

文章编号: 1000-0615(2019)01-0015-21

DOI: 10.11964/jfc.20181211602

· 综述 ·

水产生物基因组研究进展与趋势

郑先虎, 匡友谊, 吕伟华, 栾培贤, 孙志鹏, 鲁翠云, 孙效文*

(中国水产科学研究院黑龙江水产研究所, 淡水鱼类育种国家地方联合工程实验室,
农业农村部淡水水产生物技术与遗传育种重点实验室, 黑龙江 哈尔滨 150070)

摘要: 本文综述了水产生物基因组研究相关技术的发展历程和关键技术的应用。以二代测序技术的出现为分界点, 首先回顾了早期水产养殖生物在遗传学和分子生物学方面的研究结果, 以及为开展全基因组测序所做的相关基础研究, 然后重点介绍了二代测序技术应用于水产生物全基因组测序和经济性状遗传基础解析的研究进展, 最后展望了水产生物基因组研究发展趋势。水产生物经济性状遗传机制高度复杂, 从全基因组角度阐明其遗传机制仍有很多难题, 但基因决定性状是生物学法则之一, 探索这一过程的奥秘引人入胜。

关键词: 水产生物; 基因组; 下一代测序

中图分类号: Q 785;S 917.4

文献标志码: A

水产生物基因组研究的最终目的是获得经济性状的决定基因或标记以用于培育性状优异的新品种, 或是利用基因组知识进行疾病防治和营养生理调控, 最终提高水产业的经济效益和生态效益。水产生物的基因组研究开展较晚, 借助于二代测序技术, 目前已有长牡蛎(*Crassostrea gigas*)、鲤(*Cyprinus carpio*)、草鱼(*Ctenopharyngodon idellus*)、半滑舌鳎(*Cynoglossus semilaevis*)、虹鳟(*Oncorhynchus mykiss*)和大西洋鲑(*Salmo salar*)等上百种水产生物的全基因组测序完成, 为从全基因组层面探索经济性状的遗传机制提供了信息平台 and 深入研究的基础, 使水产生物的种质鉴定、品种选育等研究进入到了基因组时代。第二代测序技术的突出特点是以高通量、低成本为主, 其诞生极大地推动了基因组学及相关领域的快速发展。第二代测序技术也是目前已完成全基因组测序的水产生物所用的主要方法, 因此本文以第二代测序技术的应用为分界点, 综述水产生物基因组及相关领域研究进展及趋势。

1 DNA测序技术发展概述

从1977年第一代DNA测序技术, 发展至今四十年时间, 测序技术已取得了相当大的发展, 在2007年以前, DNA测序数据呈指数增长, 而近十年则更是达到了超指数增长。从第一代到第三代, 测序技术的每一次变革, 都对生物学、基因组、疾病医疗、育种等领域产生巨大的推动作用。不同发展阶段DNA测序平台基本特性和性能情况见表1。

第一代测序技术出现的标志是1977年Maxam等^[1]发明的DNA化学降解测序法。同年, Sanger等^[2]建立了DNA双脱氧链末端终止测序法, 并测定了第一个基因组序列(噬菌体X174)。自此, 人类步入基因组学时代。20世纪80年代, 在Sanger法理论上, 出现了荧光自动测序技术, 使DNA测序效率有了显著的提高。1987年, Martin^[3]开发出了基于荧光双脱氧链终止法的DNA快速自动测序系统。以上测序技术均被称为第一代测序技术。2001年完成的首个人类基因组图谱所

收稿日期: 2018-12-14 修回日期: 2018-12-28

资助项目: 国家自然科学基金(31672655, 31602150)

通信作者: 孙效文, E-mail: sunxw2002@163.com

表 1 测序平台基本特性和性能

Tab. 1 Basic features and performances of sequence platforms

测序平台/公司/每运行周期最大产出 platforms/company/max output per run	读长长度/bp read length per run	读长数量 no. reads per run	运行时间 time	每兆碱基费用 cost per 10 ⁶ bases	原始数据错误率/% raw error rate	测序平台费用/美元 platform cost	测序方法 sequence method
第一代测序 first generation							
Sanger/Life technologies/84 kb	800	1	2 h	2 400	0.3	95 000	双脱氧链终止法 dideoxy terminator
第二代测序 second generation							
454 GS FLX+/Roche/0.7 Gb	700	1×10 ⁶	24/48 h	10	1	500 000	焦磷酸测序 pyrosequencing
GS Junior/Roche/70 Mb	500	1×10 ⁵	18 h	9		100 000	焦磷酸测序 pyrosequencing
HiSeq/Illumina/1 500 Gb	2×150	5×10 ⁹	27/240 h	0.1	0.8	750 000	可逆终止子 reversible terminators
MiSeq/Illumina/15 Gb	2×300	3×10 ⁸	27 h	0.13	0.8	125 000	可逆终止子 reversible terminators
SOLiD/Life technologies/120 Gb	50	1×10 ⁹	14 d	0.13	0.01	350 000	连接法 ligation
retrovirology/BGI/3 000 Gb	50	1×10 ⁹	14 d	0.01	0.01	12×10 ⁶	纳米球/连接 nanoball/ligation
ion Proton/Life technologies/100 Gb	200	6×10 ⁷	2~5 h	1	1.7	215 000	质子检测 proton detection
Ion PGM/Life technologies/2 Gb	200	5×10 ⁶	2~5 h	1	1.7	80 000	质子检测 proton detection
第三代测序 third generation							
SMRT/PacBio/1 Gb	>10 000	1×10 ⁶	1~2 h	2	12.9	750 000	实时单分子测序 real-time SMS
heliscope/helicos/25 Gb	35	7×10 ⁹	8 d	0.01	0.2	1.35×10 ⁶	实时单分子测序 real-time SMS
nanopore/Oxford Nanopore technologies/1 Gb	>5 000	6×10 ⁴	48/72 h	<1	34	1 000	实时单分子测序 real-time SMS
electron microscopy/ZS	7 200		14 h	<0.01		1×10 ⁶	实时单分子测序 real-time SMS

注: 参考自文献[9], 采用人类基因组(3×10⁹ bp或3 Gb)为参考, 比较各测序平台的产出和性能
Notes: From reference[9]. For comparison of the NGS outputs, the human genome has 3×10⁹ bp or 3 Gb

采用的测序技术正是基于Sanger法。一代测序技术至今在世界范围仍在使用, 主要用于PCR产物测序、文库克隆末端测序、小片段序列分析和基因分型等研究, 在一些物种物理图谱和遗传图谱构建中发挥了重要作用。红鳍东方鲀(*Takifugu rubripes*)全基因组测序采用第一代测序技术完成^[4]。

进入21世纪, 以Roche公司的454技术、Illumina公司的Solexa, Hiseq技术和ABI公司的SOLiD技术为标志的第二代测序技术诞生^[5]。与第一代相比, 第二代测序技术不仅保持了高准确度, 而且成本大大降低, 测序速度也大幅提高。2005年Roche公司推出了基于焦磷酸测序原理的Genome SequencerFLX高通量测序系统。2006年Solexa GA分析仪面世, 使用的方法是边合成边测序以及克隆单分子阵列技术。2008年第一代SOLiD测序仪发布, 其测序技术创新之处在于“双碱基编码”的应用, 使每个碱基都被检测两次, 降低错误率, 可使准确率大于99.94%。HeliScope测序仪在第二代的基础上引入了单分子测序的概念, 被称为2.5代。其方法克服了第二代测序技术中需要用PCR扩增来增强荧光信号的技术难题^[6]。第二代测序技术极大地推进了基因组学的发展, 更多物种的基因组组装、重测序、甲基化、转录本、宏基因组等研究得以开展。大熊猫是首个采用Solexa测序技术完成基因组测序的物种^[7], 第二代测序技术也是目前已完成水产生物全基因组测序所用的主要方法。

第二代测序技术虽然已经得到了广泛的应用, 在各技术方面趋于成熟, 但是存在依赖于模板扩增、荧光分析、序列读长限制等缺点, 以及不可避免的系统误差, 这些问题催生了第三代测序技术的诞生。第三代测序技术主要包括Helicos公司的真正单分子测序技术、Oxford Nanoporetech公司的单分子纳米孔测序技术、Pacific Biosciences (PacBio)公司的单分子实时测序技术等^[8]。三代测序技术在复杂基因组测序(多倍体或大量重复序列)、甲基化识别, 及RNA直接测序等领域具有显著优势。随着第三代测序技术的不断完善和发展, 测序时间更短、成本更低廉、灵活性更强、通量更高、读取长度更长、测序质量更高等一系列目标的实现, 必将推动基因组学研究快速发展。

2 二代测序技术应用前水产生物基因组研究

在20世纪90年代, 大西洋鲑最先获得基因组研究资助, 稍后, 美国农业部组织实施了5个种类[牡蛎、虹鳟、凡纳滨对虾(*Litopenaeus vannamei*)、尼罗罗非鱼(*Oreochromis niloticus*)和斑点叉尾鲷(*Ictalurus punctatus*)]的基因组计划, 再稍后鲤、半滑舌鳎、大黄鱼(*Larimichthys crocea*)、中国明对虾(*Fenneropenaeus chinensis*)等我国的主要养殖种都开展了以全基因组测序为目标的前期研究。这些研究多数是以人类基因组研究策略为模板开展的, 如开发标记、建立高密度图谱、性状的连锁分析、BAC库和EST库构建与测序等。

2.1 遗传标记的开发

DNA分子标记有多种, 目前使用较广泛的是微卫星(simple sequence repeats, SSR)标记和单核苷酸多态性(single nucleotide polymorphism, SNP)标记。SNP标记比较适用于规模化和自动化分析, 而SSR标记更易于实验室应用。基因组中存在大量的微卫星序列是20世纪70年代在大寄居蟹(*Pagurus pollicaris*)的基因组中发现的^[10], 由于微卫星序列具有在基因组中分布广泛且中性等特征, 迅速被用于人及模式生物的遗传作图和性状连锁分析, 但用到水产生物的遗传作图是在20世纪的90年代后期, 如尼罗罗非鱼^[11]、虹鳟^[12]等遗传图谱。

由于早期微卫星标记开发比较困难且成本较高, 我国水产领域直到本世纪初才有制备微卫星标记的报道。如2001年魏东旺等^[13]报道了鲤微卫星的筛选, 徐鹏等^[14]克隆了中国明对虾的微卫星序列。在物种表达序列标签(expressed sequence tag, EST)数据丰富的前提下, I型微卫星(位于基因组外显子区域)发掘相对便捷, I型微卫星为功能基因提供直接的分子标记, 其多态性能够更好地解释个体的性状差异。因此, I型微卫星标记是非常有价值的标记资源, I型标记筛选和与性状连锁关系的报道很多, 如, Serapion等^[15]利用43 033条斑点叉尾鲷EST序列发掘了4 103个微卫星标记。I型标记在遗传作图中也很有优势, 遗传图谱中使用I型标记会提高图谱的利用价值, 如果图谱的标记全是I型标记即可称作基因图谱。

微卫星标记具有多态性高、共显性遗传、

符合孟德尔分离方式、实验室操作简便等诸多优点,尤其是基因型的分型简单方便,因此目前仍是种质鉴定和性状遗传分析的重要工具。

2.2 表达序列标签开发

EST是通过随机挑选cDNA文库进行单向测序所获得的序列(约60~500 bp),其上携带着表达基因的部分遗传信息。与基因组序列相比,EST序列具有快速、简便、易得的优势,是基因组学研究领域中不可或缺的一项内容。由于EST在研究上的重要性,1993年NCBI建立了EST数据库dbEST(database of EST),系统收集和保存所有EST数据。

