

DOI: 10.19663/j.issn2095-9869.20180115001

http://www.yykxjz.cn/

邵如玉, 许建, 江炎亮, 张瀚元, 白庆利, 杨世勇, 徐鹏, 赵紫霞. 鲑鳟通用型低通量单核苷酸多态性芯片的开发. 渔业科学进展, 2019, 40(1): 53-61

Tai RY, Xu J, Jiang YL, Zhang HY, Bai QL, Yang SY, Xu P, Zhao ZX. Development of a universal 96 single nucleotide polymorphism array for salmonid fishes. Progress in Fishery Sciences, 2019, 40(1): 53-61

鲑鳟通用型低通量单核苷酸多态性芯片的开发*

邵如玉^{1,2} 许建² 江炎亮² 张瀚元²
白庆利³ 杨世勇⁴ 徐鹏^{2,5} 赵紫霞^{2①}

- (1. 上海海洋大学水产科学国家级实验教学示范中心 上海 201306; 2. 中国水产科学研究院
农业农村部水生动物基因组学重点实验室 渔业生物技术北京市重点实验室 北京 100141;
3. 中国水产科学研究院黑龙江水产研究所 哈尔滨 150070; 4. 四川农业大学动物科技学院 雅安 625014;
5. 厦门大学海洋与地球学院 福建省海洋生物资源开发利用协同创新中心 厦门 361102)

摘要 为开发常见鲑鳟养殖物种通用的遗传分析工具,本研究利用 Affymetrix 虹鳟(*Oncorhynchus mykiss*) 57K 高通量单核苷酸多态性(Single nucleotide polymorphism, SNP)芯片,对国内代表性鲑鳟养殖群体开展了分型检测,包括山女鳟(*Oncorhynchus masou masou*)、银鲑(*Oncorhynchus kisutch*)、美洲红点鲑(*Salvelinus fontinalis*)、白斑红点鲑(*Salvelinus leucomaenis*) 4个物种,从57,501个SNP标记中筛选出96个共享多态性标记,应用 Fluidigm 96.96 动态芯片平台,构建了大麻哈鱼属(*Oncorhynchus*)和红点鲑属(*Salvelinus*)通用型低通量 SNP 芯片。该芯片分型结果准确性较高,与 Affymetrix 高通量芯片分型一致性达到96.55%。使用该芯片对来自6个家系的48尾银鲑个体及其候选亲本进行检测,应用 Cervus 3.0.7 软件进行亲权鉴定,结果能够准确重现复杂家系的真实系谱。在用于单亲本亲权鉴定时,第一亲本非排除率(Non-exclusion probability for first parent, NE-1P)为 4.120×10^{-4} ;用于双亲本亲权鉴定时,双亲非排除率(Non-exclusion probability for parent pair, NE-PP)低至 6.219×10^{-12} ,表明该芯片在鲑鳟养殖群体谱系鉴定应用中具有较高的准确性。使用该芯片开展4个鲑鳟养殖群体遗传结构分析,样本分群聚类结果与其所属的分类阶元相符,能够准确反映群体遗传组分构成和遗传关系。本研究构建的低通量 SNP 芯片在常见鲑鳟养殖物种中具有良好的通用性,将其应用于养殖群体遗传分析,能够为鲑鳟制种、育种和引种等科学决策提供基因组信息参考。

关键词 鲑科鱼类;单核苷酸多态性芯片;亲权鉴定;群体遗传

中图分类号 S917.4 文献标识码 A 文章编号 2095-9869(2019)01-0053-09

鲑鳟是鲑科(Salmonidae)多种鱼类的统称,是冷水鱼主要养殖品种,在中高端水产品市场中占据了重要份额。我国有良好的冷水资源以及多样化的鲑鳟养

殖品种,除年产量最高的虹鳟(*Oncorhynchus mykiss*)外,还有银鲑(*Oncorhynchus kisutch*) (又称“银大麻哈鱼”)、山女鳟(*Oncorhynchus masou masou*)、美洲红点

* 国家科技支撑计划项目(2015BAD25B01)和中国水产科学研究院基本科研业务费专项(2015C007)共同资助 [This work was supported by National Science and Technology Pillar Program (2015BAD25B01), and Central Public-Interest Scientific Institution Basal Research Fund, CAFS (2015C007)].

① 通讯作者: 赵紫霞, 副研究员, E-mail: zhaozx@cafs.ac.cn

收稿日期: 2018-01-15, 收修改稿日期: 2018-02-09

鲑(*Salvelinus fontinalis*)(商品名“七彩鲑”)、白斑红点鲑(商品名“白点鲑”)(*Salvelinus leucomaenis*)等均已形成了一定的养殖规模(孙大江等, 2010; 户国等, 2012)。

由于各种鲑鳟养殖品种对生境的要求类似, 我国鲑鳟养殖的地域比较集中, 往往同一育苗场或养殖场同时养殖多种鲑科鱼类, 这些养殖物品种主要集中于大麻哈鱼属(*Oncorhynchus*)和红点鲑属(*Salvelinus*), 但除虹鳟外, 其他鲑鳟养殖种现有的基因组资源和遗传分析工具十分有限, 以往的遗传分析主要基于线粒体 DNA(Yu *et al.*, 2010)和微卫星标记(Small *et al.*, 2010; Naish *et al.*, 2013)来开展, 通常标记数量较少, 基因组覆盖度较低, 并且针对不同的物种需要开发不同的标记, 通用性较低。

