

基于 GBS 技术的中华绒螯蟹的遗传特征分析

高祥刚¹, 鲍相渤¹, 高美玲²

(1. 辽宁省海洋水产科学研究院, 辽宁省海洋水产分子生物学重点实验室, 辽宁大连 116023; 2. 大连工业大学食品学院, 辽宁大连 116034)

摘要 [目的] 研究中华绒螯蟹经过多年的人工养殖、跨流域引种和增殖放流等活动可能会对其遗传特性造成的影响。[方法] 基于下一代 GBS 测序技术, 利用 SNP 标记对辽河家系、辽河野生、长江群体、海南群体、黄河群体进行遗传特征分析。[结果] 经过条件为测序深度 $4 \times$ 、Miss 0.5、次要等位基因频率 (MAF) > 0.05 的过滤后, 共获得 652 746 个高质量的 SNP 位点用于群体的遗传分析, 各群体平均观测杂合度 (H_o) 为 0.084 4~0.124 9, 平均期望杂合度 (H_e) 为 0.097 4~0.200 3, 群体平均 π 值 (P_i) 为 0.158 3~0.254 4, 群体 F_i 的平均测量值 (F_{is}) 为 0.054 2~0.238 43。两两群体间遗传分化指数 F_{st} 值结合系统发生树, 可以看出, 长江群体与海南群体的遗传距离较近, 分化较小, 表示海南的中华绒螯蟹苗种可能来自长江流域。[结论] 该研究揭示了不同水系中华绒螯蟹的遗传差异, 探讨了其遗传多样性水平, 为中华绒螯蟹资源的合理开发利用与保护提供理论依据。

关键词 中华绒螯蟹; GBS; SNP; 遗传多样性; 种群结构

中图分类号 S917.4 **文献标识码** A **文章编号** 0517-6611(2017)36-0080-03

Analysis of Genetic Diversity of Chinese Mitten Crab (*Eriocheir sinensis*) with Genotyping-by-Sequencing Technology

GAO Xiang-gang¹, BAO Xiang-bo¹, GAO Mei-ling² (1. Key Lab of Marine Fishery Molecular Biology of Liaoning Province, Liaoning Ocean and Fisheries Science Research Institute, Dalian, Liaoning 116023; 2. School of Food Science and Technology, Dalian Polytechnic University, Dalian, Liaoning 116034)

Abstract [Objective] To assess the genetic diversity and population structure of the wild and cultured *E. sinensis* populations (i. e.). [Method] Based on Genotyping-by-Sequencing technology, genetic character of Liaohe pond-reared population (LC), Liaohe wild population (LW), Changjiang pond-reared population (CJ), Hainan pond-reared population (HN), Huanghe pond-reared population (HH) was analyzed by using SNP marker. [Result] After screening with parameter of $dp4$, miss 0.5 and minor alleles frequency (MAF) > 0.05 , a total of high quality 652 746 SNP sites were retained to analyze genetic character of population. The mean observed heterozygosity (H_o) was from 0.084 4 to 0.124 9, the mean expected heterozygosity (H_e) was from 0.097 4 to 0.200 3, the mean nucleotide diversity (P_i) was from 0.158 3 to 0.254 4, the average pairwise inbreeding coefficient F_{is} ranged from 0.054 2 to 0.238 43. Both the F-statistic (F_{st}) among populations and phylogenetic analysis suggested that the CJ population had a very close relationship with HN population. It may indicate that the baby crabs of Hainan Province derived from Changjiang river. [Conclusion] The information may be beneficial for biodiversity conservation and management of *E. sinensis*.

Key words *Eriocheir sinensis*; GBS; SNP; genetic diversity; population structure

中华绒螯蟹 (*Eriocheir sinensis*) 俗称河蟹、毛蟹, 在动物分类地位上隶属于节肢动物门甲壳纲十足目爬行亚目方蟹科绒螯蟹属^[1-2], 属高等甲壳动物。该蟹广泛分布于辽河、黄河、长江、瓯江和闽江等我国沿海各大水系, 其肉味鲜美, 营养丰富^[3]。自 20 世纪 80 年代以来, 中华绒螯蟹养殖业发展迅速, 目前已成为水产养殖支柱产业之一^[4]。经过多年的人工养殖、跨流域引种和增殖放流等生产活动, 已造成不同水系种质资源严重混杂^[5-6]。因此, 为深入了解该蟹不同地理群体的种质差异和遗传特征, 更有效保护和管理利用中华绒螯蟹资源, 对不同地理群体中华绒螯蟹的遗传多样性和群体遗传结构研究势在必行。

