

DOI: 10.3724/SP.J.1118.2011.00935

全基因组选择育种策略及在水产动物育种中的应用前景

于洋^{1,2}, 张晓军¹, 李富花¹, 相建海¹

1. 中国科学院 海洋研究所 实验海洋生物学重点实验室, 山东 青岛 266071;

2. 中国科学院 研究生院, 北京 100049

摘要: 全基因组选择的概念自 2001 年由 Meuwissen 等提出后便引起了动物育种工作者的广泛关注。目前, 澳大利亚、新西兰、荷兰、美国的研究小组已经应用该方法进行了优质种牛的选择育种, 并取得了很好的效果。此外在鸡和猪的选择育种中也有该方法的应用, 但在水产动物选育中尚未见该方法使用的报道。本文对“全基因组选择育种”的概念和提出背景进行了归纳, 对全基因组选择育种的优势进行了阐述, 并详细介绍了其具体的策略, 总结了目前全基因组育种所广泛采用的方法以及取得的成果, 旨在为该方法在水产动物育种方面的应用研究提供科学参考。

关键词: 全基因组选择; 水产动物育种; SNP; QTL; 全基因组育种值估计

中图分类号: S96

文献标志码: A

文章编号: 1005-8737-(2011)04-0935-08

人类对于动物的选择育种由来已久, 最初所进行的只是简单的人工驯化。随着遗传学的发展, 尤其是“数量遗传学理论”的提出, 动物育种技术进入快速发展时期。数量遗传学理论认为生物的若干性状如生长、抗病等是由许多的微小基因共同作用的, 这些决定数量性状的微小基因称为数量性状位点(quantitative trait locus, QTL)。数量遗传学就是利用遗传学和统计学原理, 研究 QTL 所决定的数量性状^[1]。过去的数量遗传学方法把控制某一数量性状的多个基因作为一个整体研究, 因此无法准确定位到单个微小基因, 也无法准确估计出单个基因的效应^[2]。随着分子标记技术从第一代的限制性片段长度多态性(RFLP)、随机扩增 DNA 多态性(RAPD), 发展到第二代微卫星 DNA 和新一代单核苷酸多态性(single nucleotide polymorphisms, SNP), 标记辅助选育(marker assisted selection, MAS)研究也取得了更

大的进步, 人们已经能够对于某些性状的 QTL 位点进行精确定位, Grisart 等^[3]在 2002 年发现了控制牛奶乳脂含量和其他牛奶特性的 QTL 位点以及该位点处的关键基因 DGAT1, 并鉴定了起决定作用的关键位点 K232A。相对于陆生生物, 水产动物遗传育种起步较晚, 目前已有较多分子标记开发和 QTL 定位^[4-7]以及基因定位方面的报道^[8-9], 但是要实际应用仍需要更加深入的研究。

虽然标记辅助选育已经在陆生生物育种中得到广泛应用, 在水产动物中也有相关分子标记开发的报道, 但是现行的标记辅助选育仍然有很大的局限性: 从深度上看, 它所研究的某一个性状的 QTL 不够精密, 仍有许多与性状相关的 sub-QTL 未能被发现^[10], 同时目前已被证实具有显著效应的基因或标记非常有限, 而所找到的数量性状基因或标记仅能解释有限的遗传变异^[11]; 从广度上看, 由于利用 MAS 方法的工作量极为

收稿日期: 2011-03-14; 修订日期: 2011-04-10.

基金项目: 国家自然科学基金资助项目(30730071; 30972245); 农业科技成果转化资金项目(2010GB24910700).

作者简介: 于洋(1987-), 硕士研究生. E-mail: yuy8866@163.com

通信作者: 张晓军, 副研究员. E-mail: xjzhang@ms.qdio.ac.cn

巨大,发现并证实一个有效的基因需要很长的时间和很高的成本,因此一次只能研究一个或少量性状的 QTL,很难涉及不同性状 QTLs 间的相互作用。近年来,随着越来越多的生物基因组得到破译,科学家提出了全基因组选择育种的概念。本文通过对“全基因组选择育种”的概念和提出背景进行归纳,并总结目前全基因组育种所广泛采用的方法以及取得的成果,旨在为该方法在水产动物育种方面的应用研究提供科学参考。