EST在基因组学中主要应用于遗传学图谱构建(遗传图谱、物理图谱和基因图谱)、基因定位、筛选与克隆新基因、分子标记筛选和基因表达系列分析等方面^[16]。如,张晓峰等^[17]利用EST序列进行鲤生长相关SNP标记的发掘,鉴定出69个与生长相关的SNP标记;Zhang等^[18]利用EST序列鉴定到虹鳟LEAP-2基因;Schalburg等^[19]利用大西洋鲑和虹鳟的EST序列构建了16 K的鲑微阵列基因表达芯片;Naruse等^[20]利用青鳉(*Oryzias latipes*)绘制了818条基因和EST序列的图谱,研究了脊椎动物[青鳉、斑马鱼(*Danio rerio*)和人]间垂直同源的关联性,为辐鳍鱼及硬骨鱼共同祖先的基因组复制机制提供理论依据。

2.3 物理图谱构建

物理图谱是指利用限制性内切酶将染色体切成片段,再根据重叠序列确定片段间连接顺序,以及遗传标记间物理距离(以碱基对、千碱基或兆碱基为单位)的图谱。这些标记可以是酶切位点、DNA片段或者特定的基因,其反映的是DNA片段在染色体上的实际位置。物理图谱构建技术始于20世纪80年代,其最早构建技术是染色体带型分析,这是分辨率最低的物理图谱。依其产生的方法不同物理图谱分为:限制性酶切图谱、重叠群图谱和DNA序列图谱等3类。大西洋鲑^[21]的物理图谱是采内切酶酶切方法属于限制性酶切图谱。目前常见的物理图谱是在BAC文库基础上采用DNA指纹技术构建的基因组物理图谱,它能将众多片段较小的BAC克隆装配成相互叠加,反映基因组实际排列顺序的一系列重叠群^[22]。如斑点叉尾鲷^[23]、虹鳟^[24]、

鲤^[25]等物理图谱。

物理图谱对于采用传统分级鸟枪法测序进行全基因组组装来说是必不可少的,物理图谱的一个重要作用就是确定测序的最小重叠路径,随后依照这种最小重叠路径上的BAC克隆进行逐一测序并组装。虽然目前全基因组测序和组装对物理图谱和大片段文库的依赖性大大降低,但在当时测序技术条件下对全基因组装配起到了重大作用。Lewin等^[26]也研究认为好的物理图谱和高密度连锁图谱对于全基因组序列的精确装配均非常必要。

2.4 遗传连锁图谱构建

遗传连锁图谱是通过计算遗传重组率并将其换算为遗传距离标示图距,用于显示基因/标记在染色体上相对位置的图谱。图距一般用厘摩(cM,即每次减数分裂的重组频率为1%)来表示。遗传连锁图谱在基因组研究中具有重要作用,通过遗传图谱可以了解基因/标记在基因组中的大体位置,作为坐标来辅助基因组装配。图谱质量取决于定位在图谱上的标记数量,标记数量多、标记间距小,表示图谱密度越高,图谱质量越好。因此,构建高质量的遗传图谱是基因组研究的重要内容。

在20世纪90年代,水产生物的SSR标记的鉴定数量还远不能满足高密度遗传连锁图谱的构建,多数水产生物图谱采用随机扩增多态性DNA标记(random amplified polymorphic DNA, RAPD)、扩增片段长度多态性(amplified fragment length polymorphism, AFLP)等显性标记或与SSR标记联合作图,特点是标记间距较大,图谱密度较低。如,尼罗罗非鱼^[11]、虹鳟^[12]、鲤^[27-28]、长牡蛎^[29]等遗传连锁图谱,都属低密度图谱。

随着分子标记鉴定数量的增加,多种水产生物构建了更高密度的连锁图谱。如,Nichol等^[30]报道的虹鳟遗传连锁图谱含有1 359个标记(包括SSR、AFLP等标记);Lee等^[31]发表了罗非鱼图谱,含有546个SSR标记和基因标记,平均标记间隔2.4 cM。Moen等^[32]发表的大西洋鲑遗传连锁图谱,包含138个SSR标记和304个SNP标记;Xia等^[33]发表的草鱼遗传图谱,包含263个SSR标记和16个SNP标记,平均标记间隔4.2 cM。这些结果表明,在二代测序技术应用前,水产生物的图谱只达到了中等密度水平,由于鉴定多态性标

记的工作量太大、成本过高, 限制了遗传图谱密度的增加。

2.5 性状连锁分析

水产生物遗传研究的主要目标是获得与经济性状紧密连锁的基因和标记, 因此, 在第一代图谱出现时就有许多QTL的研究报道, 一些非数量性状如抗病性状也利用此技术获得了相关标记。耐热性状可能是最早利用QTL技术研究的性状, 1998年, Jackson等^[34]利用SSR标记分析了与虹鳟3个家系耐热性状相关性, 鉴定到2个位点与耐热有关。Yue^[35]和Laghari等^[36]已经对水产生物QTL的研究进行了全面的综述, 本文仅对品质、饲料转化率和抗病性状QTL研究作简要介绍。

品质的提升和饲料转化率的提高是我国水产动物需要改进的性状, 与品质有关的QTL报道较多, 如, Derayat等^[37]利用11个半同胞家系定位到了1个与大西洋鲑肌肉脂肪含量相关的位点; Baranski等^[38]定位到了3个与大西洋鲑肉色相关的QTL座位; Kuang等^[39]开展了鲤脂肪含量性状QTL定位, 获得了1个基因组水平QTL和3个染色体水平QTL。饲料转化率报道较少, 主要集中在鲤, 李鸥等^[40]定位到了7个与鲤鱼饲料转化率相关的位点, 其中3个显著相关; 张丽博等^[41]获得了与鲤饲料转化率显著相关的标记10个, 2个为EST标记; 王宣朋等^[42]鉴定到与镜鲤饲料转化率相关的标记27个。

抗病性状是阈性状中非常重要的性状, 我国水产养殖业的病害比较严重, 药物治疗和疫苗防治都非常有效, 但培育抗病品种也是一种行之有效的有效手段。日本东京海洋大学的Okamoto教授团队在抗病性状的QTL研究中成果非常突出, 他们对虹鳟^[43]、牙鲆(*Paralichthys olivaceus*)等^[44]抗病性状进行遗传定位, 利用定位的标记进行品种选育^[45], 获得了抗病品种, 这项研究成功的关键是他们在长期研究中分别建立了抗病和易感病的雌核纯系, 使鉴定到的位点是真实和准确的。我国学者在鱼类抗病QTL方面也做了大量研究, 如, Xu等^[46]鉴定了牙鲆抗病毒的基因位点, Yu等^[47]对美洲牡蛎(*Crassostrea virginica*)抗病基因进行了筛选。

2.6 模式水生生物全基因组测序

在测序技术成本较高的时期, 仅人和模式

生物等少数物种完成了全基因组测序。水产生物报道较少, 但与水产有关的几个模式生物完成了全基因组测序, 如斑马鱼、红鳍东方鲀^[44]、青鳉^[48], 其中斑马鱼只是在网上公布了基因组装配结果, 文章是利用了二代测序数据后才发表^[49,50]。这几个模式鱼类的全基因组测序结果为养殖鱼类基因组研究提供了重要参考平台, 如基因组进化、候选基因鉴定等比较基因组研究。

在世界范围, 对水产生物基础研究投入都较少, 尤其是在测序费用较高时, 只有斑点叉尾鲟、虹鳟、大西洋鲑和罗非鱼等少数物种积累了相对较多分子标记, 但构建的遗传图谱的标记密度也都比较低, 如罗非鱼第二代图谱仅有546个标记^[31]。

综上所述, 在二代测序技术应用之前, 各国科学家为了最终获得相应研究对象的全基因组信息, 做了大量的工作。即使相关数据最为丰富的大西洋鲑, 也没有完成全基因组序列的组装。总之, 水产生物全基因组测序工作的完成需要技术上的重大突破, 也需要成本上的大大降低。

3 二代测序技术应用后水产生物基因组研究

3.1 全基因组精细图谱绘制

全基因组测序是破译水产生物基因组密码、挖掘基因组资源、鉴定功能基因的关键步骤。通过全基因组测序可以获得某种生物基因组和重要功能基因序列信息, 揭示该物种进化史, 了解该物种生长、生殖、发育和环境适应的分子机制。随着以Solexa、454和SOLiD为代表的第二代测序技术的发展, 使测序的价格越来越低, 单个run的通量越来越高, 极大促进了水产基因组学的发展, 截至2018年11月, 在公共数据库中发布了210种水产生物的全基因组序列, 近5年已发表的主要水产生物精细图谱见表2。

2011年Bastiaan等^[51]利用454测序技术装配了高质量的大西洋鲑(*Gadus morhua*)的基因组, 发现了大西洋鲑具有血红蛋白基因簇的复杂缺陷和不寻常的免疫结构。为进一步研究硬骨鱼类免疫系统的进化, Malmström等^[52]采用Illumina测序技术构建了66种硬骨鱼类的基因组草图, 发现MHC II基因簇在整个鲑形目中都缺失了, 而

表2 近5年已发表的主要水产生物全基因组精细图谱一览表

Tab. 2 The list of published fine genome maps of major aquatic organisms in the past 5 years

物种 species	基因组大小 genome size	contig N50	scaffold N50	编码基因数量 coding gene no.	参考文献 reference
虹鳟 <i>Oncorhynchus mykiss</i>	1.9 Gb	7.7 kb	384 kb	46 585	[53]
罗非鱼 <i>Oreochromis niloticus</i>	815 Mb		2.8 Mb	21 437	[57]
半滑舌鲷 <i>Cynoglossus semilaevis</i>	477 Mb	26.5 kb	867 kb	21 256	[64]
鲤 <i>Cyprinus carpio</i>	1.69 Gb	68.4 kb	1 Mb	52 610	[65]
欧洲鲈 <i>Dicentrarchus labrax</i>	675 Mb	53 kb	5.1 Mb	26 719	[58]
菊黄东方鲀 <i>Takifugu flavidus</i>	390 Mb	2.8 kb	305.7 kb	30 285	[79]
南极抗冻鱼 <i>Notothenia coriiceps</i>	637 Mb	11.6 kb	219 kb	32 260	[82]
白斑狗鱼 <i>Esox lucius</i>	878 Mb	16.9 kb	700.5 kb	19 601	[83]
大黄鱼 <i>Larimichthys crocea</i>	728 Mb	25.7 kb	498.7 kb	19 362	[66]
	679 Mb	63.1 kb	1.03 Mb	25 401	[67]
草鱼 <i>Ctenopharyngodon idellus</i>	♀0.90 Gb ♂1.07 Gb	♀4.78 kb ♂18.2 kb	♀6.46 Mb ♂2.28 Mb	27 263	[68]
海带 <i>Saccharina japonica</i>	537 Mb	58.8 kb	252 kb	18 733	[73]
亚洲龙鱼 <i>Scleroopages formosus</i>	900 Mb		58.8 kb	24 274	[84]
银鲱 <i>Pampus argenteus</i>	550 Mb	0.5 kb	1.58 kb	16 322	[85]
大西洋鲑 <i>Atlantic salmon</i>	2 970	57.6 kb	2.97 Mb	37 206	[56]
斑点叉尾鲷 <i>Ictalurus punctatus</i>	783 Mb	77.2 kb	7.73 Mb	26 661	[55]
大菱鲆 <i>Scophthalmus maximus</i>	568 Mb	31.2 kb	4.3 Mb	22 751	[60]
斑点雀鳝 <i>Lepisosteus oculatus</i>	945 Mb	68.3 kb	6.9 Mb	18 328	[61]
亚洲鲈 <i>Lates calcarifer</i>	668.5 Mb	1 Mb	25 Mb	22 184	[62]
金线鲃 <i>Sinocyclocheilus grahami</i>	1.75 Gb	29 kb	1.3 Mb	42 109	[71]
中华绒螯蟹 <i>Eriocheir sinensis</i>	1.12 Gb	6.02 kb	224 kb	14 436	[80]
鲩 <i>Müichthys miiuy</i>	619.3 Mb	73.3 kb	1.15 Mb	21 026	[86]
海马 <i>Hippocampus comes</i>	501.6 Mb	34.7 kb	1.8 Mb	23 458	[72]
牙鲆 <i>Paralichthys olivaceus</i>	546 Mb	30.5 kb	3.9 Mb	21 787	[69]
雅罗鱼 <i>Leuciscus waleckii</i>	752.3 Mb	37.3 kb	447.7 kb	23 560	[70]
团头鲂 <i>Megalobrama amblycephala</i>	1 116 Mb	49.4 kb	838.7 Mb	23 696	[77]
乌鳢 <i>Channa argu</i>	615.3 Mb	81.4 kb	4.5 Mb	19 877	[78]
虾夷扇贝 <i>Patinopecten yessoensis</i>	988 Mb	38 kb	804 kb	26 415	[74]
仿刺参 <i>Apostichopus japonicus</i>	805 Mb	190 kb	486 kb	30 350	[81]