本研究基于虹鳟 57K 单核苷酸多态性(Single nucleotide polymorphism, SNP)芯片(Palti *et al.*, 2015), 致力于开发常见鲑鳟养殖种间通用的分子标记, 形成小型 SNP 芯片, 从而利用同一套分子标记和检测方法, 实现对多个鲑鳟养殖物种的遗传分析, 将有利于显著降低遗传检测成本、简化技术人员培训流程、提升科学养殖管理效率。

1 材料与方法

1.1 实验材料

山女鳟、美洲红点鲑、白斑红点鲑采自中国水产科学研究院黑龙江水产研究所渤海冷水性鱼试验站, 银鲑采自四川省都江堰市三文鱼养殖基地和中国水产科学研究院房山试验基地。每个群体或家系各采样 20 尾, 剪取尾鳍贴于干燥的滤纸上, 56℃烘干 6 h, 常温保存。

1.2 DNA 提取

使用海洋动物基因组 DNA 提取试剂盒(天根生化)提取基因组 DNA, 1%琼脂糖凝胶电泳检测 DNA 完整性, NanoDrop 8000 超微量紫外可见分光光度计(Thermo Fisher)检测 DNA 纯度和浓度。

1.3 高通量 SNP 芯片分型

对山女鳟、美洲红点鲑、白斑红点鲑、银鲑 4 个群体进行线粒体控制区测序, 检测个体单倍型, 每个群体选取 DNA 质量高、具有不同单倍型的 8 个样本用于 57K 虹鳟 SNP 芯片(Affymetrix)分型, 分别记为 OK1~8(银鲑), OM1~8(山女鳟), SF1~8(美洲红点鲑), SL1~8(白斑红点鲑)。分型实验由纽勤生物科技(上海)

有限公司完成, 使用 SNPolisher 软件(Affymetrix)读取分型数据。使用 PLINK 1.09(<http://pngu.mgh.harvard.edu/purcell/plink/>)(Purcell *et al.*, 2007)进行分型结果过滤、统计和连锁不平衡(Linkage disequilibrium, LD)检验, 个体质控标准设为 Call rate (CR)>80%, 位点质控标准设为 CR>97%。

1.4 低通量 SNP 芯片分型

96 个 SNP 位点由 Fluidigm 公司设计合成探针, 使用 EP1 平台及 96.96 动态芯片(Fluidigm)开展分型检测。检测样本共计 96 尾, 包括进行过高通量 SNP 芯片检测的样本 32 尾, 以及新增银鲑样本 64 尾, 其中, 新增样本包含银鲑养殖家系 6 个, 每个家系样本各 8 尾(记为 OK9~56), 以及养殖家系候选亲本 16 尾(记为 OK57~72), 其中, 部分家系互为半同胞, 其真实亲本为 5 尾雌鱼和 3 尾雄鱼。通过 Fluidigm SNP genotyping analysis software 4.1 软件读取分型结果, 质控阈值设为 85, 对缺失和低质量分型数据进行人工复检。使用 Cervus 3.0.7 进行遗传多态性统计, 并对 48 尾银鲑样本(OK9~56)进行亲权鉴定(Kalinowski *et al.*, 2007)。使用 Structure 2.3.4 (<http://pritch.bsd.uchicago.edu/structure.html>), 对 32 尾鲑鳟群体样本(OK1~8、OM1~8、SF1~8 和 SL1~8)进行群体遗传结构分析(Falush *et al.*, 2003; Hubisz *et al.*, 2009)。

2 结果

2.1 虹鳟高通量 SNP 芯片在国内代表性鲑鳟养殖群体中的多态性分析

使用虹鳟 57K SNP 芯片, 对国内 4 个代表性鲑鳟养殖群体样本进行分型检测, 包括大麻哈鱼属的山女鳟、银鲑, 以及红点鲑属的美洲红点鲑和白斑红点鲑, 每个群体选取 8 尾个体。该芯片共包含虹鳟多态性位点 57,501 个, 32 尾个体全部通过 CR>80%的质控标准, 即每尾个体分型成功的位点数均大于 80%。位点质控标准设为 CR>97%, 即仅统计基因型无缺失的位点。各群体多态性位点统计结果如表 1 所示, 其中, 群体间分型位点比例波动较小, 在 65.96%~74.19%范围内, 但多态性位点比例差异很大, 银鲑群体内仅有 9.01%的多态性位点, 而美洲红点鲑群体内多态性位点比例达到 45.99%。

2.2 鲑鳟群体间共享多态性位点筛选

图 1 以韦恩图的形式展示了 4 个鲑鳟群体间多态性 SNP 位点的分布, 包括群体间共享的和群体特异

性的多态位点, 其中, 4 个群体共享的多态性位点有 89 个。为了避免基因组内紧密连锁的位点同时入选, 在 32 尾鲑鳟个体中对共享多态性位点进行 LD 检验, 剔除 R^2 值高于 0.3 的位点 4 个。根据共享多态性位点侧翼序列设计 Fluidigm 分型探针, 进一步剔除探针设计失败位点 6 个, 获得用于构建鲑鳟通用 SNP 芯片的共享多态性位点 79 个。由于位点数目不足 96 个,

从山女鳟、美洲红点鳟和白斑红点鳟 3 个群体共享的 448 个多态性位点中选取补充位点 17 个, 选择标准为: 经 LD 检验, 与已入选位点间 R^2 值低于 0.3, 并能够成功设计分型探针。96 个入选位点的侧翼序列及基因组定位信息见表 1。最终入选的芯片位点中, 编号 OS-79 及以前的位点, 为 4 群体共享多态位点, 编号 OS-80 及以后的位点, 为 3 群体共享多态位点。

表 1 鲑鳟代表性群体多态性 SNP 位点数目和比例

Tab.1 Number and percentage of polymorphic SNPs in representative Salmonid populations