1 全基因组选择概念的提出

全基因组选择的概念是由 Meuwissen 等^[12]在 2001 年提出的,当时是基于一种理想的假设,即所有性状的 QTLs 都对应一个与之紧密连锁的 SNP 位点并可用该标记来代表;通过性状测定获得全基因组育种值,结合该个体所带的分子标记,应用统计学方法计算出每一个分子标记所对应的染色体片段的育种值大小;然后再对所选择的个体进行全基因组育种值估计(genomic estimated breeding value, GEBV),并进行选择。由此可见,全基因组选择是在传统 MAS 基础上的创新和改进,是用覆盖全基因组的标记进行的辅助选育^[13]。

Seidel^[14]通过一个简单的例子对全基因组选择中统计方法的应用进行了阐述。如表 1 所示,假设生物是二倍体,每个位点有两个等位基因,每个等位基因存在两种基因型,分别用 AB、CD、EF 代表位点 1、2、3(SNP-1、SNP-2、SNP-3)。表中列出了 3 个位点的 27 种组合,并给出了每种组合所对应的两个生物性状,例如牛奶中蛋白(milk protein)含量和生产周期(productive herd life)。由表可以看出,最佳的组合是 BBDDDEE 和 BBDDDEF,当 SNP-3 的基因型为 EE 或 EF 时,B 或 D 这种基因型出现的越多,奶中蛋白的含量越高,当 SNP-3 的基因型是 FF 时,SNP-1 和 SNP-2 的基因型对于牛奶蛋白的含量没有影响;同时可以看出,SNP-3 的基因型对生产周期没有影响,而 SNP-1 和 SNP-2 中的 B 和 D 基因型越多,生产周期越短,A 和 C 基因型越多,生产周期越长。这是对于简单的 3 个位点,每个位点只有两种基因型

的情况;如果是进行全基因组选择,进行选择 SNP 位点多达成千上万个,而进行分析的参考群体的个体数目又不可能达到这么多,由此会造成自由度不够的问题^[2]。因此,科学家研究了多种方法进行全基因组选择的估计,具体的估计方法将在后面进行介绍。

表 1 3 个 SNP 决定的单倍型的组合以及对应的表型性状^[14]
Tab.1 Illustration of the 27 combinations of three single nucleotide polymorphisms with phenotype^[14]

SNP-1	SNP-2	SNP-3	% milk protein	productive herd life (months)
AA	CC	EE	3.0	5.0
AB	CC	EE	3.1	4.5
BB	CC	EE	3.2	4.0
AA	CD	EE	3.3	3.5
AB	CD	EE	3.4	3.0
BB	CD	EE	3.5	2.5
AA	DD	EE	3.6	2.0
AB	DD	EE	3.7	1.5
BB	DD	EE	3.8	1.0
AA	CC	EF	3.0	5.0
AB	CC	EF	3.1	4.5
BB	CC	EF	3.2	4.0
AA	CD	EF	3.3	3.5
AB	CD	EF	3.4	3.0
BB	CD	EF	3.5	2.5
AA	DD	EF	3.6	2.0
AB	DD	EF	3.7	1.5
BB	DD	EF	3.8	1.0
AA	CC	FF	3.4	5.0
AB	CC	FF	3.4	4.5
BB	CC	FF	3.4	4.0
AA	CD	FF	3.4	3.5
AB	CD	FF	3.4	3.0
BB	CD	FF	3.4	2.5
AA	DD	FF	3.4	2.0
AB	DD	FF	3.4	1.5
BB	DD	FF	3.4	1.0

2 全基因组选择育种的优势

2.1 增加选择的准确性

通过精确的全基因组范围的 SNP 标记可以有效地提高选择的准确性^[14]。由于全基因组选择选用覆盖整个基因组的分子标记,这样可以将每个

起作用基因的效应包括在内,增加了选择的准确性。De Ross 等对全基因组选择(genomic selection)、传统选择(BLUP 法)和基因辅助选择(gene assisted selection)3 种方法的准确性进行了比较,发现在乳脂率(fat percentage)这一性状的预测准确率方面,全基因组选择能够达到 75%,而 BLUP 法的准确性只有 51%^[15]。

