

基于 RNA-seq 技术的罗氏沼虾(*Macrobrachium rosenbergii*)不同组织转录组比较分析*

李喜莲 顾志敏 慎佩晶 徐洋 张宇飞 高强 程海华 陈雪峰

(农业农村部淡水渔业健康养殖重点实验室 浙江省淡水水产遗传育种重点实验室 浙江省淡水水产研究所 湖州 313001)

摘要 罗氏沼虾(*Macrobrachium rosenbergii*)是我国重要的淡水虾养殖品种,为探讨罗氏沼虾不同组织基因表达差异,本研究使用 Illumina HiSeq 平台分析了罗氏沼虾的 7 个组织(眼柄、肝脏、卵巢、鳃、心脏、肌肉、精巢)转录组,质控后获得高质量的 clean reads 36325476、56796932、36328098、51370140、45010606、43240160、42404294 条,共计 46.2 Gb 的 clean reads 数据。测序结果经 de novo 组装共获得 95220 个 Unigenes,每个 Unigenes 的平均长度为 1064.9bp。经过生物信息学方法与五个数据库(NR, GO, COG, KEGG, SWSS-PRO)比对,共注释到 20368 个 Unigenes。其中 GO 功能注释将 Unigenes 分为生物过程、细胞组分和分子功能 3 个大类 50 个分支;在与 COG 数据库比对中,共有 15798 条比对到了同源序列,且一共被分为 25 类;KEGG 代谢通路包括 6 大类 33 个小类,可将 Unigenes 映射到 330 条代谢通路中。分子标记筛选后获得 37751 个潜在 SSR 位点,共发现 3228575 个 SNP 位点。本研究通过对罗氏沼虾 7 个不同组织进行转录组测序,数据组装及功能注释,为进一步深入挖掘和开发利用罗氏沼虾功能基因提供参考。

关键词 罗氏沼虾; 转录组; RNA-seq 技术; 不同组织; 功能基因; 比较分析

中图分类号 Q523.8; Q789; S917 **doi:** 10.11693/hyhz20200600181

罗氏沼虾(*Macrobrachium rosenbergii*), 又称淡水长臂大虾、马来西亚大虾, 动物学分类上属于节肢动物门(Arthropoda), 软甲纲(Malacostraca), 十足目(Decapoda), 长臂虾科(Palaemonidae), 沼虾属(*Macrobrachium*), 在各种类型的淡水和半咸水中都能生活, 易养殖且体型肥大, 肉质鲜美。因其适应能力强、生长速度快, 罗氏沼虾已成为我国重要的淡水养殖虾类之一。

近年来对罗氏沼虾的研究主要包括遗传多样性研究, 性腺发育、基因克隆及生长等方面的研究。其中在多样性研究中, 以微卫星为研究对象的如吕敏等(2019)研究异型雄性罗氏沼虾遗传多样性, 戴习林等(2017)对种群 SSR 分析中样本量及标记量对遗传多样性指标的影响进行分析; 或对各不同群体进行

遗传多样性研究(董丁健等, 2020; 冯艺, 2018; 孙成飞等, 2015); 周晓敏等(2020)选取 60 个 SNP 位点对养殖群体和选育群体进行多样性研究。罗氏沼虾雌激素相关受体(Estrogen-related receptor, ERR)成为研究卵巢发育的一个重要基因(赵苗鑫等, 2017; 刘金磊, 2018; 刘金磊等, 2018)。药物对性腺分化及发育的影响研究主要包括壬基酚(薛海波, 2010; 李郁娇, 2011; 朱春华等, 2017; Guo *et al.*, 2019)、十氯酮(Lafontaine *et al.*, 2016, 2017)及三丁基锡(薛海波, 2010; 李郁娇, 2011; 吴维福等, 2013)这三种药物, 以期研究对罗氏沼虾卵黄蛋白原基因表达以及性腺发育的影响。Stalin 等(2019)研究了钴-60 射线对淡水对虾生殖障碍的影响; Tan 等(2019)研究了罗氏沼虾性别逆转与雄激素腺(AG)的关系。罗氏沼虾基因研究集中在免疫相

* 国家重点研发计划蓝色粮仓科技创新专项, 2018YFD0901303 号; 浙江省农业(水产)新品种选育重大科技专项, 2016C02055-2 号; 现代农业产业技术体系专项资金资助, CARS-48 号。李喜莲, 博士, 副研究员, lixilian@126.com

通信作者: 陈雪峰, 副研究员, chenfenghailan@126.com

收稿日期: 2020-06-28, 收修改稿日期: 2020-07-27

关(刘伟利等, 2017; 江婷佳, 2017; 邱庆庆等, 2019)、性别相关(俞炎琴, 2013; 姜建萍等, 2019; Abayed *et al.*, 2019)、生长相关(叶成凯等, 2019; 邱庆庆, 2019; 杨光等, 2020; Dong *et al.*, 2020)和酶类(田荣等, 2016; 张夏青等, 2016; 卢志杰等, 2019)。罗氏沼虾生长研究集中在药物对生长的影响(程安达等, 2019; 潘忠超等, 2019; Tadese *et al.*, 2020)、微生态制剂(朱光来等, 2019; 赵臣泽, 2019)、生长环境(戴习林等, 2016; 陈建酬等, 2017; 何竺柳, 2018; 朱其建等, 2019; 张俊功, 2019; Chen *et al.*, 2019; Dong *et al.*, 2020; Manickam *et al.*, 2020)、饲料中不同配比(杨树浩等, 2018; 杨景丰等, 2019; 张剑伟等, 2019; 单凡等, 2019; 黄黎明等, 2019; Feng *et al.*, 2019)和不同品系生长对比试验(蒋飞等, 2013, 2014, 2016)。

随着高通量测序技术的发展, 转录组技术也应用于罗氏沼虾的研究中。郭梁等(2018)利用高通量测序技术和数字基因图谱等技术对感染螺原体的罗氏沼虾进行免疫通路相关基因及其差异表达分析, 获得转录本 43405 个; 严赛峰等(2018)、李俊杰等(2018)和邓泽森等(2018)对感染螺原体的罗氏沼虾高通量测序结果开展了 SSR 位点、SNP 位点和重要免疫通路相关 microRNA 研究和分析; 李威霖(2018)对罗氏沼虾肢体再生组织进行了转录组及其差异表达基因研究, 得到 Unigenes 总数目为 87783; 王传聪等(2018)对罗氏沼虾肝胰腺组织转录组测序结果进行 SSR 检测与分析, 获得 15356 个 SSR 位点; 陈雪峰等(2019)采用 Illumina HiSeq™ 4000 高通量测序研究罗氏沼虾卵巢发育四个时期卵巢组织的差异, 卵巢发育四个时期共获得 95379 个 Unigenes。Jiang 等(2019a)通过高通量测序技术研究了罗氏沼虾雌性(ZW)、雄性(ZZ)和超雌个体(WW)的性腺差异。Pasookhush 等(2019)采用转录组技术分析了罗氏沼虾幼虾感染新型冠状病毒的反应。Jiang 等(2019b)对未成年雌虾和成年雌虾的眼柄组织进行转录组分析, 共获得 53878 个 Unigenes; Cao 等(2017)研究罗氏沼虾正常样本和 WSSV 感染样本的淋巴组织转录组, 分别获得 73658 和 72374 个 Unigenes; Rao 等(2016)分别研究了罗氏沼虾正常样本和 WSSV 感染样本的肝胰腺转录组, 共获得 63584 个 Unigenes。