物种 Species	采集地 Location	样本数 Number of individuals	分型位点数 Number of genotyped SNPs	分型位点比例 Percentage of genotyped SNPs (%)	多态位点数 Number of polymorphic SNPs	多态位点比例 Percentage of polymorphic SNPs (%)
山女鳟 <i>O. masou masou</i>	黑龙江 Heilongjiang	8	40540	70.50	6699	16.52
银鳟 <i>O. kisutch</i>	四川 Sichuan	8	42658	74.19	3844	9.01
美洲红点鳟 <i>S. fontinalis</i>	黑龙江 Heilongjiang	8	37926	65.96	17442	45.99
白斑红点鳟 <i>S. leucomaenis</i>	黑龙江 Heilongjiang	8	39151	68.09	6650	16.99
合计 Total		32	21637	37.63	12163	56.21

注: 分型位点数=在本列对应样本中符合 CR>97%质控标准的位点数目; 分型位点比例=分型位点数/57,501; 多态位点数=在本列对应样本中符合 CR>97%质控标准, 且存在 2 种及以上基因型的位点数目; 多态位点比例=多态位点数/分型位点数

Note: Number of genotyped SNPs = number of SNPs which passed the CR > 97% threshold among corresponding samples; Percentage of genotyped SNPs = number of genotyped SNPs/57,501; Number of polymorphic SNPs = number of SNPs which passed the CR > 97% threshold, and had 2 or more genotypes among corresponding samples; Percentage of polymorphic SNPs = number of polymorphic SNPs/number of genotyped SNPs

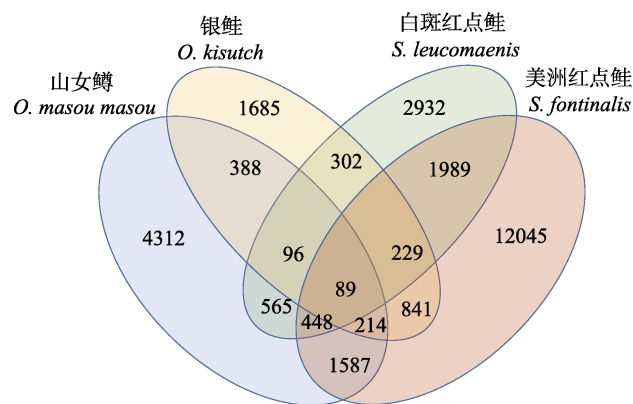


图 1 国内 4 个代表性鲑鳟养殖群体的多态性 SNP 位点韦恩图

Fig.1 Venn diagram for distribution of shared polymorphic SNPs among 4 representative salmonid populations in China

2.3 低通量 SNP 芯片位点多态性分析

使用构建的鲑鳟通用型低通量 SNP 芯片对 96 尾鲑鳟养殖个体进行分型检测, 在总计 9216 个分型反应中, 分型成功率为 97.48%, 检测样本中, 32 尾个体与高通量芯片检测样本相同, 其检测结果一致性为 96.55%。各位点遗传多态性统计结果见表 2。

2.4 基于低通量 SNP 芯片的养殖家系亲权鉴定

基于鲑鳟通用型低通量 SNP 芯片分型数据, 应用 Cervus 3.0.7 软件对来自 6 个家系的 48 尾银鳟个体进行亲权鉴定, 将同批次检测的其他 48 尾鲑鳟样本均作为待分析亲本群体, 结果表明, 无论是单亲本鉴定还是双亲本鉴定, 基于 96 个位点分型数据的亲权鉴定结果均与真实系谱相符。

根据表 2 中列出的各位点第一亲本非排除率 (Non-exclusion probability for first parent, NE-1P) 和双亲非排除率 (Non-exclusion probability for parent pair, NE-PP), 计算得该芯片用于单亲本亲权鉴定的 NE-1P 值为 4.120×10^{-4} , 用于双亲本亲权鉴定的 NE-PP 值为 6.219×10^{-12} 。图 2 展示了该芯片用于 48 尾银鳟子代亲权鉴定的对数优势比 (Logarithm of odds, LOD) 值, 由图 2 可见, 所有子代样本的检测结果都落在可靠限内。

2.5 基于低通量 SNP 芯片的养殖群体遗传结构分析

基于鲑鳟通用型低通量 SNP 芯片分型数据, 应用 Structure 软件对来自国内 4 个代表性鲑鳟养殖群

表2 鲑鳟通用型低通量 SNP 芯片位点遗传多态性统计
Tab.2 Genetic polymorphism of markers in the salmonid 96 SNP array