· 续表2 ·

物种 species	基因组大小 genome size	contig N50	scaffold N50	编码基因数量 coding gene no.	参考文献 reference
黄颡鱼 <i>Pelteobagrus fulvidraco</i>	732.8 Mb	1.1 Mb	25.8 Mb	24 552	[75]
花鲈 <i>Lateolabrax maculatus</i>	648 Mb	31 kb	1.04 Mb	22 015	[76]
黑斑原鲈 <i>Glyptosternon maculatum</i>	662.34 Mb	993.67 kb	20.9 Mb	22 066	[87]

MHC I 基因组家簇得到了扩张。Berthelot等^[53]报道了虹鳟的基因组精细图谱, 发现四倍化的虹鳟基因组具有缓慢二倍化的特征, 这与先前发现的全基因组复制后快速二倍化具有不同的机制和特点。David等^[54]同时发表5个鲷科鱼的全基因组测序精细图谱, 从全基因组水平阐述了鲷科鱼类的物种快速分化机制。Liu等^[55]报道了斑点叉尾鲷的全基因组序列和精细图谱, 阐明了斑点叉尾鲷由于缺乏分泌钙结合磷蛋白导致鳞片缺失的进化机制。Lien等^[56]报道了大西洋鲑的全基因组精细图谱, 发现新功能化的重复基因在数量上要远多于亚功能化的重复基因, 并阐述了四倍化起源后二倍化进程中重要的基因组进化事件, 即大量的基因组重复, 及转座子介导的重复序列扩张爆发性事件。此外, 还发表了罗非鱼^[57]、欧洲鲈(*Dicentrarchus labrax*)^[58]、蓝鳍金枪鱼(*Thunnus orientalis*)^[59]、大菱鲆(*Scophthalmus maximus*)^[60]、斑点雀鲷(*Lepisosteus oculatus*)^[61]、亚洲鲈(*Lates calcarifer*)^[62]等基因组精细图谱。

国内学者在水产生物基因组研究上也获得了很多高水平的成果, 如Zhang等^[63]报道了长牡蛎的全基因组精细图, 发现一系列与牡蛎抗逆能力相关的基因发生了明显扩张, 这可能是牡蛎适应潮间带极端环境的主要分子基础, 揭示了在逆境适应中发挥重要作用的贝壳的复杂形成机制。Chen等^[64]完成了半滑舌鳎全基因组精细图谱绘制, 发现*dmrt1*基因是Z染色体连锁、雄性特异表达、精巢发育必不可少的关键基因, 表现出性别决定基因的特性, 揭示了ZW性染色体起源和进化适应的分子机制。Xu等^[65]完成了鲤基因组精细图谱, 证实了鲤基因组的四倍化特征及其独特的全基因组复制事件, 首次绘制了全球鲤单碱基分辨率的基因组遗传变异图谱。此外, 识别和定位了松浦镜鲤鳞被缺失、荷包红鲤鲜红体色等性状相关的遗传位点、功能基因和调控通路。大黄鱼已有两个版本的基因组结果发表^[66-67], 阐述了先天免疫机制和压力适应

性机制。Wang等^[68]完成了草鱼的全基因组精细图谱, 并发现了有关鱼类草食性的机制。Shao等^[69]发表了牙鲆全基因组序列图谱, 证明了维甲酸对于建立不对称色素沉着以及与甲状腺激素的交互调节眼睛迁移的机制。Xu等^[70]完成了雅罗鱼(*Leuciscus waleckii*)的全基因组图谱, 并阐述了雅罗鱼对盐碱的适应性机制。Yang等^[71]报道了金线鲃属(*Sinocyclocheilus*)3个物种的基因组, 阐述了洞穴适应性进化机制。Lin等^[72]报道了海马(*Hippocampus comes*)基因组的测序结果, 发现了海马大量丢失的基因与性状功能减弱与有关, 如牙齿的退化、嗅觉减弱、腹鳍消失等。Ye等^[73]发表了海带(*Saccharina japonica*)基因组序列图谱, 发现褐藻胶合成关键酶甘露糖醛酸C-5差向异构酶及卤素代谢关键酶卤素过氧化物酶通过基因大量复制实现其功能分化。Wang等^[74]报道了虾夷扇贝(*Patinopecten yessoensis*)全基因组图谱绘制结果, 通过对扇贝基因组深度解析, 揭示扇贝呈现众多原始动物祖先基因组特征。

随着技术发展, 利用三代测序、光学图谱和HiC等新技术可以获得高可信度的染色体水平的全基因组精细图谱, 如Gong等^[75]采用PacBio三代测序技术和HiC辅助装配技术, 获得了黄颡鱼染色体水平的基因组, 对解析性别决定机制起到重要作用。此外, 我国重要养殖品种, 如, 花鲈(*Lateolabrax maculatus*)^[76]、团头鲂(*Megalobrama amblycephala*)^[77]、乌鳢(*Channa argu*)^[78]、菊黄东方鲀(*Takifugu flavidus*)^[79]、中华绒螯蟹(*Eriocheir sinensis*)^[80]、仿刺参(*Apostichopus japonicus*)^[81]等也完成了精细图谱绘制。

3.2 基因组遗传变异鉴定

基因组遗传变异包括单核苷酸多态性位点(SNP)、小片段插入和缺失位点(indel), 以及颠倒(inversion)、转座(translocation)、拷贝数变异(CNV)等结构性变异(structural variation), SNP和小片段indel可采用全基因组重测序或者简化基因组

测序(如RAD、2b-RAD、SLAF等)来鉴定,而结构性变异一般采用基因组重测序鉴定。

全基因组重测序是对已知基因组序列的物种的家系或群体样本进行基因组测序,可筛查出大量的单核苷酸多态性位点(SNP)、插入和缺失位点(indel)等变异信息,在此基础上可开展个体或群体的遗传差异性分析,如遗传图谱构建、性状定位和种质鉴定等研究。如, Xu等^[65]在获得鲤基因组序列图谱基础上,对10个群体33尾鲤进行重测序,并据此分析鲤群体间的遗传差异并得出了不同地理群的来源。Xu等^[70]对雅罗鱼的10个淡水个体和18个耐碱水个体进行全基因组重测序,鉴定到760万个差异的SNP,从中鉴定到一大批与离子稳态、酸碱调节、活性氧消除和尿素排泄等环境适应变化有关的基因。Lamichhaney等^[88]对大西洋和波罗的海不同区域的大西洋鲱(*Clupea harengus*)群体进行重测序,并将序列与外显子组装对齐,鉴定了440 817个SNP,其中有数千个SNP显示出显著的差异,并判断大西洋鲱不同群体的遗传分化主要来自各大地理群的地理适应,这也是没有参考基因组的群体全基因组测序研究报道。Hohenlohe等^[89]利用RAD测序技术对多个虹鳟群体进行测序,根据测序结果和先前鉴定的基因位点发现了鳟杂交群体具有来自候选群体的适应性超入侵的等位基因。

遗传变异检测的另一种方式是采用SNP芯片进行基因型分析。2014年发布了3个物种的高密度SNP芯片,分别为大西洋鲑^[90]、斑点叉尾鲷^[91]和鲤^[92],其中斑点叉尾鲷和鲤是同一张芯片,2个物种各有250 K SNP。但由于二代测序成本的快速下降,SNP芯片的应用也遇到了很大挑战。与重测序相比,SNP芯片的优点是基因型固定、获得有效位点的比例高、数据易处理等,缺点是灵活性不够,成本也相对高,需要有前期的研究基础,至少几十万个SNP的鉴定也是一个很庞大的工作。因此,SNP芯片在水产领域的应用比预想的速度慢,价值也不如预期,但由于每个物种与性状相关的SNP积累增加,基于研究积累的小芯片在群体基因组和品种选育等研究中应该很有前途。

3.3 表观遗传分析

高通量测序技术的发展也促进了生命科学研

究中新学科的形成和发展,如表观遗传学。Dupont等^[93]将表观遗传学定义为“对可遗传的有丝分裂或减数分裂过程中基因功能的变化且不会引起DNA序列改变的现象进行研究的学科”。采用甲基化测序、ChIP-Seq测序、DNase-Seq、ATAC-Seq等高通量测序技术,使科学家能大规模的筛选性状形成过程中基因组DNA的修饰、染色质可及性(chromatin accessibility)以及染色体构象信息。DNA甲基化是表观遗传的主要方面,是不改变基因序列的前提下改变基因表达的修饰,这种现象与基因型差异一起,结合miRNA等对基因表达、性状的最终形成进行调控。总之,在孟德尔遗传现象之外,近年来发现的表观遗传和miRNA等对基因表达的调控作用使性状与基因间的关系更为复杂,一些核遗传和细胞遗传都解释不了的现象很可能与表观遗传有关,譬如鱼类的性别决定、分化和性逆转。Jabbari等^[94]发现,变温的鱼类和两栖类动物较恒温的鸟类和哺乳类动物有更高的DNA甲基化水平;Shao等^[95]发现表观遗传调控在半滑舌鳎的雌性转变为生理雄性的过程中起着多重的关键作用。

近二十多年,表观遗传领域的进展与基因表达调控领域的进展基本上同步,2017年Cell杂志以Landmark Cell Reviews: Transcription and Epigenetics为题,收集34篇顶级学者的综述来回顾近几十年在基因表达和表观遗传领域的新进展,表观遗传现象已扩展很多,修饰包括DNA、RNA和蛋白,表观遗传不仅仅对同一细胞不同表型的现象起作用,而是对所有遗传现象都有影响,当然表观遗传引起重视的是那些经典遗传解释不清楚的现象,如植物的花色^[96],观赏鱼类的图案,半滑舌鳎等鱼类性别等,有关表观遗传进展可参考丁勇等^[97]、杨震飞等^[98]的综述和上面提到的34篇文献。