本研究以罗氏沼虾 7 个组织(眼柄、肝脏、卵巢、鳃、心脏、肌肉、精巢)为研究对象, 采用新一代高通量测序技术进行转录组测序分析, 通过 Trinity 软件组装、数据库功能注释、基因表达差异分析和 SSR、

SNP 位点筛选, 获得罗氏沼虾分子遗传信息, 以期为进一步研究罗氏沼虾遗传多样性、功能基因及基因表达差异提供理论数据, 同时为深入研究罗氏沼虾生长阻滞现象提供一定的数据基础。

1 材料与方法

1.1 实验材料

实验用罗氏沼虾(*Macrobrachium rosenbergii*)样品取自浙江省淡水水产研究所八里店综合试验基地, 体重为(5.34±1.22) g, 体长为(6.13±0.42) cm。样品经解剖后取各组织(眼柄、肝脏、卵巢、肝脏、心脏、肌肉、精巢), 共 7 个组织。每个组织取自 3 个个体(见表 1)。将分装好的样品迅速投入液氮中速冻, 置于 -80°C 冰箱中保存, 干冰运输。

表 1 组织及样品名称列表
Tab.1 List of tissue and sample names

序号	组织名称	样品命名	样品名称
1	眼柄	Mrck-E	Mrck-E1, Mrck-E2, Mrck-E3
2	肝脏	Mrck-L	Mrck-L1, Mrck-L2, Mrck-L3
3	卵巢	Mrck-O	Mrck-O1, Mrck-O2, Mrck-O3
4	鳃	Mrck-G	Mrck-G1, Mrck-G2, Mrck-G3
5	心脏	Mrck-H	Mrck-H1, Mrck-H2, Mrck-H3
6	肌肉	Mrck-M	Mrck-M1, Mrck-M2, Mrck-M3
7	精巢	Mrck-T	Mrck-T1, Mrck-T2, Mrck-T3

1.2 方法

1.2.1 RNA 提取 各组织样品均选用 3 个个体的混合样组织, 使用 TRIzol® 试剂按照制造商的说明从各组织中提取总 RNA (Invitrogen), 使用 DNase I (TaKaRa)去除基因组 DNA。使用 Nanodrop ND-2000 分光光度计(美国赛默飞)、Aglient 2100 分析仪器对总 RNA 的纯度、浓度和完整性进行检测。RIN 值>7 的 RNA 用于下游实验。

1.2.2 mRNA-seq 文库构建和 Illumina 测序 使用 mRNA-seq 样品制备试剂盒(Illumina, San Diego, CA)按照试剂说明书步骤构建了 mRNA-seq 文库和 Illumina 测序文库。

1.2.3 质量控制和从头转录组组装 使用 Fast QC 程序 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>)检查 Illumina HiSeq 2500 平台在校准前产生的原始序列读取的质量。低质量读数低于阈值质量 20; 将长度小于 50 bp 的 reads 以及包含适配序列、poly-N 和来自原始数据的测序引物的 reads 去除,

得到干净的 reads 同时, 对质控数据的误差率%、Q30、GC-含量%和 sequence 重复水平进行了评价。所有的后续分析和注释都依赖于高质量的 clean reads。

使用 Trinity (v 2.8.5)软件(<https://sourceforge.net/projects/trinityrnaseq/>)对质控后的数据进行从头组装, 获得转录本序列和 Unigenes 序列。最后, 只有长度 300 bp 的 unique contigs 才能用于组装后的下游研究。

1.2.4 生物信息学分析 功能注释。 拼接得到的 Unigenes 序列, 使用 BLASTX 比对(BLAST+2.7.1, 比对标准: E 值不大于 $1e^{-5}$)与非冗余蛋白数据库 (Non-Redundant Protein Sequence Database, NR)、SWISS-PROT 蛋白质序列数据库(Manually Annotated and Reviewed Protein Sequence Database, SWISS-PROT)、基因本体数据库(Gene Ontology, GO)、直系同源蛋白数据库(Cluster of Orthologous Groups of Proteins, COG), 京都基因和基因组百科全书 KEGG 数据库比对。使用 Trinity 软件自带的 ORF 预测模块进行开发阅读框(ORF)预测。

1.2.5 差异表达分析 通过 Bowtie2 软件将 reads 映射到组装的 Unigenes 序列, 使用 RSEM 软件根据比对结果来计算特定转录本的表达量水平。衡量基因表达水平的标准为 RPKM 值(Reads per kilobase of

exon model per million mapped reads), 即每一百万条序列中, 每个基因以一千个碱基为单位, 比对上的 reads 个数。

使用 DE-Seq 软件分析各个组织之间的差异表达, 从而找到差异基因组。以差异倍数(Foldchange) > 2 和假发现率(FDR)调整显著性值 0.05 为判断 Unigenes 表达显著性的阈值。利用 GO、egg NOG、KEGG 或 thology (KO)和 KEGG 通路富集分析对 DEGs 进行分类, 并对潜在的生物学途径中的 DEGs 进行评价。GO 和 KEGG 途径富集分析中 P 值小于 0.05 的过程、功能或成分在 DEGs 中被认为存在显著差异。

2 结果

2.1 原始测序数据质控及统计

罗氏沼虾 7 个不同组织测序共得到 344151324 条原始序列, 质控后得到有效 RNA-seq 311475706 条。每个个体的测序量为 6.65—9.04 Gb, 平均 7.38 Gb。去除接头序列, 截去连续 4 个碱基平均质量值低于 20 的部分, 舍去长度少于 50 bp 的 reads。共得到 311475706 条 reads。过滤后每个个体的测序量为 5.3—8.45 Gb, 平均 6.60 Gb(见表 2)。每个个体的测序量为每 4 个碱基的平均质量值均大于 20。

表 2 质控后数据统计表
Tab.2 Statistics of data after quality control

样品名称	质控后序列数量	质控后序列长度	Q20(%)	Q30(%)
Mrck-E	36325476	5337789559	99.26	97.60
Mrck-L	56796932	8451248007	98.99	96.68
Mrck-O	36328098	5342364405	98.39	94.47
Mrck-G	51370140	7655928756	98.36	94.31
Mrck-H	45010606	6710024684	98.36	94.31
Mrck-M	43240160	6439706561	98.36	94.41
Mrck-T	42404294	6319330011	98.34	94.33

2.2 转录本拼接

转录本拼接后共获得 95220 个 Unigenes, 总 Unigenes 长度为 101401098 bp。总的转录本数量为 145717 个, 总的转录本长度为 207379988 bp。每个 Unigenes 的平均长度为 1064.9 bp, 最长的 Unigenes 长为 36137 bp, N50 值为 1553。大部分功能基因长度分布在 401—600 bp, 占比达到 34.27%(见图 1)。

2.3 功能注释

在 NR、GO、COG、KEGG、SWSS-PROT 五个公共数据库中对获得的 95220 个 Unigenes 进行功能

注释, 其余未在 NR 数据库中的找到的 Unigenes 可能为新的蛋白。

根据 GO 数据库, 总共有 18485 个基因被归类到三个主要的 GO 类别中: 生物过程、分子功能和细胞成分。其中, 以“细胞”(14938)、“细胞部分”(14925)、“细胞器”(10827)、“生物调节”(9861)为主(见图 2)。

根据 COG 功能分类分成 26 类, 其中数量前三位的为功能未知 2172 个(13.56%), 信号转导机制 2070 个(12.92%), 翻译后修饰, 蛋白质转换, 伴侣 1660 个(10.36%)(见图 3)。