SNP 序号 SNP ID	57K 芯片对 应探针编号 57K ARRAY Probe ID	分型成功个体数 Number of genotyped individuals	表观杂合度 Observed heterozygosity (Hobs)	期望杂合度 Expected heterozygosity (Hexp)	多态信息含量 Polymorphic information content (PIC)	第一亲本非排除率 Non-exclusion probability for first parent (NE-1P)	双亲非排除率 Non-exclusion probability for parent pair (NE-PP)
OS-01	AX-89916678	93	0.624	0.454	0.350	0.898	0.734
OS-02	AX-89917717	89	0.067	0.165	0.150	0.987	0.867
OS-03	AX-89918173	96	0.667	0.502	0.375	0.875	0.719
OS-04	AX-89918910	95	0.389	0.499	0.373	0.877	0.720
OS-05	AX-89919056	93	0.613	0.488	0.367	0.882	0.723
OS-06	AX-89919812	94	0.574	0.429	0.336	0.909	0.743
OS-07	AX-89923356	92	0.174	0.160	0.146	0.987	0.870
OS-08	AX-89924099	94	0.415	0.498	0.373	0.877	0.720
OS-09	AX-89925937	92	0.402	0.323	0.270	0.948	0.783
OS-10	AX-89926458	93	0.634	0.488	0.367	0.882	0.723
OS-11	AX-89927367	94	0.213	0.208	0.185	0.979	0.840
OS-12	AX-89927655	93	0.333	0.427	0.335	0.910	0.743
OS-13	AX-89928755	94	0.245	0.232	0.204	0.973	0.827
OS-14	AX-89929078	94	0.681	0.475	0.361	0.888	0.727
OS-15	AX-89929344	96	0.375	0.355	0.291	0.938	0.770
OS-16	AX-89929576	91	0.538	0.502	0.375	0.875	0.719
OS-17	AX-89929982	93	0.376	0.307	0.259	0.953	0.790
OS-18	AX-89930877	93	0.398	0.320	0.268	0.949	0.784
OS-19	AX-89932421	93	0.161	0.333	0.277	0.945	0.779
OS-20	AX-89932513	95	0.158	0.438	0.341	0.905	0.739
OS-21	AX-89934355	87	0.195	0.417	0.329	0.914	0.747
OS-22	AX-89934998	96	0.26	0.272	0.234	0.963	0.806
OS-23	AX-89935444	92	0.554	0.502	0.375	0.875	0.719
OS-24	AX-89936054	94	0.213	0.208	0.185	0.979	0.840
OS-25	AX-89936602	90	0.289	0.308	0.259	0.953	0.789
OS-26	AX-89937397	93	0.398	0.380	0.306	0.929	0.760
OS-27	AX-89937627	93	0.581	0.472	0.359	0.890	0.728
OS-28	AX-89937660	93	0.376	0.320	0.268	0.949	0.784
OS-29	AX-89939260	90	0.211	0.241	0.211	0.971	0.822
OS-30	AX-89939691	96	0.417	0.460	0.353	0.895	0.732
OS-31	AX-89940240	94	0.543	0.407	0.323	0.918	0.750
OS-32	AX-89941041	93	0.355	0.390	0.313	0.925	0.756
OS-33	AX-89941252	93	0.366	0.447	0.346	0.901	0.737
OS-34	AX-89941864	94	0.436	0.473	0.360	0.890	0.728
OS-35	AX-89942293	94	0.457	0.501	0.374	0.876	0.719
OS-36	AX-89942762	96	0.500	0.377	0.305	0.930	0.761
OS-37	AX-89943547	94	0.670	0.503	0.375	0.875	0.719
OS-38	AX-89943578	94	0.266	0.262	0.227	0.966	0.811
OS-39	AX-89944357	93	0.516	0.414	0.327	0.915	0.748
OS-40	AX-89944600	93	0.344	0.286	0.244	0.959	0.799
OS-41	AX-89944806	91	0.495	0.497	0.372	0.878	0.721
OS-42	AX-89945237	92	0.402	0.489	0.368	0.882	0.723
OS-43	AX-89945326	92	0.413	0.330	0.274	0.946	0.780
OS-44	AX-89945455	91	0.231	0.205	0.183	0.979	0.842
OS-45	AX-89945939	93	0.645	0.486	0.367	0.883	0.724
OS-46	AX-89946091	94	0.255	0.224	0.198	0.975	0.831
OS-47	AX-89948186	94	0.181	0.461	0.354	0.895	0.732

