J]. *Genetics*, 2017, 205(4): 1551-1572.
- [38] Lindeman R E, Gearhart M D, Minkina A, *et al.* Sexual cell-fate reprogramming in the ovary by DMRT1[J]. *Current Biology*, 2015, 25(6): 764-771.
- [39] Li M H, Yang H H, Li M R, *et al.* Antagonistic roles of Dmrt1 and Foxl2 in sex differentiation via estrogen production in tilapia as demonstrated by TALENs[J]. *Endocrinology*, 2013, 154(12): 4814-4825.
- [40] Li S R, Lu M M, Zhou D Y, *et al.* GLP-1: a novel zinc finger protein required in somatic cells of the gonad for germ cell development[J]. *Developmental Biology*, 2013, 371(1): 1-11.

- 2007, 301(1): 106-116.
- [41] Strauss T J, Castrillon D H, Hammes S R. GATA-like protein-1 (GLP-1) is required for normal germ cell development during embryonic oogenesis[J]. *Reproduction*, 2011, 141(2): 173-181.
- [42] Dong Z D, Zhang N, Liu Y, *et al.* Expression analysis and characterization of *zglp1* in the Chinese tongue sole (*Cynoglossus semilaevis*)[J]. *Gene*, 2019, 683: 72-79.
- [43] Howes L, Jones R. Interactions between zona pellucida glycoproteins and sperm proacrosin/acrosin during fertilization[J]. *Journal of Reproductive Immunology*, 2002, 53(1-2): 181-192.
- [44] Zeng S, Gong Z Y. Expressed sequence tag analysis of expression profiles of Zebrafish testis and ovary[J]. *Gene*, 2002, 294(1-2): 45-53.
- [45] Liu X J, Wang H, Gong Z Y. Tandem-repeated Zebrafish *zp3* genes possess oocyte-specific promoters and are insensitive to estrogen induction[J]. *Biology of Reproduction*, 2006, 74(6): 1016-1025.
- [46] Wu T L, Cheng Y Y, Liu Z L, *et al.* Bioinformatic analyses of zona pellucida genes in vertebrates and their expression in Nile tilapia[J]. *Fish Physiology and Biochemistry*, 2018, 44(2): 435-449.
- [47] Svechnikov K, Söder O. Ontogeny of gonadal sex steroids[J]. *Best Practice & Research Clinical Endocrinology & Metabolism*, 2008, 22(1): 95-106.
- [48] Applebaum S L, Wilson C A, Holt G J, *et al.* The onset of cortisol synthesis and the stress response is independent of changes in CYP11B or CYP21 mRNA levels in larval red drum (*Sciaenops ocellatus*)[J]. *General and Comparative Endocrinology*, 2010, 165(2): 269-276.
- [49] Socorro S, Martins R S, Deloffre L, *et al.* A cDNA for European sea bass (*Dicentrarchus labrax*) 11 β -hydroxylase: gene expression during the thermosensitive period and gonadogenesis[J]. *General and Comparative Endocrinology*, 2007, 150(1): 164-173.
- [50] Liu S J, Govoroun M, D'Cotta H, *et al.* Expression of cytochrome P450_{11 β} (11 β -hydroxylase) gene during gonadal sex differentiation and spermatogenesis in rainbow trout, *Oncorhynchus mykiss*[J]. *The Journal of Steroid Biochemistry and Molecular Biology*, 2000, 75(4-5): 291-298.
- [51] Li Q Q, Du X, Pan Z X, *et al.* The transcription factor SMAD4 and miR-10b contribute to E2 release and cell apoptosis in ovarian granulosa cells by targeting CYP19A1[J]. *Molecular and Cellular Endocrinology*, 2018, 476: 84-95.
- [52] Knight P G, Satchell L, Glistler C. Intra-ovarian roles of activins and inhibins[J]. *Molecular and Cellular Endocrinology*, 2012, 359(1-2): 53-65.
- [53] Wu T T, Patel H, Mukai S, *et al.* Activin, inhibin, and follistatin in zebrafish ovary: expression and role in oocyte maturation[J]. *Biology of Reproduction*, 2000, 62(6): 1585-1592.
- [54] Lankford S E, Weber G M. Temporal mRNA expression of transforming growth factor-beta superfamily members and inhibitors in the developing rainbow trout ovary[J]. *General and Comparative Endocrinology*, 2010, 166(2): 250-258.
- [55] Lankford S E, Weber G M. The maturation-inducing hormone 17 α , 20 β -dihydroxy-4-pregnen-3-one regulates gene expression of inhibin β_A and *bambi* (bone morphogenetic protein and activin-membrane-bound inhibitor) in the rainbow trout ovary[J]. *General & Comparative Endocrinology*, 2010, 168(3): 369-376.
- [56] Wang X C, Jiang H W, Zhou W B, *et al.* Molecular cloning of a novel nuclear factor, TDRP1, in spermatogenic cells of testis and its relationship with spermatogenesis[J]. *Biochemical and Biophysical Research Communications*, 2010, 394(1): 29-35.
- [57] Yang Y X, Wang J T, Han T, *et al.* Ovarian transcriptome analysis of *Portunus trituberculatus* provides insights into genes expressed during phase III and IV development[J]. *PLoS One*, 2015, 10(10): e0138862.
- [58] Rodríguez-Marí A, Cañestro C, BreMiller R A, *et al.* Sex reversal in Zebrafish *fancl* mutants is caused by Tp53-mediated germ cell apoptosis[J]. *PLoS Genetics*, 2010, 6(7): e1001034.
- [59] Louw-du Toit R, Storbeck K H, Cartwright M, *et al.* Progestins used in endocrine therapy and the implications for the biosynthesis and metabolism of endogenous steroid hormones[J]. *Molecular and Cellular Endocrinology*, 2017, 441: 31-45.
- [60] Chambard J C, Lefloch R, Pouyssegur J, *et al.* ERK implication in cell cycle regulation[J]. *Biochimica et Biophysica Acta*, 2007, 1773(8): 1299-1310.
- [61] Lee S, Suk K, Kim I K, *et al.* Signaling pathways of bisphenol A-induced apoptosis in hippocampal neuronal cells: Role of calcium-induced reactive oxygen species, mitogen-activated protein kinases, and nuclear factor- κ B[J]. *Journal of Neuroscience Research*, 2008, 86(13): 2932-2942.
- [62] Zhao D, Zhang X Y. Selenium antagonizes the lead-induced apoptosis of chicken splenic lymphocytes *in vitro* by activating the PI3K/Akt pathway[J]. *Biological Trace Element Research*, 2018, 182(1): 119-129.