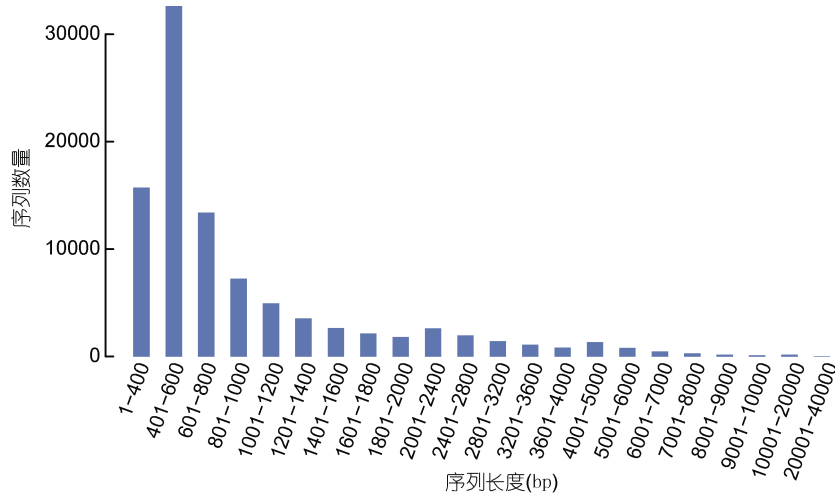


图 1 组装序列长度分布图

Fig.1 Assembly sequence length distribution diagram

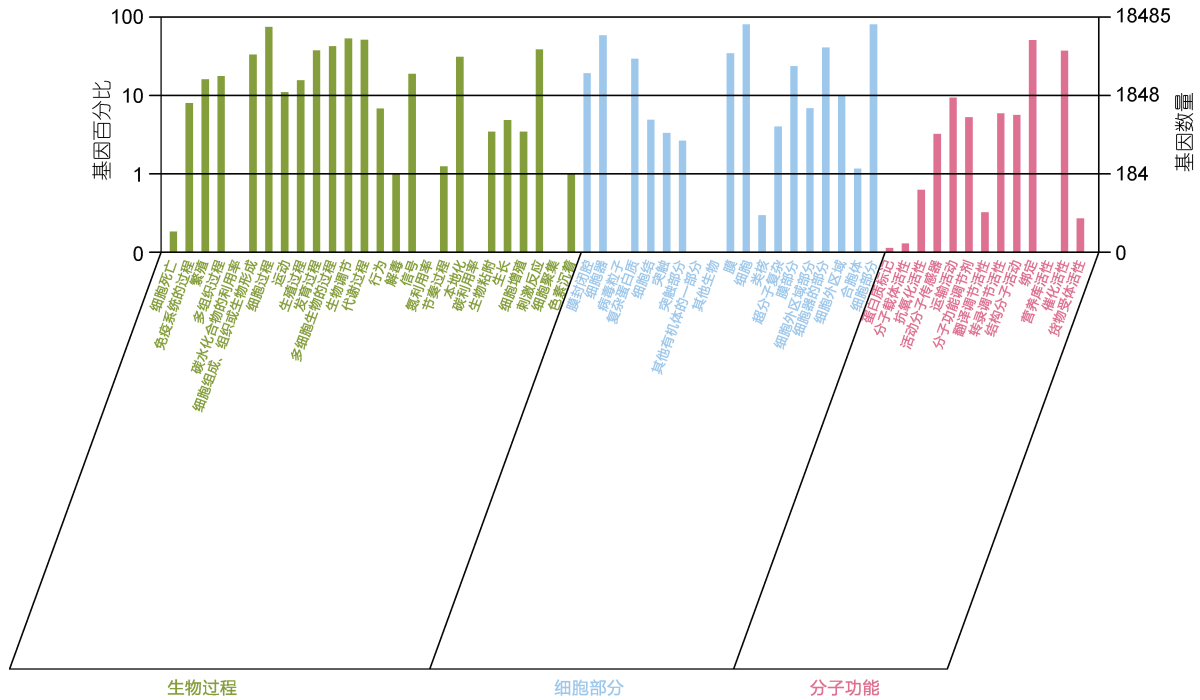


图 2 GO 统计二级统计图

Fig.2 Secondary chart of GO level 2

五个数据库 NR、GO、COG、KEGG、SWSS-PROT 分别注释到 19881、18485、15798、9147、13684 个 Unigenes, 在 NR 数据库中比对注释的信息最多, 注释 19881 个 Unigenes, 占比达 20.88%。在 NR、GO、COG、KEGG、SWSS-PROT 五个数据库中都注释到的 Unigenes 有 7848 个(见图 4)。

将基因根据参与的 KEGG 代谢通路分为 5 个分支(见图 5): 细胞过程(A, Cellular Processes), 环境信息处理(B, Environmental Information Processing), 遗

传信息处理(C, Genetic Information Processing), 代谢(D, Metabolism), 有机系统(E, Organismal Systems)。以上 5 个分支中含量的最多的类型分别为: 全局和概率地图(Global and overview maps), 转化(Translation), 单组织过程(Signal transduction), 运输和分解代谢(Transport and catabolism)、内分泌系统(Endocrine system)。

2.4 差异基因表达分析

7 个不同组织鉴定到的 mRNA 数量由多到少排

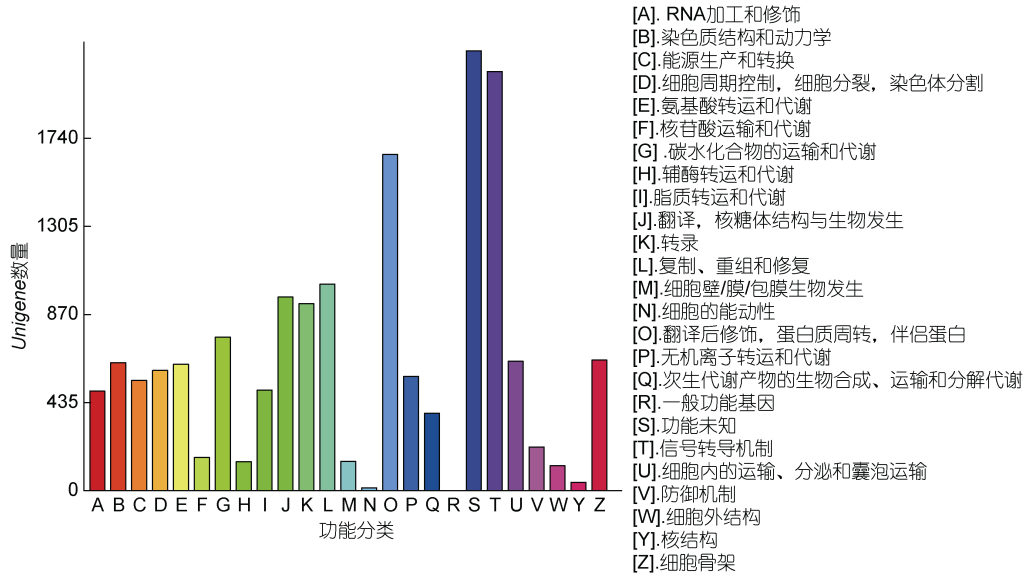


图3 COG 分类统计图
Fig.3 COG classification statistical chart

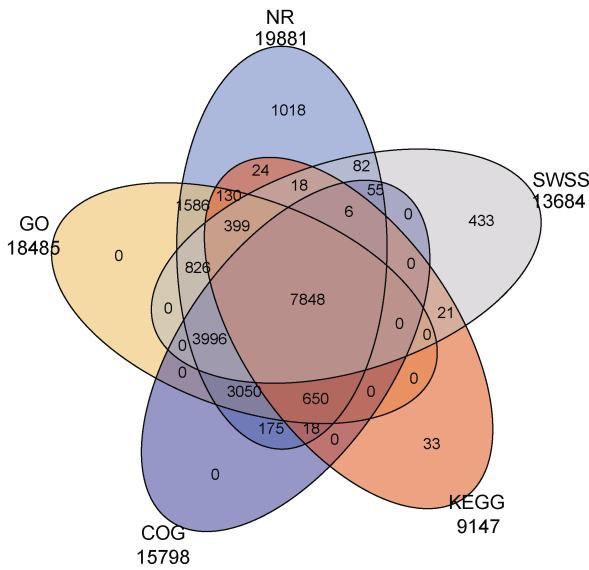


图4 注释信息统计韦恩图

Fig.4 Annotate the information with statistical Venn diagrams

序为: G(鳃) > T(精巢) > H(心脏) > L(肝脏) > E(眼柄) > O(卵巢) > M(肌肉)(见图6)。在7个组织中共同表达的基因数为15260个。