近年来, 随着高通量测序技术的发展, 基因分型的成本持续降低, GBS (Genotyping-by-sequencing) 技术作为第二代深度测序基础上发展起来的简化基因组测序技术, 通过采用酶切加标签的方法, 使多样本高通量平行测序得以实现。这不仅大大降低了基因测序的成本, 也使对大样本全基因组的基因分型成为可能, 对深入了解种质资源的遗传背景和系统演化具有重要意义^[7]。同时, GBS 获得的短读序列可通过有参或无参基因组的形式进行拼接组装, 进而获得高密度

的 SNP 标记, 利用这些 SNP 标记可进行遗传图谱构建、加密, 基因组的辅助组装, 全基因组关联分析等相关研究。目前, GBS 技术作为基因分型的重要手段, 已经在遗传多样性研究、遗传图谱构建、种质鉴定及品种识别等领域得到广泛的应用^[7-10]。

采用 GBS 测序技术对中华绒螯蟹 5 个群体的遗传多样性及遗传结构展开研究, 以揭示不同中华绒螯蟹地理群体的遗传差异, 探讨其遗传多样性水平, 为该蟹种质资源的合理开发利用与有效保护提供理论依据。

1 材料与方法

1.1 试验材料 试验分析的群体为 5 个, 均为 (2 龄) 发育成熟个体, 分别是取自盘锦辽河口的野生中华绒螯蟹群体、培育的辽河家系群体、采集自长江南通的养殖群体、采自黄河口附近东营的养殖群体, 以及购买自海口海鲜市场的海南养殖群体 (表 1)。解剖出体壁肌肉保存于 $-40\text{ }^{\circ}\text{C}$ 冰箱备用。

1.2 试验方法

1.2.1 DNA 提取和 GBS 文库构建。 对采集的肌肉组织样品, 利用天根生化科技 (北京) 有限公司的基因组 DNA 提取试剂盒 (TIANamp Genomic DNA Kit) 提取 DNA。使用紫外分光光度计 (Nanodrop, Thermo Scientific), 在 260 和 280 nm 处分别测定 DNA 的吸光值, 并通过 1.2% 琼脂糖凝胶电泳检测 DNA 的质量。

提取的样品 DNA 送上海派森诺生物科技股份有限公司

基金项目 辽宁省科学事业公益研究基金项目 (2016004005)。

作者简介 高祥刚 (1980—), 男, 山东新泰人, 副研究员, 硕士, 从事动物分子遗传学研究。

收稿日期 2017-10-11

进行 DNA 质检、建库和测序,采用限制性内切酶(HindIII + Bfal)对全基因组 DNA 进行完全酶切,回收插入片段大小在 220 ~ 450 bp 之间的酶切片段,按照 Double Digest Genotyping - by - Sequencing (dd - GBS) 的方式进行建库,利用第 2 代

测序技术(Next - Generation Sequencing, NGS),基于 Illumina Hiseq 测序平台,对文库进行双末端(Paired - end, PE)150 bp 的测序。

1.2.2 单核苷酸多态性(SNP)鉴定,每条测序数据按照条码

表 1 5 个中华绒螯蟹群体的基本采样信息

Table 1 Sources of 5 populations of *E. sinensis* used in experiment

群体名称 Population name	缩写 Abbreviation	性质 Properties	采样地点 Sampling site	采样个体数 Sample number	采样时间 Sampling time
辽河家系 Liaohe pond-reared population	LC	养殖	盘锦	4	2016 - 10 - 10
辽河野生 Liaohe wild population	LW	野生	盘锦	5	2016 - 10 - 10
长江群体 Changjiang pond-reared population	CJ	养殖	南通	5	2016 - 09 - 27
海南群体 Hainan pond-reared population	HN	养殖	海口	5	2016 - 11 - 09
黄河群体 Huanghe pond-reared population	HH	养殖	东营	5	2016 - 09 - 29