表观遗传研究结果目前在水产遗传育种和病害防治等研究领域还没有起到明显作用,但可预期的是,一些长期难以解释的遗传现象将通过表观遗传的研究获得合理的解释并最终进行有效的遗传控制,前提是要鉴定到可稳定遗传的位点。

3.4 基因表达分析

根据中心法则,mRNA是基因的转录产物,携带基因的信息指导蛋白质的合成,是基因与

蛋白质间的桥梁。由于不同个体的同一基因在同一发育阶段表达出的蛋白质有结构和数量的差别,而这些差别是物种中个体间差异的核心。因此,研究mRNA是阐明基因与蛋白质关系的关键,也是性状研究的关键。

转录组测序即RNA-seq,是用第二代或第三代测序技术对转录的RNA进行测序以探索基因的转录情况。这项技术一经出现就被同行誉为是功能基因组研究技术的革命性进步^[99-100],也快速地应用到生物学的各领域。RNA-Seq Blog网站(<https://www.rna-seqblog.com>)上公布了2017年前每年文章数量,2012年超过500篇,以后每年增加500多篇,2017年超过3 000篇,第一篇文章^[99]被引用已达到9 330多次,第一篇综述^[100]被引用也达到8 552次以上,反映了这项技术应用速度非常快。在水产方面也得到了广泛的应用,如Robledo等^[101]利用RNA-seq技术对大西洋鲑皮肤感染海虱基因应答进行了研究;Xu等^[102]利用RNA-seq技术探索了我国内蒙地区碱水湖泊(达里湖)雅罗鱼的耐碱机制和适应进化的遗传基础。转录组测序技术已经应用到生长发育、疾病发生等遗传机制及相关的遗传基础的探索之中,从分子生物角度看这项技术有如下用途:①在全基因组范围筛查差异表达的基因,这是转录组测序的主要任务,多数水产性状相关基因的鉴定基本上也是通过这个方法获得;②批量鉴定转录组SNP、SSR和其他indel等变异位点;③获得全长cDNA序列,全长cDNA序列一直是完整、全面注释基因的基础;④表观转录组分析^[103],研究表明腺苷的甲基化属转录后修饰且广泛存在于哺乳动物中,且此种修饰富集在终止密码子附近,也与3'UTR内的微小RNA结合位点有关联。随着此技术的推广,将有更多的创新技术被发明出来。

必须指出,获得差异表达基因只是研究的开始,要真正获得可信度高的决定性状的关键基因和阐明性状的形成机制,还有很多深入的研究要做,下面简介几个深入研究的策略^[104],①通路分析,鉴定出候选基因所在的代谢通路,或基因表达通路是最终确定基因在性状形成中贡献的依据。②基因共表达网络分析,如果有2个以上样本的表达谱,通过聚类得到表达特征相似的成对基因,计算每对基因的共表达的分值,确定显著性阈值,如果选择的基因足

够多,样本针对性强,可以获得包括多种生物学过程的共表达网络,具有相似表达谱的不同基因很可能具有相近的功能,由此,通过基因共表达网络可以帮助我们更全面地认识同一性状形成的复杂的遗传机制,但同样功能的基因可能有不同的表达谱,不同功能的基因也可能有相似的表达谱,这些也同样表现出性状形成的生物学过程是极其复杂的。③系统生物学分析,系统生物学提出多年,但只有基因芯片和二代测序等高通量检测技术出现后才有了实际性进展^[105],核心内容就是通过生物信息分析技术将各种组学包括基因组、转录组、蛋白质组、代谢组、表观组等的检测结果进行系统研究,分析各个层面对性状形成或疾病发生所起的作用,准确鉴定基因和通路所起到的关键作用,从而筛选出可以用于品种选择的基因或疾病治疗的候选靶点。

3.5 遗传连锁图谱和QTL定位

随着DNA测序技术的发展及其在水产生物中的应用,通过二代测序技术许多水产生物获得了大量多态性标记,为高密度遗传连锁图谱构建奠定了基础。2010年以前SSR是构建遗传图谱最常用的标记类型,自2011年开始,SNP取代了SSR标记成为主流,到2015年,绝大多数图谱采用SNP构建。近5年来,超过30种水产生物构建了以SNP为主的高密度遗传连锁图谱,其中大西洋鲑^[106]图谱密度最高,包含96 396个SNP,雌、雄图谱平均标记间隔分别0.05和0.07 cM。在高密度图谱构建基础上,同时也对重要性状进行了精细QTL定位。如,Peng等^[107]构建了黄河鲤(*Cyprinus carpio haematopterus*)高密度SNP图谱,含有28 194个SNP,平均标记间隔为0.38 cM,并获得了22个生长相关QTL和7个性别决定QTL;Liu等^[108]构建了鲫(*Carassius auratus*)高密度SNP图谱,含有8 487个SNP,平均标记间隔0.44 cM,定位到8个体重QTL。Liu等^[109]构建了亚洲鲈高密度遗传图谱,包含2 852个SNP和148个SSR标记,平均标记间隔为1.27 cM,并定位到4个病毒性神经坏死病相关QTL,确定*Pcdhac2*为候选基因,进一步分析发现该基因内含子上6 bp缺失与该病抗性显著相关。Shao等^[110]构建了牙鲆高密度遗传图谱,含12 712个SNP,平均标记间隔为0.47 cM,定位到4个与鳃弧菌病抗性相关QTL,利用基因

组装信息在QTL区间筛选出12个免疫相关候选基因,其中4个基因(*tap1*、*satb1*、*cd40*和*cd69*)与鳃弧菌病抗性显著相关。Yu等^[111]构建了凡纳滨对虾高密度遗传图谱,包含6 146个SNP,平均标记间隔为0.7 cM,定位到了11个体长和7个体质量QTL。Jiao等^[112]构建了栉孔扇贝(*Chlamys farreri*)的高密度遗传图谱,包含3 806个SNP,平均标记间隔为0.41 cM,定位到5个生长性状主成分因子1(PC1)的QTL区间2个,分别位于LG1和LG3,同时发现性别相关QTL也位于LG1,置信区间仅为0.37 cM,区间里27个标记与性别显著相关,其中1个标记位于*ZNFXI*基因。总之,基于高密度图谱的QTL研究,由于标记多,使QTL定位更加精细,较小的QTL区间,获得区间内的候选基因的数目少,更有利于候选基因的筛选与鉴定。本文整理出了近5年重要水产养殖生物高密度图谱及QTL定位情况(表3),更详细的水产生物QTL定位研究进展参见Yue等^[113]和张晓峰等^[114]综述。

3.6 性状关联分析

随着二代测序技术的应用,许多动植物的全基因组测序已经完成,全基因组关联分析(genome-wide association study, GWAS)成为了研究复杂性状的重要方法,并在人类和动植物复杂性状遗传研究中已取得诸多成果。在全基因组测序和海量SNP标记鉴定的基础上,GWAS在水产生物多种复杂性状的遗传研究中也得到了应用。大西洋鲑^[90, 131]、斑点叉尾鲟^[91]、鲤^[92]、虹鳟^[132]、牡蛎^[133-134]等重要水产养殖生物高通量SNP芯片已经被开发,基于这些SNP芯片,GWAS研究在这些鱼类的生长、抗病、品质等性状中逐渐被开展。目前,GWAS研究主要集中在鱼类(表4)。大西洋鲑中,Sodeland等^[135]对鱼肉品质性状进行了GWAS研究,将与鱼片脂肪含量显著相关的SNPs定位于9和10号染色体,鱼片质地显著相关的SNPs定位于3和11号染色体;Tasi等^[136]对早期生长性状进行了GWAS研究,1个与体长相关的SNP达到了染色体水平显著性;Gonen等^[137]开展了鲑鱼胰腺病性状的GWAS和QTL分析,鉴定到6个该病抗性QTL,其中1个位于3号染色体;Correa等^[138]对大西洋鲑海虱病进行GWAS分析发现,位于21号染色体上的1个SNP达到染色体水平显著性,且该位点位于Collagen alpha-1基

因内含子中。Campbell等^[139]对冷水性细菌病(CWD)和传染性造血坏死病毒病(IHNV) GWAS研究发现,12个SNP与CWD抗性显著相关,19个SNP与IHNV抗性显著相关;Palti等^[140]进行了CWD抗性的GWAS分析发现,有6个显著位点与之前QTL结果一致,2个新位点位于11和14号染色体;Gonzalez-Pena等^[141]开展了虹鳟鱼片产量、体质量等性状GWAS研究。斑点叉尾鲟中,利用SNP芯片开展了多个性状的GWAS研究,如,Geng等^[142]对斑点叉尾鲟柱形病毒抗性进行了GWAS研究,在LG7上定位到一个620 kb的QTL区间,进一步发现该区间包含10个基因,其中有5个免疫相关基因,可能与柱形病毒抗性相关;Jin等^[143]开展了热应激性状GWAS研究,定位到3个SNP与性状显著相关,其中1个SNP位于LG14,2个SNP位于LG16,与斑点叉尾鲟^[91]基因组序列信息比对,在显著相关区域发现14个基因与热应激相关;Zhong等^[144]进行了斑点叉尾鲟^[91]和蓝斑点叉尾鲟杂交子代低氧耐受性状的GWAS分析,在LG2和LG23上存在性状显著相关QTL,区间内基因位于VEGF、MAPK、mTOR、PI3K-Akt等信号通路。此外,在体型^[145]、生长^[146]等性状也有相关GWAS研究报道。在鲤中,Zheng等^[147]利用SNP芯片开展了鲤品质性状GWAS研究,获得18个SNP与肌肉脂肪含量、腹部脂肪重和腹脂率性状显著相关,与鲤基因组注释信息比对发现10个候选基因,进一步对5个脂肪含量性状的候选基因进行验证,其中4个基因(*ankrd10a*、*tanc2*、*jx1*和*chka*)表达量与脂肪含量呈现显著相关;Chen等^[148]对鲤头尺寸相关性状进行了GWAS分析。

GWAS具有高通量(即在全基因组范围有效检测性状与基因位点的关联性)、精度高和花费时间少等显著优点。但诸多限制性因素^[149],如,GWAS的假阳性或假阴性高,多群体、大样本的重复验证研究才是提高检验效能,确保真正性状关联位点的关键。

4 水产生物基因组研究发展趋势

4.1 构建更为准确的全基因组精细图谱

构建准确的全基因组精细图谱是下个阶段水产生物基因组研究的重要任务,因为已发表的精细图谱基本上是覆盖度比较低的,如鲤基因组的精细图谱仅有56%左右的序列整合在连

表 3 近5年水产生物高密度遗传连锁图谱和QTL定位一览表

Tab. 3 The list of high-density genetic linkage map and QTL mapping of aquaculture species in nearly 5 years