2.5 KEGG 通路分析

对罗氏沼虾转录组的 KEGG 分析显示, 9148 个 Unigenes 被注释到 330 条 KEGG 通路中, 其中代谢途径(1578 个)、核糖体(540 个)、次生代谢物的生物合成(528 个)通路数量居前三(见图7)。

330 条通路中, 其中信号通路包括: PI3K-Akt 信号通路、Rap1 信号通路、催产素信号通路、Hippo

信号通路、MAPK 信号通路、胰岛素信号通路、雌激素信号通路、Ras 信号通路、甲状腺激素信号通路、钙信号通路、HIF-1 信号通路、FoxO 信号通路、生成信号通路、Wnt 信号通路、PPAR 信号通路、促性腺激素信号通路、趋化因子信号通路、Adipocytokine 信号通路、点样受体信号通路、ErbB 信号通路、mTOR 信号通路、Jak-STAT 信号通路、T 细胞受体信号通路、p53 信号通路、肿瘤坏死因子信号通路、VEGF 信号通路、B 细胞受体信号通路、Notch 信号通路、催乳激素信号通路、NF-kappa B 信号通路、Fc epsilon RI 信号通路等。

2.6 SSR 分析

在得到的 95220 条序列中筛选 SSR 位点共找到 SSR 位点 37751 个, 这些位点存在于 25520 条序列中, SSR 发生频率为 26.80%。其中单条序列中含多个 SSR 位点的序列有 7962 条, 复合型的 SSR 位点 3384 个。SSR 位点中单碱基重复、二碱基重复、三碱基重复、四碱基重复、五碱基重复、六碱基重复个数分别为 14919(39.52%)、14715(38.98%)、7577、488、32 和 20, 其中单碱基和二碱基重复含量居第一和第二位(见图8)。

2.7 SNP 分析

对获得的序列进行 SNP 位点分析, 共发现 3228575 个 SNP 标记(见表3), 其中包括 C:G->A:T、C:G->G:C、C:G->T:A、T:A->A:T、T:A->C:G、T:A->G:C6 中类型, 其中 T:A->C:G 和 C:G->T:A 这两种碱基替换占比较高, 分别占总数的 33.45%和 33.36%。7 个不同组