续表 2

SNP 序号 SNP ID	57K 芯片对 应探针编号 57K ARRAY Probe ID	分型成功个体数 Number of genotyped individuals	表观杂合度 Observed heterozygosity (Hobs)	期望杂合度 Expected heterozygosity (Hexp)	多态信息含量 Polymorphic information content (PIC)	第一亲本非排除率 Non-exclusion probability for first parent (NE-1P)	双亲非排除率 Non-exclusion probability for parent pair (NE-PP)
OS-48	AX-89949006	96	0.552	0.496	0.372	0.878	0.721
OS-49	AX-89950618	95	0.253	0.267	0.231	0.965	0.808
OS-50	AX-89950801	93	0.645	0.493	0.370	0.880	0.722
OS-51	AX-89951792	94	0.404	0.421	0.331	0.913	0.746
OS-52	AX-89952308	93	0.333	0.307	0.259	0.953	0.790
OS-53	AX-89952765	95	0.168	0.409	0.324	0.917	0.750
OS-54	AX-89954998	95	0.284	0.260	0.225	0.967	0.812
OS-55	AX-89955274	94	0.255	0.224	0.198	0.975	0.831
OS-56	AX-89955318	93	0.344	0.502	0.375	0.875	0.719
OS-57	AX-89957360	96	0.281	0.411	0.325	0.916	0.749
OS-58	AX-89958032	93	0.172	0.158	0.145	0.988	0.871
OS-59	AX-89958585	96	0.365	0.450	0.348	0.900	0.735
OS-60	AX-89958649	90	0.522	0.470	0.358	0.891	0.729
OS-61	AX-89958745	93	0.366	0.352	0.289	0.939	0.771
OS-62	AX-89959154	96	0.552	0.500	0.374	0.876	0.720
OS-63	AX-89961084	94	0.394	0.318	0.266	0.950	0.785
OS-64	AX-89963184	95	0.653	0.503	0.375	0.875	0.719
OS-65	AX-89963523	95	0.084	0.400	0.319	0.921	0.753
OS-66	AX-89963568	95	0.642	0.459	0.352	0.896	0.733
OS-67	AX-89963593	95	0.337	0.346	0.285	0.941	0.773
OS-68	AX-89964518	94	0.234	0.360	0.294	0.936	0.768
OS-69	AX-89969612	94	0.617	0.470	0.358	0.891	0.729
OS-70	AX-89973928	92	0.163	0.151	0.139	0.989	0.876
OS-71	AX-89974037	94	0.245	0.216	0.192	0.977	0.835
OS-72	AX-89974888	94	0.553	0.480	0.363	0.886	0.726
OS-73	AX-89975150	94	0.170	0.174	0.158	0.985	0.860
OS-74	AX-89976123	92	0.293	0.393	0.314	0.924	0.755
OS-75	AX-89976174	91	0.549	0.481	0.364	0.886	0.726
OS-76	AX-89976394	93	0.258	0.226	0.200	0.975	0.830
OS-77	AX-89976445	96	0.271	0.279	0.239	0.961	0.803
OS-78	AX-89976740	95	0.663	0.459	0.352	0.896	0.733
OS-79	AX-89976936	94	0.309	0.397	0.317	0.922	0.754
OS-80	AX-89916591	95	0.526	0.502	0.375	0.875	0.719
OS-81	AX-89927162	94	0.511	0.502	0.375	0.875	0.719
OS-82	AX-89927505	91	0.187	0.170	0.155	0.986	0.863
OS-83	AX-89932853	93	0.559	0.414	0.327	0.915	0.748
OS-84	AX-89933870	94	0.053	0.232	0.204	0.973	0.827
OS-85	AX-89950286	93	0.505	0.497	0.372	0.878	0.721
OS-86	AX-89916505	92	0.598	0.496	0.371	0.879	0.721
OS-87	AX-89918155	93	0.441	0.443	0.344	0.903	0.738
OS-88	AX-89922603	95	0.389	0.328	0.273	0.947	0.781
OS-89	AX-89923232	96	0.281	0.243	0.212	0.971	0.821
OS-90	AX-89936048	95	0.568	0.426	0.334	0.910	0.743
OS-91	AX-89940938	96	0.177	0.162	0.148	0.987	0.868
OS-92	AX-89944836	94	0.277	0.402	0.320	0.920	0.752
OS-93	AX-89952599	93	0.527	0.469	0.358	0.891	0.729
OS-94	AX-89957593	93	0.376	0.463	0.355	0.894	0.731
OS-95	AX-89961734	96	0.427	0.450	0.348	0.900	0.735
OS-96	AX-89976669	96	0.583	0.466	0.356	0.893	0.730

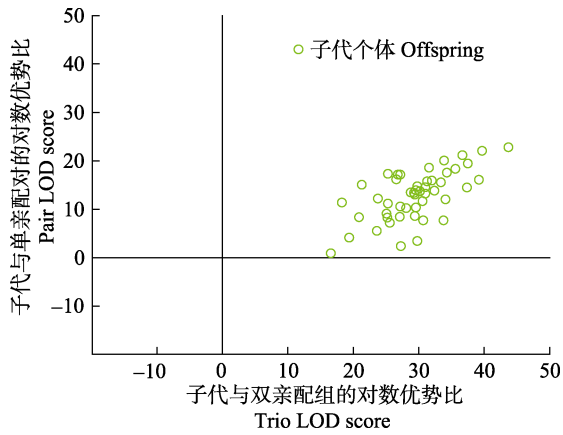


图2 48尾银鲑家系样本的亲权鉴定准确性分析
Fig.2 Parent assignments for 48 *O. kisutch* offsprings

亲权鉴定基于鲑鳟通用型低密度 SNP 芯片分型数据。每个点代表一尾子代个体。横坐标表示双亲本亲权鉴定数据，纵坐标表示单亲本亲权鉴定数据
The parent assignments analysis was based on genotyping data determined by the salmonid 96 SNP array. Each dot represents one offspring. Horizontal ordinate indicates data for parent pair assignment. Longitudinal ordinate indicates data for single parent assignment

体的 32 尾个体开展遗传结构分析。图 3 展示了假定祖源群体数(K)分别为 2 和 4 时的 Structure 遗传结构图，图中不同颜色代表不同的祖源成分。

$K=2$ 的假设下，所有样本以属为分类阶元进行聚类，而 $K=4$ 的假设下，样本以种为分类阶元进行聚类。每尾样本中占主导地位的遗传组分与个体所属的分类阶元相符，同时包含少量其他遗传组分，2 组分析结果都显示，银鲑样本 OK2、OK7，山女鲡样本 OM1、OM7，美洲红点鲑样本 SF8、白斑红点鲑样本 SL2、SL6 为非主导遗传组分占比较高的个体。

3 讨论

各种鲑科鱼类物种间亲缘关系较近，基因组序列相似性较高，因而具有大量共享的分子标记。梁利群等(2004)从 50 个虹鳟微卫星标记中筛选出 12 个在黑龙江乌苏里白鲑(*Coregonus ussuriensis*)自然群体中具有多态性的标记用于遗传多样性分析。鲁翠云等(2016)从 150 个虹鳟微卫星标记中筛选出 48 个在白斑红点鲑养殖群体中具有多态性的标记，并开展了群体遗传结构分析。目前，虹鳟的全基因组测序已经完成，