划分到对应的样本中,对每条序列进行严格的质控筛选,去除原始测序序列中的接头序列和低质量的短读序列,包括单端测序中含有的 N 数量超过该条序列长度的 10%,或者单端测序中低质量($Q \leq 5$)碱基数超过该条序列长度 50% 的序列。SNP 检测采用 Samtools 软件进行,对齐的匹配文件用 Samtools 软件转换为 BAM 文件后进行变异调用, Samtools 收集 BAM 文件中的汇总信息,计算可能基因型的似然值,再利用 Bcftools 应用先验值进行变异调用,采用贝叶斯模型检测群体中的多态性位点获得 VCFs 文件。

1.2.3 遗传数据分析。对 5 个群体的样本,采用 stacks 程序包中的 populations 命令进行群体遗传多样性各参数、Fst 值计算,分析各群体的遗传多样性水平、群体间的遗传分化程度^[11-12]。采用 FastTree 软件中的 ML 算法,利用 P - distance 计算样品间的遗传距离,并构建系统发育树。建树完成后,对系统发育树分支的可靠性进行验证(bootstrap, 1000 replications)。

2 结果与分析

2.1 测序质量 GBS 测序得到的原始图像数据文件经碱基识别分析转化为原始测序序列,24 个中华绒螯蟹样本的总测序数据量为 44.7 Gb,去除低质量的序列后,产生的高质量序列

数据量为 38.5 Gb,平均每个样本数据量为 1 604 Mb。测序质量较高($Q_{20} \geq 94.67\%$, $Q_{30} \geq 85.38\%$),GC 分布正常,测序的峰值在 Q40 左右,测序质量较高。经过条件为测序深度 $4 \times$ 、Miss0.5、次要等位基因频率(MAF) > 0.05 的过滤后,最后共获得 652 746 个高质量的 SNP 位点用于群体的遗传分析。

2.2 群体遗传多样性分析 采用 stacks 程序包获得各群体的遗传多样性结果,辽河家系、辽河野生、长江群体、海南群体、黄河群体的平均观测杂合度(H_o)分别为:0.109 6、0.122 2、0.084 4、0.124 9、0.110 5,平均期望杂合度(H_e)分别为:0.145 0、0.097 4、0.172 5、0.190 7、0.200 3,群体平均 π 值(P_i)分别为:0.191 1、0.158 3、0.203 8、0.241 3、0.254 4,群体 F_i 的平均测量值(F_{is})分别是:0.125 9、0.054 2、0.234 8、0.193 7、0.238 4(表 2)。

2.3 群体遗传结构分析 两两群体间遗传分化指数 Fst 值显示,辽河家系与辽河野生群体之间的 Fst 值较大,长江群体与海南群体之间的 Fst 值较小。各群体间的 Fst 值差异不显著,为 0.134 375 ~ 0.278 688。由 ML 系统发生树(图 1)可以看出,辽河家系首先和辽河的野生中华绒螯蟹群体聚在一起,长江群体和海南群体聚在一起,最后这 4 个群体再与黄河群体聚在一起。

表 2 中华绒螯蟹遗传多样性结果统计

Table 2 Genetic polymorphic of 5 populations of *E. sinensis*

群体 Population	观测杂合度(H_o)	期望杂合度(H_e)	群体平均 π 值(P_i)	近交系数(F_{is})
LC	0.109 6	0.145 0	0.191 1	0.125 9
LW	0.122 2	0.097 4	0.158 3	0.054 2
CJ	0.084 4	0.172 5	0.203 8	0.234 8
HN	0.124 9	0.190 7	0.241 3	0.193 7
HH	0.110 5	0.200 3	0.254 4	0.238 4