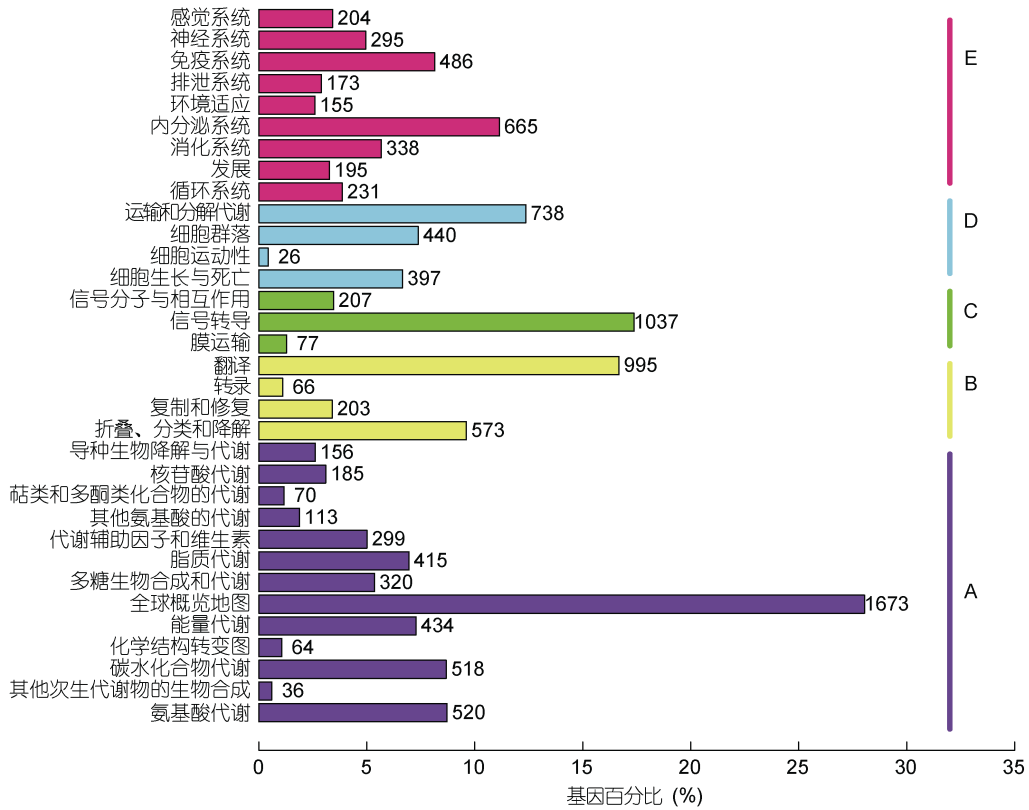


图 5 KEGG 注释

Fig.5 The KEGG annotation

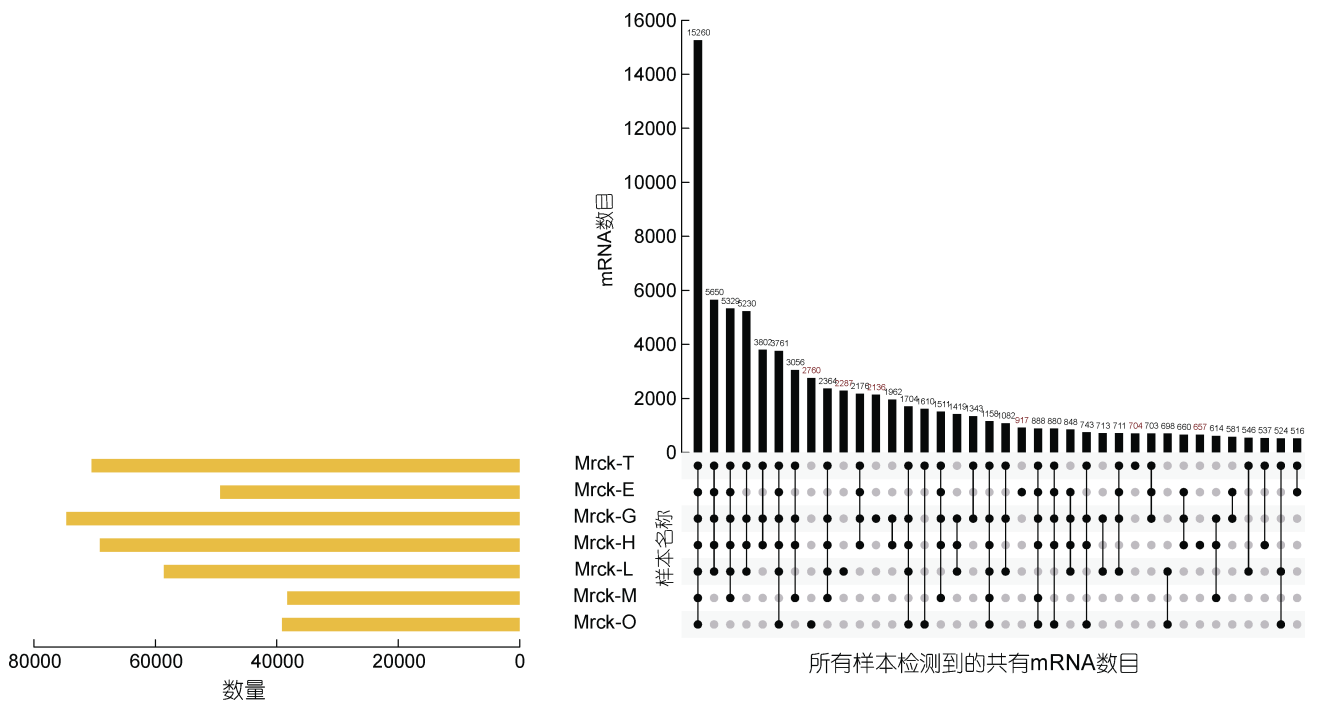


图 6 每个样本鉴定得到的 mRNA Upset 图

Fig.6 The upset map of mRNA from each sample

注: 柱顶数字表示每个样本鉴定到的全部 mRNA 的数目; 横坐标所有点的连线表示所有样本鉴定到的共有 mRNA 的数目, 其余的单点或者多个点连线, 均表示相关样本中鉴定到的特有 mRNA 的数目

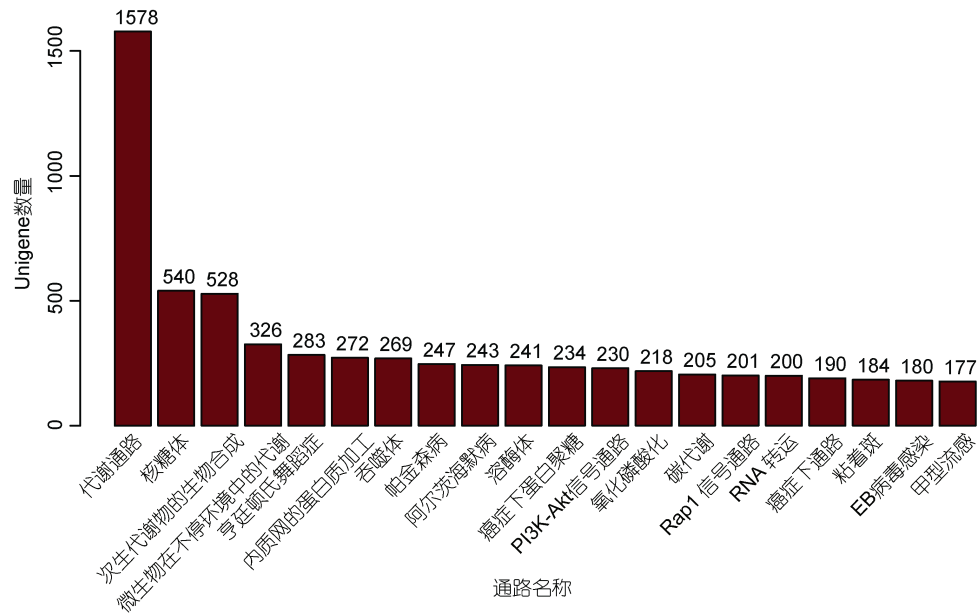


图 7 KEGG 通路列表(前 20)
Fig.7 The KEGG pathway list (Top 20)

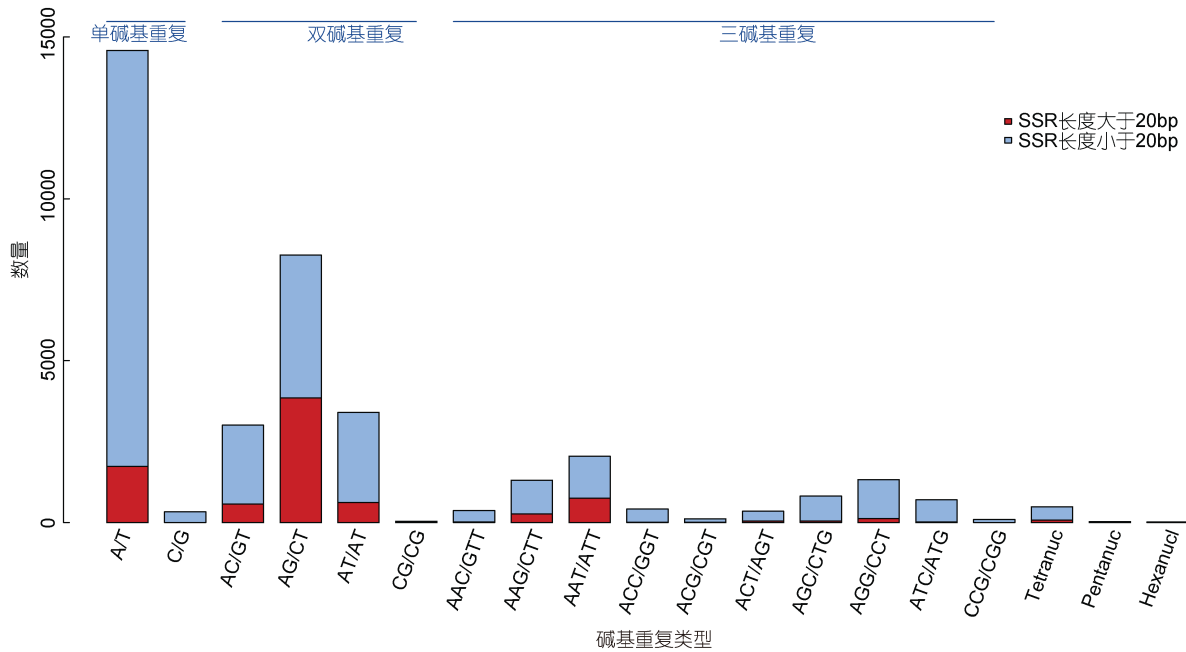


图 8 SSR 位点的分布情况
Fig.8 The distribution of SSR sites

织中发现的SNP位点数量在鳃(23.12%)上最高,其次是心脏(17.89%)、精巢(17.51%)和肝脏(16.28%),在肌肉组织中的SNP位点数量最少(5.43%)。

3 讨论

3.1 转录组技术在其他虾上的应用

随着转录组技术的迅速发展,转录组技术被应

用于水产动物的研究中。其中虾类转录组研究主要集中在差异基因的筛选和候选基因的发掘上。

红、黄和透明3种纯色米虾皮肤组织的转录组获得45434条Unigenes(林师, 2017)。波纹龙虾肝胰腺和卵巢组织进行了转录组测序获得Unigenes共74124个(李斌等, 2016)。脊尾白虾成虾样品及胚胎样品共获得47574条Unigenes(孙政, 2012)。

表 3 不同 SNP 类型在各组织之间的分布情况
Tab.3 Distribution of different SNP types among different tissues