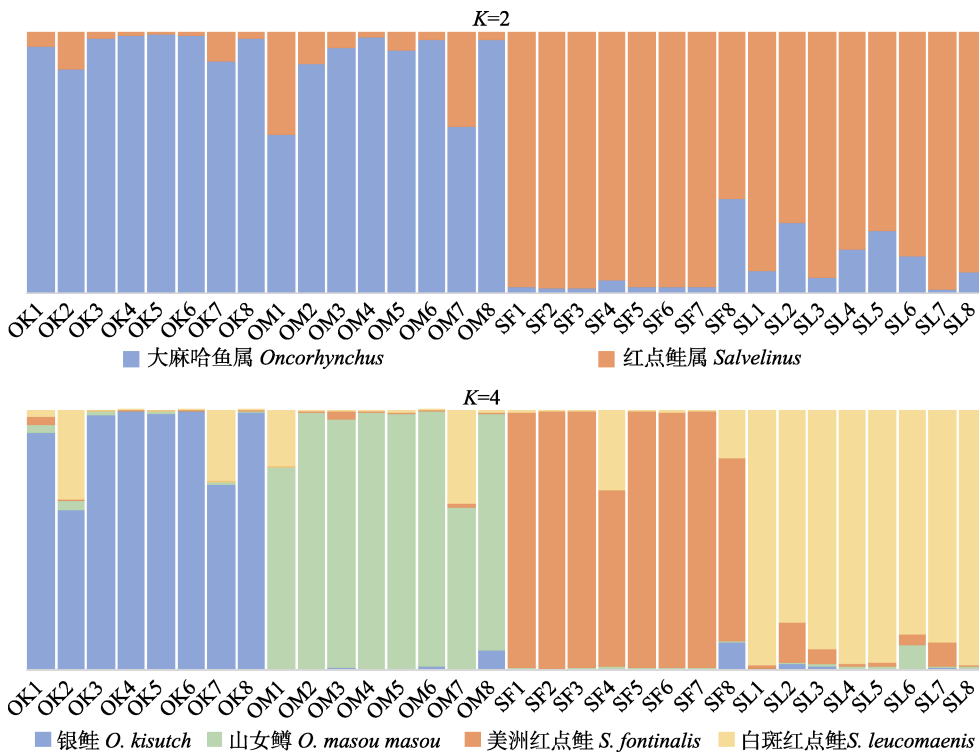


图3 国内 4 个代表性鲑鳟养殖群体遗传结构分析
Fig.3 Structure analysis for 4 representative salmonid populations in China

遗传结构分析基于鲑鳟通用型低密度 SNP 芯片分型数据。每种颜色代表一种假定的祖源成分。每条柱形代表一尾个体
The structure analysis was based on genotyping data determined by the salmonid 96 SNP array. Each color represents one putative ancestral fraction. Each bar represents one individual

并开发了多种通量的 SNP 分型芯片等遗传分析工具,建立了基于基因组信息的现代化育种、育苗体系。本研究基于这些新开发的虹鳟基因组信息资源,开展了共享标记筛查,使用虹鳟高通量 SNP 芯片,在国内代表性鲢鳟养殖群体中对 57,501 个标记进行了分型检测。

结果显示,在跨物种检测中,标记分型成功比例和多态比例都明显低于虹鳟样本。在相同的质控标准(CR > 97%)下,该芯片在不同虹鳟群体中的分型位点数为 49,299~49,468 个(Palti *et al.*, 2015),而在其他鲢鳟养殖物种中的分型位点数仅为 37,926~42,658 个;各虹鳟群体中多态性位点比例在 21.5%~92.8% 范围内波动(Palti *et al.*, 2015),而在其他鲢鳟养殖群体中的多态性位点比例仅为 9.01%~45.99%。但由于 SNP 标记的总量巨大,基因组覆盖度高,在基因组编码区和非编码区都有大量分布,这些物种间共享的 SNP 仍能提供较为丰富的遗传多态信息,鲤高通量 SNP 芯片也曾被成功应用于鲫(*Carassius auratus*)、草鱼(*Ctenopharyngodon idellus*)等 8 种近缘鱼类遗传分析中(Xu *et al.*, 2014)。由表 1 可见,跨物种分型成功位点数目与目标物种间的亲缘关系相关,山女鳟、银鲑与虹鳟同属于大麻哈鱼属,亲缘关系较近,因此,分型位点数略高于美洲红点鲑和白斑红点鲑;而多态性位点数目主要取决于群体本身的遗传多样性,虽然银鲑与虹鳟亲缘关系更近,但其多态位点比例远低于与虹鳟亲缘关系较远的美洲红点鲑,表明采样的银鲑群体遗传多样性较低,亟待改善。

Fluidigm 低通量分型平台有 48.48 和 96.96 两种芯片,分别适用于 48 和 96 个位点的同时分型。Xu 等(2017)使用 48.48 芯片构建了鲤(*Cyprinus carpio*)低密度 SNP 芯片,该芯片在用于家系亲权鉴定时的准确率为 94%。Liu 等(2016)构建虹鳟低密度 SNP 芯片则使用了 96.96 芯片,结果表明,当使用其中 68 个位点开展家系亲权鉴定时,结果准确率为 100%,仅使用 48 个位点时,准确率为 99.2%,当位点数下降到 36 个时,准确率降至 92.5%。Harlizius 等(2011)就 SNP 芯片在杜洛克猪商品群体亲权鉴定中的应用开展过更为细化的统计分析,当候选亲本中仅包含真实亲本($n=66$)时,60 个 SNP 位点即可实现 100% 准确的亲权鉴定,而当候选亲本中混入了大量干扰个体($n=304$)时,则需要 80 个 SNP 位点来实现 100% 准确的亲权鉴定。鉴定的准确性受到子代样本数目、候选亲本数目,以及群体遗传结构、分型位点在群体中的多态性等多种因素影响,而本研究试图构建多物种通用的 SNP 芯片,还需要考虑部分位点可能在待测物