物种 species	标记类型 type of marker	标记数量 no. of markers	图谱大小/cM map length/cM	连锁群数 no. of linkage groups	平均标记间隔/cM average marker interval/cM	QTL性状 QTL trait	参考文献 reference
大西洋鲑 <i>Salmo salar</i>	SNP	96 396	4 769(♂)/7 153(♀)	29	0.05(♂)/0.07(♀)		[106]
	SNP	6 458	2 190.3	29	0.34		[115]
鲤 <i>Cyprinus carpio</i>	SNP	28 194	10 595.94	50	0.38	生长、性别	[107]
	SSR、SNP	8 115	4 586.56	50	0.57	生长、性别	[116]
鲫 <i>Carassius auratus</i>	SNP	8 521	5 252	50	0.62	生长	[117]
	SNP	8 487	3 762.88	50	0.44	生长	[108]
鳊 <i>Hypophthalmichthys nobilis</i>	SNP	3 121	2 341.27	24	0.75	生长	[118]
鳊 <i>Siniperca chuatsi</i>	SNP	3 283	1 972.01	24	0.61	生长、性别	[119]
团头鲂 <i>Megalobrama amblycephala</i>	SNP	14 648	3 258.38	24	0.57	生长、性腺发育、性别	[120]
杂交斑点叉尾鲴 <i>Ictalurus punctatus</i> × <i>Ictalurus furcatus</i>	SNP	26 238	3 240	29	0.25		[121]
大黄鱼 <i>Larimichthys crocea</i>	SNP	10 150	5 451	24	0.54		[122]
亚洲鲈 <i>Lates calcarifer</i>	SNP	3 321	1 577.67	24	0.52	生长	[123]
	SSR、SNP	3 000	2 957.79	24	1.27	病毒性神经坏死病 鳃弧菌病抗性	[109]
牙鲆 <i>Paralichthys olivaceus</i>	SNP	12 712	3 497.29	24	0.47		[110]
大菱鲆 <i>Scophthalmus maximus</i>	SNP	6 647	2 262	22	0.395	生长、性别	[124]
欧洲鲈 <i>Dicentrarchus labrax</i>	SNP	6 706	4 816	24	0.72	性别	[125]
凡纳滨对虾 <i>Litopenaeus vannamei</i>	SNP	6 146	4 271.43	44	0.7	生长	[111]
中华绒螯蟹 <i>Eriocheir sinensis</i>	SNP、SSR	18 309	14 894.9	73	0.81	生长、性别	[126]
杂交牡蛎 <i>Crassostrea gigas</i> × <i>Crassostrea angulata</i>	SNP	3 367	1 084.3	10	0.8	生长	[127]
栉孔扇贝 <i>Chlamys farreri</i>	SNP	3 806	1 543.4	19	0.41	生长、性别	[112]
缢蛏 <i>Sinonovacula constricta</i>	SNP	7 516	2 383.85	19	0.32	生长	[128]
海带 <i>Saccharina japonica</i>	SNP	7 627	5 232.42	31	0.69	产量	[129]
海胆 <i>Strongylocentrotus intermedius</i>	SNP	4 378	1 907.21	21	0.44	生长、性腺品质	[130]

锁图上, 有了三代测序技术和光学图谱等新技术, 可以使已建立的精细图的覆盖度提高; 另

外, 基因注释质量也需要提高, 毕竟准确的基因注释结果对鉴定经济性状的决定基因和基因

表 4 鱼类全基因组关联分析研究情况一览表

Tab. 4 Summary of GWAS applications in fish

物种 species	性状 trait	SNP数量 no. of SNPs	样本数 no. of samples	参考文献 reference
大西洋鲑 <i>Salmo salar</i>	鱼片质地、脂肪含量 fillet texture, fat content	250,000	770	[135]
	生长 frowth	122,000	622	[136]
	生长、早熟 growth, early maturation	6,500	480	[150]
	胰腺疾病 pancreas disease	6,000	1,182	[137]
虹鳟 <i>Oncorhynchus mykiss</i>	海虱抗性 sea lice resistance	50,000	2,628	[138]
	冷水性细菌病、传染性造血坏死病毒病 bacterial cold water disease (CWD), infectious hematopoietic necrosis virus (IHNV)	4,600	312	[139]
	冷水性细菌病 bacterial cold water disease (CWD)	5,000	240	[140]
斑点叉尾鲷 <i>Ictalurus punctatus</i>	鱼片产量、生长 fillet yield, growth	38,000	1,447	[141]
	柱状病毒抗性 columnaris disease resistance	250,000	340	[142]
	热刺激 heat stress	250,000	315	[143]
	低氧耐受 low oxygen tolerance	250,000	630	[144]
	头骨形态 Skull morphology	250,000	556	[145]
	生长 Growth	250,000	556	[146]
鲤 <i>Cyprinus caprio</i>	肠败血症抗性 enteric septicemia resistance	250,000	192	[151]
	脂肪含量相关性状 fat related traits	250,000	220	[147]
银大麻哈鱼 <i>Oncorhynchus kisutch</i>	头尺寸 head size	250,000	474	[148]
	脂肪含量相关性状 fat related traits	250,000	220	[147]
欧洲鲈 <i>Dicentrarchus labrax</i>	鲑鱼立克次体综合征 salmon rickettsial syndrome (SRS)	4,100	764	[152]
亚洲鲈 <i>Lates calcarifer</i>	性别决定 sex determination	6,700	175	[153]
	病毒性神经坏死病 viral nervous necrosis (VNN) disease	44,500	986	[154]

所在信号通路是必须的，但准确的鉴定功能基因是一件困难的事情，人类基因总数虽然在基因组完成时确定约有24 000个编码蛋白基因，有2个团队继续开展人类基因数量的研究，但直至目前人类基因的准确数目仍有争论^[155]，新的编码基因、RNA基因和总的转录本是水产生物的重要参考基因。

4.2 质量性状决定基因及信号通路的鉴定

质量性状(如性别、抗性等)的决定基因相对于数量性状基因是比较容易鉴定的，也可以通过比较基因组方法获得，虽然质量性状在物种间也有差异，如染色体上的位置不同、基因和转录本的数量不同等，但基本序列应该是相似的，基因所在的信号通路和代谢通路也应该基本相同，这些是开展比较基因组研究的基础，同理，候选基因法在基因和基因组资源越来越

丰富的条件下会更有使用价值，尤其是人、模式生物和畜禽等有很多相同的性状，有很多的研究结果可以借鉴。

4.3 数量性状连锁基因及信号通路在群体间的变异

典型的数量性状其QTL和GWAS结果有很大的差异，而且两者分别在不同群体间有很大变动，这既是研究的难点也是生物学引人入胜之处。虽然QTL和GWAS表现出极其复杂的特点，但其反应的是群体的遗传差异，这些遗传差异是通过基因的表达来影响性状的。因此，可以通过基因表达数量性状基因座(eQTL)和主效基因附近的功能基因所在信号通路与QTL和GWAS建立联系^[156]，探索形成群体间QTL和GWAS相同和不同所代表的遗传基础及与环境互作的生理生化机制，从而在基因组层面理解数量性状相关

的遗传学和育种学问题。同时, 基因组水平的种质鉴定研究也将是下一个阶段的重要研究内容。

4.4 多种组学结合阐明经济性状的形成机制

性状与基因的关系极其复杂, 多种组学结合起来探索性状的形成机制是近年来的发展趋势之一, 人及模式生物的多种组学技术平台已基本建立起来^[157], 近年来微生物组学研究进展迅速, 研究平台基本成型^[158]。对于水产生物来说多种组学结合起来研究目前还有困难, 但2个以上的组学数据整合起来分析是可能的, 也可以更全面地了解性状形成的生理机制和遗传基础, 获得更可信的结果。

参考文献:

- [1] Maxam A M, Gilbert W. A new method for sequencing DNA[J]. Proceedings of the National Academy of Sciences of the United States of America, 1977, 74(2): 560-564.
- [2] Sanger F, Nicklen S, Coulson A R. DNA sequencing with chain-terminating inhibitors[J]. Proceedings of the National Academy of Sciences of the United States of America, 1977, 74(12): 5463-5467.
- [3] Martin G R. The roles of FGFs in the early development of vertebrate limbs[J]. Genes & Development, 1998, 12(11): 1571-1586.
- [4] Aparicio S, Chapman J, Stupka E, et al. Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*[J]. Science, 2002, 297(5585): 1301-1310.
- [5] Metzker M L. Sequencing technologies-the next generation[J]. Nature Reviews Genetics, 2010, 11(1): 31-46.
- [6] Pushkarev D, Neff N F, Quake S R. Single-molecule sequencing of an individual human genome[J]. Nature Biotechnology, 2009, 27(9): 847-850.
- [7] Li R Q, Fan W, Tian G, et al. The sequence and *de novo* assembly of the giant panda genome[J]. Nature, 2010, 463(7279): 311-317.
- [8] Treffer R, Deckert V. Recent advances in single-molecule sequencing[J]. Current Opinion in Biotechnology, 2010, 21(1): 4-11.
- [9] Kulski J K. Next-generation sequencing-an overview of the history, tools, and “omic” applications[M]//Kulski J K. Next Generation Sequencing-Advances, Applications and Challenges. InTech, 2016: 3-60.
- [10] Skinner D M, Beattie W G, Blattne F R, et al. Repeat sequence of a hermit crab satellite deoxyribonucleic acid is (-T-A-G-G)_n. (-A-T-C-C)_n[J]. Biochemistry, 1974, 13(19): 3930-3937.
- [11] Kocher T D, Lee W J, Sobolewska H, et al. A genetic linkage map of a cichlid fish, the tilapia (*Oreochromis niloticus*)[J]. Genetics, 1998, 148(3): 1225-1232.
- [12] Young W P, Wheeler P A, Coryell V H, et al. A detailed linkage map of rainbow trout produced using doubled haploids[J]. Genetics, 1998, 148(2): 839-850.
- [13] 魏东旺, 楼允东, 孙效文, 等. 鲤鱼微卫星分子标记的筛选[J]. 动物学研究, 2001, 22(3): 238-241.
- Wei D W, Lou Y D, Sun X W, et al. Isolation of microsatellite markers in the common carp (*Cyprinus carpio*)[J]. Zoological Research, 2001, 22(3): 238-241 (in Chinese).
- [14] 徐鹏, 周岭华, 相建海. 中国对虾微卫星DNA的筛选[J]. 海洋与湖沼, 2001, 32(3): 255-259.
- Xu P, Zhou L H, Xiang F H. Isolating microsatellite DNA of Chinese shrimp *Penaeus chinensis*[J]. Oceanologia et Limnologia Sinica, 2001, 32(3): 255-259.
- [15] Serapion J, Kucuktas H, Feng J N, et al. Bioinformatic mining of type I microsatellites from expressed sequence tags of channel catfish (*Ictalurus punctatus*)[J]. Marine Biotechnology, 2004, 6(4): 364-377.
- [16] 万海伟, 杜立新. 表达序列标签(EST)在基因组学中的应用[J]. 生物技术通报, 2004(1): 35-38, 34.
- Wan H W, Du L X. Application of EST in the study of genomics[J]. Biotechnology Bulletin, 2004(1): 35-38, 34 (in Chinese).
- [17] 张晓峰, 杨晶, 孙效文. 基于EST序列的鲤鱼生长相关 SNP发掘[J]. 水产学杂志, 2009, 22(4): 1-7.
- Zhang X F, Yang J, Sun X W. Exploitation of SNPs related to growth traits of common carp (*Cyprinus carpio* L.) based on EST sequences[J]. Chinese Journal of Fisheries, 2009, 22(4): 1-7 (in Chinese).
- [18] Zhang Y A, Zou J, Chang C I, et al. Discovery and characterization of two types of liver-expressed antimicrobial peptide 2(LEAP-2) genes in rainbow trout[J]. Veterinary Immunology and Immunopathology, 2004, 101(3-4): 259-269.