类型	Mrck-E	Mrck-G	Mrck-H	Mrck-L	Mrck-M	Mrck-O	Mrck-T	合计
C:G->A:T	24923	62549	48000	42950	14379	29946	46525	269272
C:G->G:C	16852	41514	31809	28520	9995	19613	31305	179608
C:G->T:A	100152	250593	194474	178447	58143	106974	188259	1077042
T:A->A:T	32833	81389	62392	54151	19734	38621	61688	350808
T:A->C:G	102892	247796	192971	178419	58617	109014	190091	1079800
T:A->G:C	26037	62621	47996	43274	14450	30211	47456	272045
合计	303689	746462	577642	525761	175318	334379	565324	3228575

红螯螯虾肝脏、精巢和卵巢组织共获得了 67369 个 Unigenes(李喜莲等, 2019); 次级卵黄发生时期卵巢组织共得到 69261 条 Unigenes(康鹏飞, 2017)。

日本沼虾正常性成熟的和性早熟的卵巢组织中共获得 63336 个 Unigenes(江红霞, 2017); 感染白斑综合征病毒(WSSV)个体的肝胰腺转录组共获得 64049 个 Unigenes(赵才源等, 2018); 亚硝酸盐胁迫下肝脏共获得 19022 个 Unigenes, 氨氮胁迫下肝脏组织共获得 63453 个 Unigenes (于杰伦, 2019)。

凡纳滨对虾(*Litopenaeus vannamei*)五个早期发育时期(卵裂期、囊胚期、原肠期、肢芽幼体期发育至膜内无节幼体期)转录组测序, 共得到 66815 条 Unigenes(隗健凯, 2015); 性腺组织获得 Unigenes 65218 个(韦媛媛, 2017); 低温胁迫下肝胰腺组织获得 50921 条 Unigenes(董丽君等, 2019)。

本研究从正常个体 7 个组织转录组文库中共获得 99520 个 Unigenes, 这个数量较前人对于虾类转录组研究得到的 Unigenes 数量都高, 这可能与分析的组织数量大、覆盖面广有一定的关系。

3.2 功能注释

本研究从七个组织共获得 Unigenes 99520 个, N501553, 平均长度为 1064.9 bp。在五个数据库中注释到的 Unigenes 共有 20368 个, 占到总数的 21.39%, 这比前人研究的转录组结果 Unigenes 的注释率都低。肝胰腺注释率 31%和 35.31%, 卵巢 31.46%和 54.44%, 眼柄 29.3%, 淋巴器官 29.46%, 再生肢体基部注释率 37.23%。估计与本研究首次对罗氏沼虾的心脏、肌肉、精巢和鳃等组织进行转录组高通量测序, 引入了较多在这几个组织中特异表达的基因, 这些新获得的序列在以上五个数据库中得不到注释, 从而降低了注释率。还有部分序列未能被注释, 可能与序列的长度有关, 过短的序列也会造成无法注释和分类; 无

法注释的另一种可能是近缘物种序列信息的缺乏, 导致无法通过同源序列比对得到注释。

3.3 KEGG 通路

对罗氏沼虾转录组的 KEGG 分析显示, 9148 个 Unigenes 被注释到 KEGG 数据库中并分布在 330 个已知途径中, 与免疫相关的通路如 FoxO 信号通路和 Rap1 信号通路。FoxO 转录因子家族在细胞生理事件中调控基因的表达, 包括凋亡、细胞周期控制、葡萄糖代谢、氧化应激抵抗和寿命。FoxO 蛋白的一个主要调控机制是对磷脂酰肌醇 3 激酶(PI3K)下游的丝氨酸苏氨酸激酶 Akt/蛋白激酶 B(Akt/PKB)进行磷酸化, 这是对胰岛素或几种生长因子的反应。FoxO 信号通路在本研究结果中涉及 93 个 Unigenes。Rap1 是一种小型 GTPase, 它控制多种过程, 如细胞黏附, 细胞-细胞连接的形成和细胞极性。Rap1 通过调节多种细胞类型中整合素等黏附分子的功能, 在细胞-细胞和细胞-基质相互作用的控制中发挥主导作用。Rap1 还以高度依赖于细胞类型的方式调控 MAPK 活性。Rap1 信号通路在本研究中共涉及 201 个 Unigenes。这些结果的获得都将为进一步研究罗氏沼虾抗性相关基因提供科学依据。

4 结论

本研究对罗氏沼虾 7 个不同组织的转录组数据进行分析, 获得 99520 个 Unigenes, 这比单一组织或较少组织受到环境胁迫或细菌、病毒感染获得数据具有较高的可信度, 这一结果将大大丰富罗氏沼虾的基因数据库资源。与此同时, 20368 个 Unigenes 在五大数据库中得到注释。各个组织间差异基因也得到进一步的分析, 本研究还筛选得到大量的 SSR 位点和 SNP 位点, 这些分子标记也将在后续罗氏沼虾分析标记的开发中起到重要的作用, 为进一步深入挖掘