种或群体中分型失败的可能性,因此,采用 96.96 芯片,通过检测略高于必要数目的 SNP 位点,来保障分析结果的可靠性。

本研究经过高通量 SNP 芯片分型筛查,获得 4 个鲢鳟养殖群体共享的多态性标记共 89 个,涵盖了大麻哈鱼属和红点鲑属各 2 个物种,推测这些位点在鲑科常见养殖物种中都具有较高的通用性,可以作为鲢鳟通用型 SNP 芯片构建的候选位点。考虑到共享位点筛查所用的银鲑群体遗传多样性较低,可能导致多态性位点发掘不充分,从其他 3 个群体共享的多态性位点中选取了补充位点,补足 96 个位点用于芯片构建。

为了验证所构建的低通量 SNP 芯片分型结果准确性,使用该芯片对高通量 SNP 芯片检测过的 32 尾样本进行了重复检测,结果一致性为 96.55%,不一致的位点主要表现为分型失败,结果缺失。由此可见,Fluidigm 低通量 SNP 芯片在鲢鳟样本检测中的分型准确性较高。

高通量芯片分型结果显示,OK1~8 样本所属的银鲑群体遗传多样性不足,因此,在低通量芯片样本选择时,从另一个银鲑养殖群体中采集了亲本及子代样本共计 64 尾,即 OK9~72,以验证该芯片在银鲑中的适用性。分型结果显示,初筛银鲑群体中不具有多态性的 17 个位点中,有 12 个在新的银鲑样本中表现出多态性,证明芯片的位点选择是成功的,入选位在鲑科不同物种中均具有较高的多态性。

虹鳟低通量芯片的研究结果显示(Liu *et al.*, 2016),在用于亲权鉴定时,Cervus 软件的准确率高于其他同类软件。将同批次分型的其他鲢鳟样本均作为候选亲本,应用 Cervus 3.0.7 软件开展亲权鉴定,结果表明,基于该芯片分型结果能够准确重现复杂家系的真实系谱。亲权鉴定的准确性通过 LOD 值来反映,表示假设为真与假设为假的概率之比的 Log_{10} 对数值,当 LOD 值为 0 时,假设成立的概率为 50%,当 LOD 为正值时,认为假设成立,即存在亲子关系。由图 2 可见,多数子代个体与软件预测的亲本间配对或配组的 LOD 值均远大于零,特别是双亲鉴定时,所有配组的 LOD 值均高于 18 以上,表明亲权鉴定结果准确性很高。

除系谱鉴定外,该芯片还能够应用于群体遗传资源分析,如图 3 中的示例,基于 96 个 SNP 位点的基因频率分布,反映群体间的遗传组构成和遗传关系,筛查群体内部具有独特遗传组成的离群个体。在水产养殖业中,准确的系谱鉴定为科学制种和育种方案的规划奠定了基础(李东宇等, 2016),基于分子标记的群体遗传分析还可应用于种质资源评估(王军等,

2018)、增殖放流效果评估(司飞等, 2017)等多个领域, 因此, 本研究构建的鲑鳟通用型低密度 SNP 具有较为广泛的应用前景, 能够为鲑鳟制种、育种和引种等科学决策提供基因组信息参考。

参 考 文 献

- Falush D, Stephens M, Pritchard JK. Inference of population structure using multi locus genotype data: Linked loci and correlated allele frequencies. *Genetics*, 2003, 164(4): 1567–1587
- Harlizius B, Lopes MS, Duijvesteijn N, *et al.* A single nucleotide polymorphism set for paternal identification to reduce the costs of trait recording in commercial pig breeding. *Journal of Animal Science*, 2011, 89(6): 1661–1668
- Hu G, Gu W, Bai QL, *et al.* Research advances in genetics and breeding in salmonids. *Chinese Journal of Fisheries*, 2012, 25(3): 58–62 [户国, 谷伟, 白庆利, 等. 主要养殖鲑科鱼类遗传育种的研究进展. *水产学杂志*, 2012, 25(3): 58–62]
- Hubisz MJ, Falush D, Stephens M, *et al.* Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, 2009, 9(5): 1322–1332
- Kalinowski ST, Taper ML, Marshall TC. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, 2007, 16(5): 1099–1106
- Li DY, Kong J, Meng XH, *et al.* Development of multiplex PCR systems of microsatellite markers for Pacific white shrimp (*Litopenaeus vannamei*) and its application for parentage identification. *Progress in Fishery Sciences*, 2016, 37(3): 58–67 [李东宇, 孔杰, 孟宪红, 等. 凡纳滨对虾(*Litopenaeus vannamei*)微卫星多重 PCR 体系的建立及其在家系亲权鉴定中的应用. *渔业科学进展*, 2016, 37(3): 58–67]
- Liang LQ, Chang YM, Dong CZ. Analysis of genetic diversity for *Coregonus ussruensis* Berg in Heilongjiang River. *Journal of Fishery Sciences of China*, 2004, 11(6): 501–505 [梁利群, 常玉梅, 董崇智. 黑龙江乌苏里白鲑遗传多样性分析. *中国水产科学*, 2004, 11(6): 501–505]
- Liu S, Palti Y, Gao G, *et al.* Development and validation of a SNP panel for parentage assignment in rainbow trout. *Aquaculture*, 2016, 452: 178–182
- Lu CY, Cao DC, Li YF, *et al.* Evaluation of genetic potential of Siberian char *Salvelinus leucomaenis* cultured population using microsatellite markers from rainbow trout *Oncorhynchus mykiss*. *Chinese Journal of Fisheries*, 2016, 29(1): 1–7 [鲁翠云, 曹顶臣, 李永发, 等. 利用虹鳟微卫星标记评价白斑红点鲑养殖群体的遗传结构. *水产学杂志*, 2016, 29(1): 1–7]
- Naish KA, Phillips RB, Briec MS, *et al.* Comparative genome mapping between Chinook salmon (*Oncorhynchus tshawytscha*) and rainbow trout (*O. mykiss*) based on homologous microsatellite loci. *G3 (Bethesda)*. 2013, 3(12): 2281–2288
- Palti Y, Gao G, Liu S, *et al.* The development and characterization of a 57K single nucleotide polymorphism array for rainbow trout. *Molecular Ecology Resources*, 2015, 15(3): 662–672
- Purcell S, Neale B, Todd-Brown K, *et al.* PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 2007, 81(3): 559–575
- Si F, Sun ZH, Yu SS, *et al.* Analysis of the genetic variability of the captured population of Japanese flounder *Paralichthys olivaceus* in the coastal releasing area of Qinhuangdao. *Progress in Fishery Sciences*, 2017, 38(6): 18–24 [司飞, 孙朝徽, 于姗姗, 等. 秦皇岛近海增殖放流区牙鲆(*Paralichthys olivaceus*)回捕群体的遗传多样性分析. *渔业科学进展*, 2017, 38(6): 18–24]
- Small MP, Beacham TD, Withler RE, *et al.* Discriminating coho salmon (*Oncorhynchus kisutch*) populations within the Fraser River, British Columbia, using microsatellite DNA markers. *Molecular Ecology*, 2010, 7(2): 141–155
- Sun DJ, Wang BQ. Aquaculture of salmonids in China. *Chinese Journal of Fisheries*, 2010, 23(2): 56–63 [孙大江, 王炳谦. 鲑科鱼类及其养殖状况. *水产学杂志*, 2010, 23(2): 56–63]
- Wang J, Wang QY, Kong J, *et al.* SSR analysis on genetic diversity in breeding and wild populations of *Fenneropenaeus chinensis*. *Progress in Fishery Sciences*, 2017, 38(6): 18–24 [王军, 王清印, 孔杰, 等. 中国对虾人工选育群体与野生群体遗传多样性的 SSR 分析. *渔业科学进展*, 2018, 39(2): 1–8]
- Xu J, Feng J, Peng W, *et al.* Development and evaluation of a high-throughput single nucleotide polymorphism multiplex assay for assigning pedigrees in common carp. *Aquaculture Research*, 2017, 48(4): 1866–1876
- Xu J, Zhao Z, Zhang X, *et al.* Development and evaluation of the first high-throughput SNP array for common carp (*Cyprinus carpio*). *BMC Genomics*, 2014, 15(1): 307
- Yu JN, Azuma N, Yoon M, *et al.* Genetic structure and phylogeography of masu salmon (*Oncorhynchus Masou masou*) inferred from mitochondrial and microsatellite DNA Analyses. *Zoological Science*, 2010, 27(5): 375–385