- [19] von Schalburg K R, Rise M L, Cooper G A, *et al.* Fish and chips: various methodologies demonstrate utility of a 16,006-gene salmonid microarray[J]. *BMC Genomics*, 2005, 6: 126.
- [20] Naruse K, Tanaka M, Mita K, *et al.* A medaka gene map: the trace of ancestral vertebrate proto-chromosomes revealed by comparative gene mapping[J]. *Genome Research*, 2004, 14(5): 820-828.
- [21] Ng S H S, Artieri C G, Bosdet I E, *et al.* A physical map of the genome of Atlantic salmon, *Salmo salar*[J]. *Genomics*, 2005, 86(4): 396-404.
- [22] 汤飞宇, 张天真. 重叠群物理图谱的构建及其应用[J]. *基因组学与应用生物学*, 2009, 28(1): 195-201.
Tang F Y, Zhang T Z. Construction and application of large-insert clones-based physical map[J]. *Genomics and Applied Biology*, 2009, 28(1): 195-201 (in Chinese).
- [23] Xu P, Wang S L, Liu L, *et al.* A BAC-based physical map of the channel catfish genome[J]. *Genomics*, 2007, 90(3): 380-388.
- [24] Palti Y, Luo M C, Hu Y Q, *et al.* A first generation BAC-based physical map of the rainbow trout genome[J]. *BMC Genomics*, 2009, 10: 462.
- [25] Xu P, Wang J, Wang J T, *et al.* Generation of the first BAC-based physical map of the common carp genome[J]. *BMC Genomics*, 2011, 12: 537.
- [26] Lewin H A, Larkin D M, Pontius J, *et al.* Every genome sequence needs a good map[J]. *Genome Research*, 2009, 19(11): 1925-1928.
- [27] 孙效文, 梁利群. 鲤鱼的遗传连锁图谱(初报)[J]. *中国水产科学*, 2000, 7(1): 1-5.
Sun X W, Liang L Q. A genetic linkage map of common carp[J]. *Journal of Fishery Sciences of China*, 2000, 7(1): 1-6 (in Chinese).
- [28] Sun X W, Liang L Q. A genetic linkage map of common carp (*Cyprinus carpio* L.) and mapping of a locus associated with cold tolerance[J]. *Aquaculture*, 2004, 238(1-4): 165-172.
- [29] Hubert S, Hedgecock D. Linkage maps of microsatellite DNA markers for the pacific oyster *Crassostrea gigas*[J]. *Genetics*, 2004, 168(1): 351-362.
- [30] Nichols K M, Young W P, Danzmann R G, *et al.* A consolidated linkage map for rainbow trout (*Oncorhynchus mykiss*)[J]. *Animal Genetics*, 2003, 34(2): 102-115.
- [31] Lee B Y, Lee W J, Streeleman J T, *et al.* A second-generation genetic linkage map of tilapia (*Oreochromis spp.*)[J]. *Genetics*, 2005, 170(1): 237-244.
- [32] Moen T, Hayes B, Baranski M, *et al.* A linkage map of the Atlantic salmon (*Salmo salar*) based on EST-derived SNP markers[J]. *BMC Genomics*, 2008, 9: 223.
- [33] Xia J H, Liu F, Zhu Z Y, *et al.* A consensus linkage map of the grass carp (*Ctenopharyngodon idella*) based on microsatellites and SNPs[J]. *BMC Genomics*, 2010, 11: 135.
- [34] Jackson T R, Ferguson M M, Danzmann R G, *et al.* Identification of two QTL influencing upper temperature tolerance in three rainbow trout (*Oncorhynchus mykiss*) half-sib families[J]. *Heredity*, 1998, 80(2): 143-151.
- [35] Yue G H. Recent advances of genome mapping and marker-assisted selection in aquaculture[J]. *Fish and Fisheries*, 2014, 15(3): 376-396.
- [36] Laghari M Y, Lashari P, Zhang X F, *et al.* QTL mapping for economically important traits of Common carp (*Cyprinus carpio* L.)[J]. *Journal of Applied Genetics*, 2015, 56(1): 65-75.
- [37] Derayat A, Houston R D, Guy D R, *et al.* Mapping QTL affecting body lipid percentage in Atlantic salmon (*Salmo salar*)[J]. *Aquaculture*, 2007, 272(S1): S250-S251.
- [38] Baranski M, Moen T, Våge D I. Mapping of quantitative trait loci for flesh colour and growth traits in Atlantic salmon (*Salmo salar*)[J]. *Genetics Selection Evolution*, 2010, 42: 17.
- [39] Kuang Y Y, Zheng X H, Lv W H, *et al.* Mapping quantitative trait loci for flesh fat content in common carp (*Cyprinus carpio*)[J]. *Aquaculture*, 2014, 435: 100-105.
- [40] 李鸥, 曹顶臣, 张研, 等. 利用EST-SSR分子标记研究鲤的饲料转化率性状[J]. *水产学报*, 2009, 33(4): 624-631.
Li O, Cao D C, Zhang Y, *et al.* Studies on feed conversion ratio trait of common carp (*Cyprinus carpio* L.) using EST-SSR marker[J]. *Journal of Fisheries of China*, 2009, 33(4): 624-631 (in Chinese).
- [41] 张丽博, 张晓峰, 曹顶臣, 等. 利用SSR及EST标记对鲤鱼饲料转化率的QTL分析[J]. *农业生物技术学报*,

- 2010, 18(5): 963-967.
- Zhang L B, Zhang X F, Cao D C, *et al.* QTL analysis related to feed conversion efficiency in common carp (*Cyprinus carpio*) using SSR and EST markers[J]. Journal of Agricultural Biotechnology, 2010, 18(5): 963-967 (in Chinese).
- [42] 王宣朋, 张晓峰, 李文升, 等. 鲤饲料转化率性状的 QTL 定位及遗传效应分析[J]. 水生生物学报, 2012, 36(2): 177-196.
- Wang X P, Zhang X F, Li W S, *et al.* Mapping and genetic effect analysis on quantitative trait loci related to feed conversion ratio of common carp (*Cyprinus carpio* L.)[J]. Acta Hydrobiologica Sinica, 2012, 36(2): 177-196 (in Chinese).
- [43] Barroso R M, Wheeler P A, Lapatra S E, *et al.* QTL for IHNV resistance and growth identified in a rainbow (*Oncorhynchus mykiss*) \times yellowstone cutthroat (*Oncorhynchus clarki bouvieri*) trout cross[J]. Aquaculture, 2008, 277(3-4): 156-163.
- [44] Fuji K, Kobayashi K, Hasegawa O, *et al.* Identification of a single major genetic locus controlling the resistance to lymphocystis disease in Japanese flounder (*Paralichthys olivaceus*)[J]. Aquaculture, 2006, 254(1-4): 203-210.
- [45] Fuji K, Hasegawa O, Honda K, *et al.* Marker-assisted breeding of a lymphocystis disease-resistant Japanese flounder (*Paralichthys olivaceus*)[J]. Aquaculture, 2007, 272(1-4): 291-295.
- [46] Xu T J, Chen S L, Ji X S, *et al.* MHC polymorphism and disease resistance to *Vibrio anguillarum* in 12 selective Japanese flounder (*Paralichthys olivaceus*) families[J]. Fish & Shellfish Immunology, 2008, 25(3): 213-221.
- [47] Yu H Y, He Y, Wang X X, *et al.* Polymorphism in a serine protease inhibitor gene and its association with disease resistance in the eastern oyster (*Crassostrea virginica* Gmelin)[J]. Fish & Shellfish Immunology, 2011, 30(3): 757-762.
- [48] Kasahara M, Naruse K, Sasaki S, *et al.* The medaka draft genome and insights into vertebrate genome evolution[J]. Nature, 2007, 447(7145): 714-719.
- [49] Howe K, Clark M D, Torroja C F, *et al.* The zebrafish reference genome sequence and its relationship to the human genome[J]. Nature, 2013, 496(7446): 498-503.
- [50] Kettleborough R N W, Busch-Nentwich E M, Harvey S A, *et al.* A systematic genome-wide analysis of zebrafish protein-coding gene function[J]. Nature, 2013, 496(7446): 494-497.
- [51] Star B, Nederbragt A J, Jentoft S, *et al.* The genome sequence of Atlantic cod reveals a unique immune system[J]. Nature, 2011, 477(7363): 207-210.
- [52] Malmstrøm M, Matschiner M, Tørresen O K, *et al.* Evolution of the immune system influences speciation rates in teleost fishes[J]. Nature Genetics, 2016, 48(10): 1204-1210.
- [53] Berthelot C, Brunet F, Chalopin D, *et al.* The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates[J]. Nature Communications, 2014, 5: 3657.
- [54] Brawand D, Wagner E C, Li Y I, *et al.* The genomic substrate for adaptive radiation in African cichlid fish[J]. Nature, 2014, 513(7518): 375-381.
- [55] Liu Z J, Liu S K, Yao J, *et al.* The channel catfish genome sequence provides insights into the evolution of scale formation in teleosts[J]. Nature Communications, 2016, 7: 11757.
- [56] Lien S, Koop B F, Sandve S R, *et al.* The Atlantic salmon genome provides insights into rediploidization[J]. Nature, 2016, 533(7602): 200-205.
- [57] Brawand D, Wagner C E, Li Y I, *et al.* The genomic substrate for adaptive radiation in African cichlid fish[J]. Nature, 2014, 513(7518): 375-381.
- [58] Tine M, Kuhl H, Gagnaire P A, *et al.* European sea bass genome and its variation provide insights into adaptation to euryhalinity and speciation[J]. Nature Communications, 2014, 5: 5770.
- [59] Nakamura Y, Mori K, Saitoh K, *et al.* Evolutionary changes of multiple visual pigment genes in the complete genome of pacific bluefin tuna[J]. Proceedings of the National Academy of Sciences of the United States of America, 2013, 110(27): 11061-11066.
- [60] Figueras A, Robledo D, Corvelo A, *et al.* Whole genome sequencing of turbot (*Scophthalmus maximus*; Pleuronectiformes): a fish adapted to demersal life[J]. DNA Research, 2016, 23(3): 181-192.
- [61] Braasch I, Gehrke A R, Smith J J, *et al.* The spotted gar genome illuminates vertebrate evolution and facilitates