和开发利用罗氏沼虾功能基因提供参考。

参 考 文 献

- 于杰伦, 2019. 日本沼虾免疫应答相关基因的筛选与鉴定. 泰安: 山东农业大学博士学位论文
- 王传聪, 唐修阳, 项杰等, 2018. 罗氏沼虾转录组 SSR 标记信息分析. 江苏农业科学, 46(22): 56—59
- 韦媛媛, 2017. 凡纳滨对虾(*Litopenaeus vannamei*)亲虾性腺转录组分析及雌雄性腺差异表达基因鉴定. 南宁: 广西大学硕士学位论文
- 邓泽森, 欧江涛, 王资生, 2018. 罗氏沼虾重要免疫通路相关 microRNA 的转录组系统解析. 见: 中国生物工程学会第十二届学术年会暨 2018 年全国生物技术大会论文集. 长沙: 中国生物工程学会, 252—253
- 卢志杰, 叶成凯, Sarath Babu V 等, 2019. 罗氏沼虾谷氨酰胺合成酶基因的克隆与表达. 水产学报, 43(3): 549—562
- 田荣, 许婷, 潘晓艺等, 2016. 罗氏沼虾幼体精氨酸激酶基因的 cDNA 克隆及其在双顺反子病毒感染后的表达特征. 水生生物学报, 40(5): 908—913
- 冯艺, 2018. 罗氏沼虾养殖群体的形态学及多样性研究. 佛山: 佛山科学技术学院硕士学位论文, 92
- 吕敏, 黄光华, 李旻等, 2019. 异型雄性罗氏沼虾遗传多样性的微卫星分析. 水产科学, 38(3): 355—360
- 朱光来, 齐富刚, 王东博等, 2019. 益生菌对罗氏沼虾生长性能和水质的影响. 中国饲料, (20): 97—101
- 朱其建, 鞠波, 戴习林, 2019. 亚硝态氮对罗氏沼虾存活率、生长及能量代谢的影响. 水产科技情报, 46(1): 14—18
- 朱春华, 刘金磊, 李郁娇等, 2017. 壬基酚对罗氏沼虾卵黄蛋白原基因表达以及性腺发育的影响. 见: 2017 年中国水产学会学术年会论文摘要集. 南昌: 中国水产学会, 205
- 刘伟利, 刘红, 丁福江等, 2017. 罗氏沼虾 TLRs 基因表达差异分析. 南方农业学报, 48(4): 721—727
- 刘金磊, 2018. 雌激素相关受体(ERR)对雌性罗氏沼虾生殖的影响. 湛江: 广东海洋大学硕士学位论文
- 刘金磊, 邓思平, 江东能等, 2018. ERR-dsRNA 对罗氏沼虾卵巢中 ERR 及生殖相关基因表达的影响. 广东海洋大学学报, 38(3): 8—16
- 江红霞, 2017. 日本沼虾雌性早熟相关基因的筛选、克隆、表达与功能分析. 杨凌: 西北农林科技大学博士学位论文
- 江婷佳, 2017. 罗氏沼虾四种不同 DnaJ 基因克隆及表达分析. 杭州: 中国计量大学硕士学位论文
- 孙政, 张晓军, 张继泉等, 2012. 脊尾白虾转录组分析及重要功能基因的发掘. 见: 中国海洋湖沼学会第十次全国会员代表大会暨学术研讨会. 青岛: 中国海洋湖沼学会, 中国科学院海洋研究所, 119
- 孙成飞, 叶星, 董浚键等, 2015. 罗氏沼虾 6 个养殖群体遗传多样性的微卫星分析. 南方水产科学, 11(2): 20—26
- 严赛峰, 欧江涛, 王资生, 2018. 利用转录组技术系统挖掘罗氏沼虾免疫相关 SNP 标记. 见: 中国生物工程学会第十二届学术年会暨 2018 年全国生物技术大会论文集. 长沙: 中国生物工程学会, 245—246
- 李斌, 梁华芳, 陈兆明, 2016. 波纹龙虾转录组分析与 C-型凝集素的初步研究. 见: 2016 年中国水产学会学术年会论文摘要集. 成都: 中国水产学会, 四川省水产学会, 17
- 李郁娇, 2011. 壬基酚与三丁基锡对罗氏沼虾免疫因子、卵黄蛋白原基因表达和性腺发育的影响. 湛江: 广东海洋大学硕士学位论文
- 李威霖, 2018. 罗氏沼虾肢体再生转录组及其差异表达基因的基础研究. 湛江: 广东海洋大学硕士学位论文
- 李俊杰, 欧江涛, 王资生, 2018. 利用转录组技术系统挖掘罗氏沼虾 SSR 标记信息. 见: 中国生物工程学会第十二届学术年会暨 2018 年全国生物技术大会. 长沙: 中国生物工程学会, 243—244
- 李喜莲, 郭建林, 黄振远等, 2019. 红螯螯虾转录组高通量测序及分析. 四川农业大学学报, 37(2): 241—246, 265
- 杨光, 秦真东, 赵丽娟等, 2020. 罗氏沼虾细胞色素 P450 家族 CYP302a1 基因克隆及其在蜕皮周期中的表达. 水产学报, 44(04): 562—574
- 杨树浩, 冯艺, 陈建酬等, 2018. 不同比例的生物饲料对罗氏沼虾生长、消化酶和免疫酶活性的影响. 饲料工业, 39(2): 18—25
- 杨景丰, 郭子好, 蔡修兵等, 2019. 大豆抗原蛋白对罗氏沼虾生理生化及免疫的影响. 中国水产科学, 26(2): 322—332
- 吴维福, 李郁娇, 黎东等, 2013. 三丁基锡对罗氏沼虾卵黄蛋白原(VTG)表达和性腺发育的影响. 渔业科学进展, 34(4): 77—83
- 邱庆庆, 袁翔, 黄光华等, 2019. 罗氏沼虾热休克蛋白 HSP90 基因的克隆及表达分析. 黑龙江畜牧兽医, (19): 153—157, 185
- 何竺柳, 2018. 低盐度胁迫对罗氏沼虾生长和肉质的影响. 舟山: 浙江海洋大学硕士学位论文
- 张俊功, 2019. 温度及雌雄比对罗氏沼虾能量代谢和生长的影响. 上海: 上海海洋大学硕士学位论文
- 张剑伟, 王恒志, 张新党等, 2019. 饲料中添加不同梯度珍珠露水草提取物对罗氏沼虾生长、蛋白质代谢和肌肉营养成分的影响. 中国农学通报, 35(7): 158—164
- 张夏青, 许建和, 潘茜等, 2016. 罗氏沼虾组织蛋白酶 D 基因的克隆及其组织表达分析. 淮海工学院学报(自然科学版), 25(4): 71—77
- 陈建酬, 高腾, 张济培等, 2017. 养殖密度和投饵策略对罗氏沼虾幼体生长发育的影响. 江西农业学报, 29(1): 92—95, 101
- 陈雪峰, 王春琳, 顾志敏等, 2019. 罗氏沼虾(*Macrobrachium rosenbergii*)卵巢发育不同时期转录组分析. 海洋与湖沼, 50(2): 398—408
- 林师, 2017. 中华锯齿米虾色素沉积相关基因的发掘与证明. 厦门: 集美大学硕士学位论文
- 周晓敏, 戴习林, 2020. 罗氏沼虾 SNP 标记筛选及不同群体的遗传多样性. 水产学报, <https://www.cnki.net/KCMS/detail/31.1283.S.20200331.1205.006.html>
- 单凡, 刘波, 孙存鑫等, 2019. 饲料维生素 E 含量对罗氏沼虾生长性能、血清生化指标和肌肉品质的影响. 动物营养学报, 31(1): 214—225
- 赵才源, 钱珺, 傅洪拓等, 2018. 白斑综合症病毒(WSSV)感染日本沼虾肝胰腺的转录组研究. 见: 2018 年中国水产学会学术年会论文摘要集. 西安: 中国水产学会, 199
- 赵巨泽, 2019. 微生态制剂对罗氏沼虾生长、免疫及肠道健康的影响研究. 扬州: 扬州大学硕士学位论文
- 赵苗鑫, 陈华谱, 刘金磊等, 2017. 罗氏沼虾雌激素相关受体(ERR)基因原核表达与纯化. 广东海洋大学学报, 37(1): 108—112