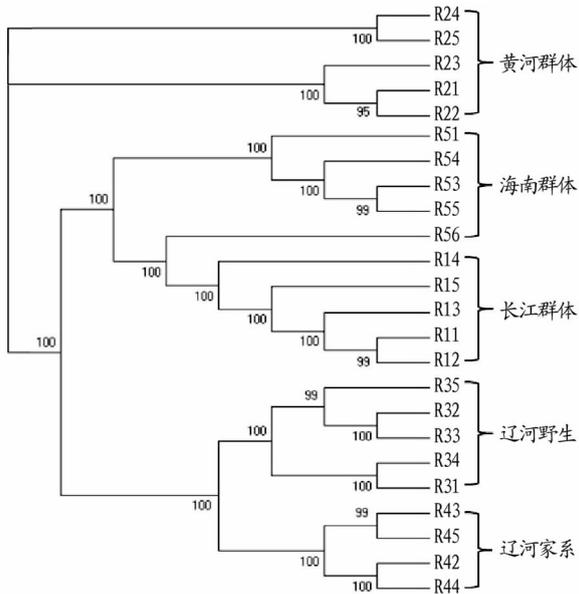
表 3 中华绒螯蟹群间 Fst 结果统计

Table 3 F-statistic (Fst) results of *E. sinensis*

LC	LW	CJ	HN	HH
LC	0.278 688	0.152 666	0.184 865	0.185 867
LW		0.186 55	0.234 189	0.239 758
CJ			0.134 375	0.139 275
HN				0.161 355

3 结论与讨论

基于 GBS 技术的辽河家系、辽河野生、长江群体、海南群体、黄河群体 5 个群体的 SNP 遗传多样性,其 H_o 为 0.084 4 ~ 0.124 9, H_e 为 0.097 4 ~ 0.200 3, P_i 为 0.158 3 ~ 0.254 4。而闫龙等^[6]利用线粒体 D_Loop 控制区基因对取自盘锦、丹东、南京、垦利的野生群体的遗传多样性分析,表明其 P_i 在 0.010 ~ 0.019,其遗传多样性处于较高的水平;刘青等^[13]采用微卫星标记分析取自长江、辽河、黄河的 6 个中华绒螯蟹



注:节点上标注的数字(bootstrap 值)代表该分枝中所有物种聚成一个分枝的可信性

Note:Numbers on the nodes show the confidence level from bootstrap analysis

图1 中华绒螯蟹 ML 系统发生树

Fig.1 Phylogenetic tree of *E. sinensis* by ML method

群体遗传水平,均表现出较高的遗传杂合度水平($H_o = 0.702 \sim 0.744$);熊良伟等^[14]利用微卫星标记分析中华绒螯蟹兴化、盘锦、宣城养殖群体的遗传水平,发现其 H_o 为 $0.70 \sim 0.75$; H_e 为 $0.87 \sim 0.90$,该蟹养殖群体也具有较高的遗传多样性水平。该研究与闫龙等^[6]相比,群体平均 π 值 (P_i)要高一个数量级,与刘青等^[13]、熊良伟等^[14]的微卫星分析结果相比,则遗传多样性水平要低很多,而产生这样的分析结果可能是分析方法不同造成的,该研究采用的第2代高通量测序技术,产生的数据能更真实,误差更小。

当群体内观测杂合度小于期望杂合度,同时 F_{is} 为正值,则表示群体内近交程度较为严重^[15],该研究的5个群体的 F_{is} 均为正值,除辽河野生群体的 $H_o > H_e$,其他群体的 $H_o < H_e$,说明我国中华绒螯蟹养殖群体内的近交程度可能较严重,存在种质衰退的风险。

群体的遗传分化指数 (F_{st})能揭示种群间的遗传分化程度, F_{st} 值为 $0 \sim 0.05$,群体的遗传分化较弱; F_{st} 值为 $0.05 \sim 0.15$ 时,遗传分化水平中等; F_{st} 值大于 0.15 时,遗传分化较大^[16]。该研究的几个群体间,除了长江群体与海南群体、黄河群体的 F_{st} 值分别为: $0.134\ 375$ 、 $0.139\ 275$,小于 0.15 ,说明长江群体与这2个群间的遗传分化较小;其余群体间的 F_{st} 值均大于 0.15 ,说明各群间的遗传分化较大。闫龙等^[6]基于线粒体控制区对4个野生群体的遗传多样性分析,两两群体间遗传分化指数 F_{st} 值显示在 $0.020\ 00 \sim 0.129\ 09$ 。熊良伟等^[14]对4个中华绒螯蟹群体在10个微卫星位点上平