- human-teleost comparisons[J]. *Nature Genetics*, 2016, 48(4): 427-437.
- [62] Vij S, Kuhl H, Kuznetsova I S, *et al.* Chromosomal-level assembly of the Asian seabass genome using long sequence reads and multi-layered scaffolding[J]. *PLOS Genetics*, 2016, 12(4): e1005954.
- [63] Zhang G F, Fang X D, Guo X M, *et al.* The oyster genome reveals stress adaptation and complexity of shell formation[J]. *Nature*, 2012, 490(7418): 49-54.
- [64] 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 lifestyle[J]. *Nature Genetics*, 2014, 46(3): 253-260.
- [65] Xu P, Zhang X F, Wang X M, *et al.* Genome sequence and genetic diversity of the common carp, *Cyprinus carpio*[J]. *Nature Genetics*, 2014, 46(11): 1212-1219.
- [66] Wu C W, Zhang D, Kan M Y, *et al.* The draft genome of the large yellow croaker reveals well-developed innate immunity[J]. *Nature Communication*, 2014, 5: 5227.
- [67] Ao J Q, Mu Y N, Xiang L X, *et al.* Genome sequencing of the perciform fish *Larimichthys crocea* provides insights into molecular and genetic mechanisms of stress adaptation[J]. *PLoS Genetics*, 2015, 11(4): e1005118.
- [68] Wang Y P, Lu Y, Zhang Y, *et al.* The draft genome of the grass carp (*Ctenopharyngodon idellus*) provides insights into its evolution and vegetarian adaptation[J]. *Nature Genetics*, 2015, 47(6): 625-631.
- [69] Shao C W, Bao B L, Xie Z Y, *et al.* The genome and transcriptome of Japanese flounder provide insights into flatfish asymmetry[J]. *Nature Genetics*, 2017, 49(1): 119-124.
- [70] Xu J, Li J T, Jiang Y L, *et al.* Genomic basis of adaptive evolution: the survival of Amur ide (*Leuciscus waleckii*) in an extremely alkaline environment[J]. *Molecular Biology and Evolution*, 2017, 34(1): 145-159.
- [71] Yang J X, Chen X L, Bai J, *et al.* The *Sinocyclocheilus* cavefish genome provides insights into cave adaptation[J]. *BMC Biology*, 2016, 14: 1.
- [72] Lin Q, Fan S H, Zhang Y H, *et al.* The seahorse genome and the evolution of its specialized morphology[J]. *Nature*, 2016, 540(7633): 395-399.
- [73] Ye N H, Zhang X W, Miao M, *et al.* *Saccharina* genomes provide novel insight into kelp biology[J]. *Nature Communications*, 2015, 6: 6986.
- [74] Wang S, Zhang J B, Jiao W Q, *et al.* Scallop genome provides insights into evolution of bilaterian karyotype and development[J]. *Nature Ecology & Evolution*, 2017, 1(5): 120.
- [75] Gong G R, Dan C, Xiao S J, *et al.* Chromosomal-level assembly of yellow catfish genome using third-generation DNA sequencing and Hi-C analysis[J]. *GigaScience*, 2018, 7(11): giy120.
- [76] Shao C W, Li C, Wang N, *et al.* Chromosome-level genome assembly of the spotted sea bass, *Lateolabrax maculatus*[J]. *GigaScience*, 2018, 7(11): giy114.
- [77] Liu H, Chen C H, Gao Z X, *et al.* The draft genome of blunt snout bream (*Megalobrama amblycephala*) reveals the development of intermuscular bone and adaptation to herbivorous diet[J]. *GigaScience*, 2017, 6(7): 1-13.
- [78] Xu J, Bian C, Chen K C, *et al.* Draft genome of the Northern snakehead, *Channa argus*[J]. *GigaScience*, 2017, 6(4): 1-5.
- [79] Gao Y, Gao Q, Zhang H, *et al.* Draft sequencing and analysis of the genome of pufferfish *Takifugu flavidus*[J]. *DNA Research*, 2014, 21(6): 627-637.
- [80] Song L S, Bian C, Luo Y J, *et al.* Draft genome of the Chinese mitten crab, *Eriocheir sinensis*[J]. *GigaScience*, 2016, 5: 5.
- [81] Zhang X J, Sun L N, Yuan J B, *et al.* The sea cucumber genome provides insights into morphological evolution and visceral regeneration[J]. *PLoS Biology*, 2017, 15(10): e2003790.
- [82] Shin S C, Ahn D H, Kim S J, *et al.* The genome sequence of the Antarctic bullhead notothen reveals evolutionary adaptations to a cold environment[J]. *Genome Biology*, 2014, 15(9): 468.
- [83] Rondeau E B, Minkley D R, Leong J S, *et al.* The genome and linkage map of the northern pike (*Esox lucius*): conserved synteny revealed between the salmonid sister group and the Neoteleostei[J]. *PLoS One*, 2014, 9(7): e102089.
- [84] Austin C M, Tan M H, Croft L J, *et al.* Whole genome sequencing of the Asian arowana (*Scleropages formosus*) provides insights into the evolution of ray-

- fanned fishes[J]. *Genome Biology and Evolution*, 2015, 7(10): 2885-2895.
- [85] Xu T J, Xu G L, Che R B, *et al.* The genome of the miiuy croaker reveals well-developed innate immune and sensory systems[J]. *Scientific Reports*, 2016, 6: 21902.
- [86] AlMomin S, Kumar V, Al-Amad S, *et al.* Draft genome sequence of the silver pomfret fish, *Pampus argenteus*[J]. *Genome*, 2016, 59(1): 51-58.
- [87] Liu H P, Liu Q Y, Chen Z Q, *et al.* Draft genome of *Glyptosternon maculatum*, an endemic fish from Tibet plateau[J]. *GigaScience*, 2018, 7(9): giy104.
- [88] Lamichhaney S, Barrio A M, Rafati N, *et al.* Population-scale sequencing reveals genetic differentiation due to local adaptation in Atlantic herring[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2012, 109(47): 19345-19350.
- [89] Hohenlohe P A, Day M D, Amish S J, *et al.* Genomic patterns of introgression in rainbow and westslope cutthroat trout illuminated by overlapping paired-end RAD sequencing[J]. *Molecular Ecology*, 2013, 22(11): 3002-3013.
- [90] Housto R D, Taggart J B, Cézard T, *et al.* Development and validation of a high density SNP genotyping array for Atlantic salmon (*Salmo salar*)[J]. *BMC Genomics*, 2014, 15: 90.
- [91] Liu S K, Sun L Y, Li Y, *et al.* Development of the catfish 250K SNP array for genome-wide association studies[J]. *BMC Research Notes*, 2014, 7: 135.
- [92] Xu J, Zhao Z X, Zhang X F, *et al.* Development and evaluation of the first high-throughput SNP array for common carp (*Cyprinus carpio*)[J]. *BMC Genomics*, 2014, 15: 307.
- [93] Dupont C, Armant D R, Brenner C A. Epigenetics: definition, mechanisms and clinical perspective[J]. *Seminars in Reproductive Medicine*, 2009, 27(5): 351-357.
- [94] Jabbari K, Cacciò S, Païs de Barros J P, *et al.* Evolutionary changes in CpG and methylation levels in the genome of vertebrates[J]. *Gene*, 1997, 205(1-2): 109-118.
- [95] Shao C W, Li Q Y, Chen S L, *et al.* Epigenetic modification and inheritance in sexual reversal of fish[J]. *Genome Research*, 2014, 24(4): 604-615.
- [96] Niederhuth C E, Bewick A J, Ji L X, *et al.* Widespread natural variation of DNA methylation within angiosperms[J]. *Genome Biology*, 2016, 17: 194.
- [97] 丁勇, 许超, 吴季辉, 等. 表观遗传学研究进展[J]. *中国科学: 生命科学*, 2017, 47(1): 3-15.
Ding Y, Xu C, Wu J H, *et al.* Recent progress in epigenetics[J]. *Scientia Sinica Vitae*, 2017, 47(1): 3-15 (in Chinese).
- [98] 杨震飞, 刘波, 戈贤平, 等. 水产养殖环境胁迫对鱼类表观遗传的影响研究进展[J]. *大连海洋大学学报*, 2018, 33(2): 270-282.
Yang Z F, Liu B, Ge X P, *et al.* Research progress of aquaculture environmental stress on epigenetic regulation in fish: a review[J]. *Journal of Dalian Ocean University*, 2018, 33(2): 270-282 (in Chinese).
- [99] Mortazavi A, Williams B A, McCue K, *et al.* Mapping and quantifying mammalian transcriptomes by RNA-Seq[J]. *Nature Methods*, 2008, 5(7): 621-628.
- [100] Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics[J]. *Nature Reviews Genetics*, 2009, 10(1): 57-63.
- [101] Robledo D, Gutiérrez A P, Barria A, *et al.* Gene expression response to sea lice in Atlantic salmon skin: RNA sequencing comparison between resistant and susceptible animals[J]. *Frontiers in Genetics*, 2018, 9: 287.
- [102] Xu J, Li Q, Xu L M, *et al.* Gene expression changes leading extreme alkaline tolerance in Amur ide (*Leuciscus waleckii*) inhabiting soda lake[J]. *BMC Genomics*, 2013, 14: 682.
- [103] Meyer K D, Saletore Y, Zumbo P, *et al.* Comprehensive analysis of mRNA methylation reveals enrichment in 3'UTRs and near stop codons[J]. *Cell*, 2012, 149(7): 1635-1646.
- [104] Han Y X, Gao S G, Muegge K, *et al.* Advanced applications of RNA sequencing and challenges[J]. *Bioinformatics and Biology Insights*, 2015, 9(S1): 29-46.
- [105] Suravajhala P, Kogelman L J A, Kadarmideen H N. Multi-omic data integration and analysis using systems genomics approaches: methods and applications in animal production, health and welfare[J]. *Genetics Selection Evolution*, 2016, 48(1): 38.

- [106] Tsai H Y, Robledo D, Lowe N R, *et al.* Construction and annotation of a high density SNP linkage map of the Atlantic salmon (*Salmo salar*) genome[J]. G3: Genes, Genomes, Genetics, 2016, 6(7): 2173-2179.
- [107] Peng W Z, Xu J, Zhang Y, *et al.* An ultra-high density linkage map and QTL mapping for sex and growth-related traits of common carp (*Cyprinus carpio*)[J]. Scientific Reports, 2016, 6: 26693.
- [108] Liu H Y, Fu B D, Pang M X, *et al.* A high-density genetic linkage map and QTL fine mapping for body weight in crucian carp (*Carassius auratus*) using 2b-RAD sequencing[J]. G3: Genes, Genomes, Genetics, 2017, 7(8): 2473-2487.
- [109] Liu P, Wang L, Wong S M, *et al.* Fine mapping QTL for resistance to VNN disease using a high-density linkage map in Asian seabass[J]. Scientific Reports, 2016, 6: 32122.
- [110] Shao C W, Niu Y C, Rastas P, *et al.* Genome-wide SNP identification for the construction of a high-resolution genetic map of Japanese flounder (*Paralichthys olivaceus*): applications to QTL mapping of *Vibrio anguillarum* disease resistance and comparative genomic analysis[J]. DNA Research, 2015, 22(2): 161-170.
- [111] Yu Y, Zhang X J, Yuan J B, *et al.* Genome survey and high-density genetic map construction provide genomic and genetic resources for the Pacific White Shrimp *Litopenaeus vannamei*[J]. Scientific Reports, 2015, 5: 15612.
- [112] Jiao W Q, Fu X T, Dou J Z, *et al.* High-resolution linkage and quantitative trait locus mapping aided by genome survey sequencing: building up an integrative genomic framework for a bivalve mollusc[J]. DNA Research, 2014, 21(1): 85-101.
- [113] Yue G H, Wang L. Current status of genome sequencing and its applications in aquaculture[J]. Aquaculture, 2017, 468: 337-347.
- [114] 张晓峰, 耿波. 重要养殖鱼类经济性状QTL定位研究进展[J]. 水产学杂志, 2018, 31(1): 52-66.
Zhang X F, Geng B. Advances in QTL mapping of important economic traits in principal cultured fishes[J]. Chinese Journal of Fisheries, 2018, 31(1): 52-66 (in Chinese).
- [115] Gonen S, Lowe N R, Cezard T, *et al.* Linkage maps of the Atlantic salmon (*Salmo salar*) genome derived from RAD sequencing[J]. BMC Genomics, 2014, 15: 166.
- [116] Feng X, Yu X M, Fu B D, *et al.* A high-resolution genetic linkage map and QTL fine mapping for growth-related traits and sex in the Yangtze River common carp (*Cyprinus carpio haematopterus*)[J]. BMC Genomics, 2018, 19(1): 230.
- [117] Kuang Y Y, Zheng X H, Li C Y, *et al.* The genetic map of goldfish (*Carassius auratus*) provided insights to the divergent genome evolutions in the Cyprinidae family[J]. Scientific Reports, 2016, 6: 34849.
- [118] Fu B D, Liu H Y, Yu X M, *et al.* A high-density genetic map and growth related QTL mapping in bighead carp (*Hypophthalmichthys nobilis*)[J]. Scientific Reports, 2016, 6: 28679.
- [119] Sun C F, Niu Y C, Ye X, *et al.* Construction of a high-density linkage map and mapping of sex determination and growth-related loci in the mandarin fish (*Siniperca chuatsi*)[J]. BMC Genomics, 2017, 18: 446.
- [120] Wan S M, Liu H, Zhao B W, *et al.* Construction of a high-density linkage map and fine mapping of QTLs for growth and gonad related traits in blunt snout bream[J]. Scientific Reports, 2017, 7: 46509.
- [121] Liu S, Li Y, Qin Z, *et al.* High-density interspecific genetic linkage mapping provides insights into genomic incompatibility between channel catfish and blue catfish[J]. Animal Genetics, 2016, 47(1): 81-90.
- [122] Ao J Q, Li J, You X X, *et al.* Construction of the high-density genetic linkage map and chromosome map of large yellow croaker (*Larimichthys crocea*)[J]. International Journal of Molecular Sciences, 2015, 16(11): 26237-26248.
- [123] Wang L, Wan Z Y, Bai B, *et al.* Construction of a high-density linkage map and fine mapping of QTL for growth in Asian seabass[J]. Scientific Reports, 2015, 5: 16358.
- [124] Wang W J, Hu Y L, Ma Y, *et al.* High-density genetic linkage mapping in turbot (*Scophthalmus maximus* L.) based on SNP markers and major sex- and growth-related regions detection[J]. PLoS ONE, 2015, 10(3): e0120410.
- [125] Palaiokostas C, Bekaert M, Taggart J B, *et al.* A new SNP-based vision of the genetics of sex determination in European sea bass (*Dicentrarchus labrax*)[J].