(编辑 冯小花)

Development of a Universal 96 Single Nucleotide Polymorphism Array for Salmonid Fishes

TAI Ruyu^{1,2}, XU Jian², JIANG Yanliang², ZHANG Hanyuan²,
BAI Qingli³, YANG Shiyong⁴, XU Peng^{2,5}, ZHAO Zixia^{2①}

(1. National Demonstration Center for Experimental Fisheries Science Education, Shanghai Ocean University, Shanghai 201306;

2. Key Laboratory of Aquatic Genomics, Ministry of Agriculture and Rural Affairs, Beijing Key Laboratory of Fishery Biotechnology, Chinese Academy of Fishery Sciences, Beijing 100141;

3. Heilongjiang River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Haerbin 150070;

4. College of Animal Science and Technology, Sichuan Agricultural University, Yaan 625014;

5. Fujian Collaborative Innovation Center for Exploitation and Utilization of Marine Biological Resources, College of Ocean and Earth Sciences, Xiamen University, Xiamen 361102)

Abstract This study aimed to develop a universal genotyping tool for common salmonid species cultured in China. Using an Affymetrix rainbow trout (*Oncorhynchus mykiss*) 57 K single nucleotide polymorphism (SNP) array, four representative aquaculture salmonid populations were genotyped, including *Oncorhynchus masou masou*, *Oncorhynchus kisutch*, *Salvelinus fontinalis*, and *Salvelinus leucomaenis*. A total of 96 shared polymorphism markers was screened out of 57,501 SNPs, then a low-density SNP array was constructed based on a Fluidigm 96.96 dynamic array, which is universal for the genera *Oncorhynchus* and *Salvelinus*. Reliability of the genotyping results was high because the identity of results with those of the Affymetrix 57 K array reached 96.55%. Parentage assignment for 48 *Oncorhynchus kisutch* offspring was performed with Cervus 3.0.7 software, based on genotyping data by the 96 SNP array. The results matched perfectly with the true pedigrees among complicated families, proving high accuracy in the pedigree assignment application of the 96 SNP array. The non-exclusion probability for the first parent (NE-1P) was 4.120×10^{-4} during single parent assignment, whereas the non-exclusion probability for the parent pair (NE-PP) was as low as 6.219×10^{-12} . Genetic structure analysis for four salmonid populations was also performed using genotyping data of the 96 SNP array. Clustering results were in accordance with the taxonomic categories of the samples, whereas certain degrees of genetic introgression were observed among the populations, indicating that the genetic composition and relationships among populations were properly demonstrated in the analysis. All these results suggested that the developed 96 SNP array has great potential in genetic analyses for salmonid populations, which might provide helpful genome-wide information for plasm resource evaluation, selective breeding, high-quality fry production, and reasonable foreign strain introduction in the Chinese salmonid aquaculture.

Key words Salmonid fishes; Single nucleotide polymorphism array; Pedigree assignment; Population genetics

① Corresponding author: ZHAO Zixia, E-mail: zhaozx@cafs.ac.cn