均配对遗传分化系数统计量 F_{st} 介于 $0.012\ 3 \sim 0.020\ 7$,处于低等程度分化,该研究与它们相比,遗传分化较高。结合图1的系统发生树,可以看出,长江群体与海南群体的遗传距离较近,分化较小,表明海南的中华绒螯蟹苗种可能来自长江流域。黄河口收集的中华绒螯蟹养殖群体与长江流域的群体遗传分化也较近,表示其苗种有可能来自南方。Sui等^[17]和Chang等^[18]运用微卫星对中华绒螯蟹群体遗传学度研究表明,辽河水系与黄河水系、长江水系中华绒螯蟹群体之间存在较显著的遗传差异,而黄河水系与长江水系中华绒螯蟹群体间遗传差异不明显,该研究结果与上述研究结果相近。

参考文献

- [1] 王瑶,杨志刚,郭子好,等.中华绒螯蟹 RXR 基因全长 cDNA 克隆及表达分析[J].水产学报,2013,37(12):1761-1769.
- [2] LI X G, XU Z Q, ZHOU G, et al. Molecular characterization and expression analysis of five chitinases associated with molting in the Chinese mitten crab, *Eriocheir sinensis* [J]. Comp Biochem Physiol, Part B, 2015, 187: 110-120.
- [3] 孔晓瑜,喻子牛,刘亚军,等.中华绒螯蟹与日本绒螯蟹线粒体 COI 基因片段的序列比较研究[J].青岛海洋大学学报,2001,31(6):861-866.
- [4] 宋林生,季延宾,蔡中华,等.温度骤升对中华绒螯蟹(*Eriocheir sinensis*)几种免疫化学指标的影响[J].海洋与湖沼,2004,35(1):74-77.
- [5] 谷孝鸿,赵福顺.长江中华绒螯蟹的资源与养殖现状及其种质保护[J].湖泊科学,2001,13(3):267-271.
- [6] 闫龙,宋娜,王俊,等.基于线粒体控制区的中华绒螯蟹群体遗传多样性分析[J].水生生物学报,2015,39(3):615-620.
- [7] 王小柯,江东,孙珍珠.利用GBS技术研究240份宽皮柑橘的系统演化[J].中国农业科学,2017,50(9):1666-1673.
- [8] LIN M, CAI S B, WANG S, et al. Genotyping-by-sequencing (GBS) identified SNP tightly linked to QTL for pre-harvest sprouting resistance [J]. Theoretical and applied genetics, 2015, 128(7):1385-1395.
- [9] LU F, LIPKA A E, GLAUBITZ J, et al. Switchgrass genomic diversity, ploidy, and evolution: Novel insights from a network-based SNP discovery protocol [J]. PLoS Genetics, 2013, 9(1):1-14.
- [10] BAJAJ D, DAS S, BADONI S, et al. Genome-wide high-throughput SNP discovery and genotyping for understanding natural (functional) allelic diversity and domestication patterns in wild chickpea [J]. Sci Rep, 2015, 5:12468.
- [11] CATCHEN J M, AMORES A, HOHENLOHE P, et al. Stacks: Building and genotyping Loci de novo from short-read sequences [J]. G3 (Bethesda), 2011, 1(3):171-182.
- [12] CATCHEN J, HOHENLOHE P A, BASSHAM S, et al. Stacks: An analysis tool set for population genomics [J]. Mol Ecol, 2013, 22(11):3124-3140.
- [13] 刘青,刘皓,吴旭干,等.长江、黄河和辽河水系中华绒螯蟹野生和养殖群体遗传变异的微卫星分析[J].海洋与湖沼,2015,46(4):958-968.
- [14] 熊良伟,李真,马克异,等.利用微卫星 DNA 分子标记分析中华绒螯蟹养殖群体遗传分化[J].农业生物技术学报,2012,20(12):1441-1448.
- [15] FRANKHAM R, BALLOU J D, BRISCOE D A. Introduction to conservation genetics [M]. Cambridge: Cambridge University Press, 2010: 260-308.
- [16] BALLOUX F, LUGON-MOULIN N. The estimation of population differentiation with microsatellite markers [J]. Mol Ecol, 2002, 11(2):155-165.
- [17] SUI L Y, ZHANG F M, WANG X M, et al. Genetic diversity and population structure of the Chinese mitten crab *Eriocheir sinensis* in its native range [J]. Mar Biol, 2009, 156(8):1573-1583.
- [18] CHANG Y M, LIANG L Q, MA H T, et al. Microsatellite analysis of genetic diversity and population structure of Chinese mitten crab (*Eriocheir sinensis*) [J]. J Genet Genomics, 2008, 35(3):171-176.