- Genetics Selection Evolution, 2015, 47(1): 68.
- [126] Qiu G F, Xiong L W, Han Z K, *et al.* A second generation SNP and SSR integrated linkage map and QTL mapping for the Chinese mitten crab *Eriocheir sinensis*[J]. Scientific Reports, 2017, 7: 39826.
- [127] Wang J P, Li L, Zhang G F. A high-density SNP genetic linkage map and QTL analysis of growth-related traits in a hybrid family of oysters (*Crassostrea gigas*×*Crassostrea angulata*) using genotyping-by-sequencing[J]. G3: Genes, Genomes, Genetics, 2016, 6(5): 1417-1426.
- [128] Niu D H, Du Y C, Wang Z, *et al.* Construction of the first high-density genetic linkage map and analysis of quantitative trait loci for growth-related traits in *Sinonovacula constricta*[J]. Marine Biotechnology, 2017, 19(5): 488-496.
- [129] Wang X L, Chen Z H, Li Q Y, *et al.* High-density SNP-based QTL mapping and candidate gene screening for yield-related blade length and width in *Saccharina japonica* (Laminariales, Phaeophyta)[J]. Scientific Reports, 2018, 8: 13591.
- [130] Chang Y Q, Ding J, Xu Y H, *et al.* SLAF-based high-density genetic map construction and QTL mapping for major economic traits in sea urchin *Strongylocentrotus intermedius*[J]. Scientific Reports, 2018, 8(1): 820.
- [131] Yáñez J M, Naswa S, López M E, *et al.* Genomewide single nucleotide polymorphism discovery in Atlantic salmon (*Salmo salar*): validation in wild and farmed American and European populations[J]. Molecular Ecology Resources, 2016, 16(4): 1002-1011.
- [132] Palti Y, Gao G, Liu S, *et al.* The development and characterization of a 57K single nucleotide polymorphism array for rainbow trout[J]. Molecular Ecology Resources, 2015, 15(3): 662-672.
- [133] Qi H G, Song K, Li C Y, *et al.* Construction and evaluation of a high-density SNP array for the Pacific oyster (*Crassostrea gigas*)[J]. PLoS One, 2017, 12(3): e0174007.
- [134] Gutierrez A P, Turner F, Gharbi K, *et al.* Development of a medium density combined-species SNP array for Pacific and European oysters (*Crassostrea gigas* and *Ostrea edulis*)[J]. G3: Genes, Genomes, Genetics, 2017, 7(7): 2209-2218.
- [135] Sodeland M, Gaarder M, Moen T, *et al.* Genome-wide association testing reveals quantitative trait loci for fillet texture and fat content in Atlantic salmon[J]. Aquaculture, 2013, 408-409: 169-174.
- [136] Tsai H Y, Hamilton A, Tinch A E, *et al.* Genome wide association and genomic prediction for growth traits in juvenile farmed Atlantic salmon using a high density SNP array[J]. BMC Genomics, 2015, 16: 969.
- [137] Gonen S, Baranski M, Thorland I, *et al.* Mapping and validation of a major QTL affecting resistance to pancreas disease (Salmonid Alphavirus) in Atlantic salmon (*Salmo salar*)[J]. Heredity, 2015, 115(5): 405-414.
- [138] Correa K, Lhorente J P, Bassini L, *et al.* Genome wide association study for resistance to *Caligus rogercresseyi* in Atlantic salmon (*Salmo salar* L.) using a 50K SNP genotyping array[J]. Aquaculture, 2017, 472(S1): 61-65.
- [139] Campbell N R, LaPatra S E, Overturf K, *et al.* Association mapping of disease resistance traits in rainbow trout using restriction site associated DNA sequencing[J]. G3: Genes, Genomes, Genetics, 2014, 4(12): 2473-2481.
- [140] Palti Y, Vallejo R L, Gao G T, *et al.* Detection and validation of QTL affecting bacterial cold water disease resistance in rainbow trout using restriction-site associated DNA sequencing[J]. PLoS One, 2015, 10(9): e0138435.
- [141] Gonzalez-Pena D, Gao G T, Baranski M, *et al.* Genome-wide association study for identifying loci that affect fillet yield, carcass, and body weight traits in rainbow trout (*Oncorhynchus mykiss*)[J]. Frontiers in Genetics, 2016, 7: 203.
- [142] Geng X, Sha J, Liu S K, *et al.* A genome-wide association study in catfish reveals the presence of functional hubs of related genes within QTLs for columnaris disease resistance[J]. BMC Genomics, 2015, 16(1): 196.
- [143] Jin Y, Zhou X, Geng X, *et al.* A Genome-wide association study of heat stress-associated SNPs in catfish[J]. Animal Genetics, 2017, 48(2): 233-236.
- [144] Zhong X X, Wang X Z, Zhou T, *et al.* Genome-wide association study reveals multiple novel QTL associated with low oxygen tolerance in hybrid catfish[J]. Marine Biotechnology, 2017, 19(4): 379-

- 390.
- [145] Geng X, Liu S K, Yuan Z H, *et al.* A Genome-wide association study reveals that genes with functions for bone development are associated with body conformation in catfish[J]. *Marine Biotechnology*, 2017, 19(6): 570-578.
- [146] Li N, Zhou T, Geng X, *et al.* Identification of novel genes significantly affecting growth in catfish through GWAS analysis[J]. *Molecular Genetics and Genomics*, 2018, 293(3): 587-599.
- [147] Zheng X H, Kuang Y Y, Lv W H, *et al.* Genome-wide association study for muscle fat content and abdominal fat traits in common carp (*Cyprinus carpio*)[J]. *PLoS One*, 2016, 11(12): e0169127.
- [148] Chen L, Peng W Z, Kong S N, *et al.* Genetic mapping of head size related traits in common carp (*Cyprinus carpio*)[J]. *Frontiers in Genetics*, 2018, 9: 448.
- [149] Pearson T A, Manolio T A. How to interpret a genome-wide association study[J]. *JAMA*, 2008, 299(11): 1335-1344.
- [150] Gutierrez A P, Yáñez J M, Fukui S, *et al.* Genome-wide association study (GWAS) for growth rate and age at sexual maturation in Atlantic salmon (*Salmo salar*)[J]. *PLoS One*, 2015, 10(3): e0119730.
- [151] Zhou T, Liu S K, Geng X, *et al.* GWAS analysis of QTL for enteric septicemia of catfish and their involved genes suggest evolutionary conservation of a molecular mechanism of disease resistance[J]. *Molecular Genetics and Genomics*, 2017, 292(1): 231-242.
- [152] Barria A, Christensen K A, Yoshida G M, *et al.* Genomic predictions and genome-wide association study of resistance against *Piscirickettsia salmonis* in coho salmon (*Oncorhynchus kisutch*) using ddRAD sequencing[J]. *G3: Genes, Genomes, Genetics*, 2018, 8(4): 1183-1194.
- [153] Palaiokostas C, Bekaert M, Taggart J B, *et al.* A new SNP-based vision of the genetics of sex determination in European sea bass (*Dicentrarchus labrax*)[J]. *Genetics Selection Evolution*, 2015, 47(1): 68.
- [154] Wang L, Liu P, Huang S Q, *et al.* Genome-wide association study identifies loci associated with resistance to viral nervous necrosis disease in Asian seabass[J]. *Marine Biotechnology*, 2017, 19(3): 255-265.
- [155] Salzberg S L. Open questions: how many genes do we have?[J]. *BMC Biology*, 2018, 16(1): 94.
- [156] Wu C, Pan W. Integrating eQTL data with GWAS summary statistics in pathway-based analysis with application to schizophrenia[J]. *Genetic Epidemiology*, 2018, 42(3): 303-316.
- [157] Suravajhala P, Kogelman L J A, Kadarmideen H N. Multi-omic data integration and analysis using systems genomics approaches: methods and applications in animal production, health and welfare[J]. *Genetics Selection Evolution*, 2016, 48(1): 38.
- [158] Bardozzo F, Lió P, Roberto T. A study on multi-omic oscillations in *Escherichia coli* metabolic networks[J]. *BMC Bioinformatics*, 2018, 19(S7): 194.

Progress and perspective of the genome research in aquatic organisms

ZHENG Xianhu, KUANG Youyi, LÜ Weihua, LUAN Peixian,
SUN Zhipeng, LU Cuiyun, SUN Xiaowen*

(National and Local Joint Engineering Laboratory for Freshwater Fish Breeding; Key Laboratory of Freshwater Aquatic Biotechnology and Breeding, Ministry of Agriculture and Rural Affairs; Heilongjiang River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Harbin 150070, China)

Abstract: In this paper, the development and application of related technologies in aquatic genome research are reviewed. Taking the appearance of the next-generation sequencing (NGS) technology as a demarcation point, we firstly looked back to the researches of genetics and molecular biology of aquatic organisms in the early stages, and the basic researches for the whole-genome sequencing. And then we emphatically introduced the development of the NGS technology application for whole-genome sequencing and economic traits analysis. At last, the development trend of aquatic organisms whole-genome researches is prospected. The genetic mechanism of economic traits of aquatic organisms is highly complex, thus there are still a lot of difficulties in explaining the genetic mechanism at the whole-genome level. However, it is a basic law that trait is determined by genes, so the exploration will still be fascinating.

Key words: aquatic organisms; genome; next generation sequencing

Corresponding author: SUN Xiaowen. E-mail: sunxw2002@163.com

Funding projects: National Natural Science Foundation of China (31672655, 31602150)