Comparative transcriptome analysis of the gonad of the lacustrine goby (*Gobiopsis lacustris*)

DONG Zhongdian, LI Xueyou, HUANG Chengqin, ZHANG Hairui, HUANG Shunkai,
ZHANG Ning, GUO Yusong, WANG Zhongduo *

(Key Laboratory of Aquaculture in South China Sea for Aquatic Economic Animal of Guangdong Higher Education Institutes,
Fisheries College, Guangdong Ocean University, Zhanjiang 524088, China)

Abstract: Lacustrine goby (*Gobiopsis lacustris*) belongs to a genus of gobies that are small in size and endemic to freshwater, brackish waters or coastal environments around the Indian and Pacific oceans. *G. lacustris* has almost transparent skin and clearly visible internal organs, making it an ideal fish model for potential research in development, reproduction and other fields. Here, we constructed gonadal transcriptomes of *G. lacustris* using Illumina sequencing for the first time and identified genes that may be involved in gonadal development, gametogenesis and reproduction. Raw reads were assembled into 62 573 unigenes with N50 value of 3 082 bp and a mean length of 1 869 bp. 41 480, 32 848, 37 500, 35 394, 18 318, 35 394 and 27 009 unigenes were successfully annotated in NR, NT, SwissProt, PFAM, KOG, GO and KEGG databases, respectively. Gene expression patterns in the ovary and testis were compared, and 10 954 differentially expressed genes (DEGs) were identified. Among these genes, 8 571 were up-regulated in the ovary, and 2 383 were up-regulated in the testis. qPCR analysis of 14 selected genes showed patterns consistent with the transcriptome results. Numerous DEGs involved in gonadal development and gametogenesis were identified, including *foxl2*, *dmrt1*, *cyp19a1a*, *inha*, *inhib*, *sycp2*, *zglp1*, *tdrp*, *zps* and *esra*. Using GO and KEGG enrichment analyses, pathways involved regulation of gonadal development and gametogenesis, such as “ubiquitin mediated proteolysis”, “oocyte meiosis”, “progesterone-mediated oocyte maturation”, “p53 signaling pathway” and “PI3K-Akt signaling pathway”, were also identified. In addition, 38 550 simple sequence repeats were identified from 20 517 SSR containing sequences, and 192 450 single nucleotide polymorphisms were detected. This study denotes the first gonadal transcriptomic analysis of *G. lacustris* and provides a valuable dataset for future functional analysis of sex-associated and molecular marker-assisted selections in *G. lacustris*.

Key words: *Gobiopsis lacustris*; RNA-Seq; gametogenesis; reproduction

Corresponding author: WANG Zhongduo. E-mail: aduofa@hotmail.com

Funding projects: National Natural Science Foundation of China (41806195; 31201996); Technology Planning Project of Guangdong Province, China (2017A030303075); Guangdong Ocean University Featured Innovation Project (230419069; 230419055); Start-up Fund from GDOU