- 姜建萍, 袁翔, 邱庆庆等, 2019. 罗氏沼虾性别相关基因研究进展及其单性化养殖现状. 南方农业学报, 50(09): 2111—2118
- 俞炎琴, 2013. 罗氏沼虾中性别发育相关基因 Sxl 和 Dmrt 基因的分子特征和功能研究. 杭州: 浙江大学博士学位论文
- 郭梁, 季文博, 郝军等, 2018. 感病罗氏沼虾免疫通路关键基因的转录组分析. 见: 中国生物工程学会第十二届学术年会暨 2018 年全国生物技术大会论文集. 长沙: 中国生物工程学会, 中国生物工程学会, 138—148
- 黄黎明, 吕敏, 杨慧赞等, 2019. 不同糖源对罗氏沼虾生长性能和糖代谢的影响. 动物营养学报, 31(12): 5635—5644
- 康鹏飞, 2017. 克氏原螯虾卵巢转录组分析及 RDH13 基因在发育过程中的作用. 武汉: 华中师范大学硕士学位论文
- 魏健凯, 2015. 组学解析凡纳滨对虾早期发育相关重要生物学过程的分子机制. 北京: 中国科学院研究生院(海洋研究所)博士学位论文
- 董丁健, 戴习林, 2020. 罗氏沼虾不同群体世代遗传多样性的 SSR 分析. 南方农业学报, 51(2): 421—428
- 董丽君, 孟宪红, 孔杰等, 2019. 基于转录组分析筛选凡纳滨对虾低温胁迫下的差异表达基因. 中国水产科学, 26(1): 161—171
- 蒋飞, 戴习林, 丁福江等, 2016. 罗氏沼虾 4 个不同群体的生长和存活率比较. 河北渔业, (6): 12—14
- 蒋飞, 戴习林, 朱其建等, 2013. 罗氏沼虾 5 个专门化品系选择系 F₂ 生长比较及其遗传结构分析. 见: 2013 年中国水产学会学术年会论文集. 合肥: 中国水产学会, 49
- 蒋飞, 戴习林, 朱其建等, 2014. 罗氏沼虾 5 个专门化品系选择系生长比较及其遗传结构分析. 上海海洋大学学报, 23(3): 329—337
- 程安达, 吕慧源, 姚嘉赞等, 2019. 一种中药复方制剂在饲料中长期添加对罗氏沼虾生长及免疫功能的影响. 饲料研究, 42(6): 12—16
- 潘忠超, 李志华, 孙凤刚等, 2019. 植物甾醇对罗氏沼虾生长性能、消化酶、肌肉品质及氨基酸组成的影响. 江西水产科技, (6): 3—7
- 薛海波, 2010. 环境内分泌干扰物对罗氏沼虾生长发育及组织结构的影响. 湛江: 广东海洋大学硕士学位论文
- 戴习林, 刘洁, 李晶晶等, 2017. 罗氏沼虾种群 SSR 分析中样本量及标记量对遗传多样性指标的影响. 水产学报, 41(7): 1083—1095
- 戴习林, 周梦颖, 鞠波等, 2016. 养殖密度对罗氏沼虾生长、性别分化与性腺发育的影响. 水产学报, 40(12): 1874—1882
- Abayed F A A, Manor R, Aflalo E D *et al*, 2019. Screening for Dmrt genes from embryo to mature *Macrobrachium rosenbergii* prawns. General and Comparative Endocrinology, 282: <https://doi.org/10.1016/j.ygcen.2019.06.009>
- Cao J, Wu L, Jin M *et al*, 2017. Transcriptome profiling of the *Macrobrachium rosenbergii* lymphoid organ under the white spot syndrome virus challenge. Fish & Shellfish Immunology, 67: 27—39
- Chen T, Lv M, Wu Y G *et al*, 2019. A new growth curve model for giant freshwater prawn *Macrobrachium rosenbergii* in a prawn-plant symbiotic system. Ecological Modelling, 411: 108801, doi: 10.1016/j.ecolmodel.2019.108801
- Dong X X, Liu Q G, Kan D Q *et al*, 2020. Effects of ammonia-N exposure on the growth, metabolizing enzymes, and metabolome of *Macrobrachium rosenbergii*. Ecotoxicology and Environmental Safety, 189: 110046, doi: 10.1016/j.ecoenv.2019.110046
- Feng P F, He J Z, Lv M *et al*, 2019. Effect of dietary *Tenebrio molitor* protein on growth performance and immunological parameters in *Macrobrachium rosenbergii*. Aquaculture, 511: 734247, doi: 10.1016/j.aquaculture.2019.734247
- Guo H, Chen L L, Li G L *et al*, 2019. Accumulation and depuration of nonylphenol and its effect on the expressions of vitellogenin and vitellogenin receptor in freshwater prawn *Macrobrachium rosenbergii*. Bulletin of Environmental Contamination and Toxicology, 103(5): 729—733
- Jiang J P, Yuan X, Qiu Q Q *et al*, 2019a. Comparative transcriptome analysis of gonads for the identification of sex-related genes in giant freshwater prawns (*Macrobrachium rosenbergii*) using RNA sequencing. Genes, 10(12): 1035
- Jiang Q, Min Y, Yang H *et al*, 2019b. De novo transcriptome analysis of eyestalk reveals ovarian maturation related genes in *Macrobrachium rosenbergii*. Aquaculture, 505: 280—288
- Lafontaine A, Gismondi E, Dodet N *et al*, 2017. Bioaccumulation, distribution and elimination of chlordecone in the giant freshwater prawn *Macrobrachium rosenbergii*: Field and laboratory studies. Chemosphere, 185: 888—898
- Lafontaine A, Hanikenne M, Boulangé-Lecomte C *et al*, 2016. Vitellogenin and vitellogenin receptor gene expression and 20-hydroxyecdysone concentration in *Macrobrachium rosenbergii* exposed to chlordecone. Environmental Science and Pollution Research, 23(20): 20661—20671
- Manickam N, Bhavan P S, Santhanam P *et al*, 2020. Influence of wild mixed zooplankton on growth and muscle biochemical composition of the freshwater prawn *Macrobrachium rosenbergii* post larvae. Aquaculture, 522: 735110, doi: 10.1016/j.aquaculture.2020.735110
- Pasookhush P, Hindmarch C, Sithigorngul P *et al*, 2019. Transcriptomic analysis of *Macrobrachium rosenbergii* (giant fresh water prawn) post-larvae in response to *M. rosenbergii* nodavirus (MrNV) infection: de novo assembly and functional annotation. BMC Genomics, 20(1): 762
- Rao R, Bhassu S, Bing R Z Y *et al*, 2016. A transcriptome study on *Macrobrachium rosenbergii* hepatopancreas experimentally challenged with white spot syndrome virus (WSSV). Journal of Invertebrate Pathology, 136: 10—22
- Stalin A, Suganthi P, Mathivani S *et al*, 2019. Effect of cobalt-60 gamma radiation on reproductive disturbance in freshwater prawn *Macrobrachium rosenbergii* (De Man, 1879). Toxicology Reports, 6: 1143—1147
- Tadese D A, Sun C X, Liu B *et al*, 2020. Combined effects of emodin and *Clostridium butyricum* on growth and non-specific immunity of giant freshwater prawns, *Macrobrachium rosenbergii*. Aquaculture, 525: 735281, doi: 10.1016/j.aquaculture.2020.735281
- Tan K, Jiang H G, Jiang D H *et al*, 2019. Sex reversal and the androgenic gland (AG) in *Macrobrachium rosenbergii*: A review. Aquaculture and Fisheries, 4(6), doi: 10.1016/j.aaf.2019.11.004

COMPARATIVE TRANSCRIPTOME ANALYSIS OF DIFFERENT TISSUES IN GIANT FRESHWATER PRAWNS (*MACROBRACHIUM ROSENBERGII*) BASED ON RNA-SEQ

LI Xi-Lian, GU Zhi-Min, SHEN Pei-Jing, XU Yang, ZHANG Yu-Fei, GAO Qiang,
CHENG Hai-Hua, CHEN Xue-Feng

(Key Laboratory of Healthy Freshwater Aquaculture of Ministry of Agriculture and Rural Development, Key Laboratory of Freshwater Aquatic Animal Genetic and Breeding of Zhejiang Province, Zhejiang Institute of Freshwater Fisheries, Huzhou 313001, China)

Abstract The giant freshwater prawns (*Macrobrachium rosenbergii*) is an important freshwater shrimp breeding species in China. By using Illumina Miseq platform, transcriptome from seven tissues/organs (eyestalk, liver, ovary, gill, heart, muscle, and testis) were analyzed, from which high quality clean reads (36325476, 56796932, 36328098, 51370140, 45010606, 43240160, and 42404294) were obtained, respectively, in total of 46.2Gb data after quality control. A total of 95220 Unigenes were obtained by the Denovo assembly. The average length of each Unigenes was 1064.9 bp. After the bioinformatics method was compared with five databases (NR, GO, COG, KEGG, SWSS-PROT), 20368 Unigenes were annotated. Among them, GO functional annotation divided Unigenes into three categories: biological process, cell component and molecular function, and 50 branches. In comparison with COG database, 15798 items were compared to homologous sequences, and were divided into 25 categories. The KEGG metabolic pathway includes 6 categories and 33 subcategories, which can map Unigenes to 330 metabolic pathways. After molecular marker screening, 37751 potential SSR sites were obtained. In addition, 3228575 SNP markers were found. Transcriptome sequencing, data assembly, and functional annotation were performed on 7 different tissues of *M. rosenbergii*, providing references for further exploration and development of functional genes of *M. rosenbergii*.

Key words *Macrobrachium rosenbergii*; transcriptome; RNA-seq; different tissues; functional genes; comparative analysis