2.2 提高选择的效率

该方法可以同时多个性状进行选择,并显著地提高选择的效率。Illumina 公司推出的 50K 的牛基因组 SNP 芯片可以同时检测 50 000 个 SNP 位点,分析多个性状相关的 SNP 位点。VanRaden 等应用全基因组选择的方法,同时考虑 5 种生产性状(yield traits)、5 种健康性状(fitness traits)和 16 种体型性状,对北美荷尔斯坦因公牛进行了选择^[11]。

2.3 缩短代与代的间隔,降低生产成本

由于在动物的幼体时就可以进行 SNP 基因型的检测,因此大大缩短了育种周期,节约了育种成本。据估计,应用全基因组育种对奶牛进行筛选可以节约 90%的成本^[16]。

2.4 适合低遗传力性状的选择

对于低遗传力性状采用全基因组选择的方法可以明显地提高选择准确性^[17]。这是因为对于低遗传力性状,通过表型所获得的用于估计的信息较少,这会导致估计选择效应时准确率较低,而采用全基因组选择的方法,对于低遗传力的标记能进行很好地估计,因此所得出的估计准确性能够得到明显提高^[18]。

3 全基因组育种策略

全基因组选择提供了一种新的 MAS 育种策略,这种方法充分利用了目前越来越精确的分子标记。但是对于 Meuwissen 等于 2001 年提出的应用高密度分子标记进行全基因组育种来说,大部分育种生物的分子标记密度还没有达到 Meuwissen 论文中“1cM 一个标记”的要求,同时目前对如此多的标记进行分型花费巨大,所以许多科学家,包括 Meuwissen 本人也在对原来的方案进行改

进^[10]。事实上,目前广泛采用的方案如下:

第一步:进行全基因组 SNP 标记的筛选。SNP 具有数量多、分布广的特点,该特点使得筛选覆盖全基因组的标记成为可能。

第二步:使用筛选得到的 SNP 对参考群体进行分型,同时测定参考群体的表型性状,通过对 SNP 分型数据和表型数据进行关联分析,计算出带有相应标记的染色体片段的效应。

第三步:通过获得的这些信息,利用 SNP 芯片对选择个体进行育种值估计,筛选出育种值较高的个体进行强化培育(图 1)。

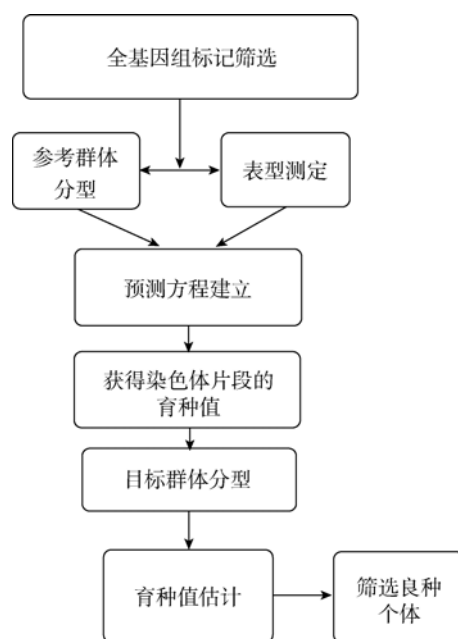


图 1 全基因组选择过程路线图

Fig.1 Flow chart of whole genome selection

3.1 全基因组 SNP 筛选

全基因组育种需要覆盖全基因组的 SNP 标记信息。对于研究基础较好的物种,由于已经积累了大量 SNP 标记,进行全基因组选育时可以直接利用这些已有的 SNP 信息;但是对于 SNP 发掘较少的物种来说,寻找覆盖全基因组的 SNP 仍是需要首先解决的一个问题。通过传统的凝胶电泳法和荧光检测法很难在短时间内筛选得到如此多的 SNP 标记,于是研究人员寻找高通量筛选 SNP 的新方法。

测序法直接寻找 SNP 是 SNP 发掘的最直接最准确的方式。一些物种在完成基因组测序后相继启动了 HapMap 计划,该计划通过基因组重测序的方法,比较同一物种不同个体碱基序列上的差别,发现了大量的 SNP。如通过鸡的 HapMap 计划,发现了 280 万个 SNP 位点^[19],而牛的 HapMap 计划发现了 220 万个 SNP 位点^[20]。

基因组重测序的方法虽然直接准确,但是花费较高,因此 Altshuler^[21]于 2000 年提出了一种改良的测序方法,即 Reduced Representation Shotgun (RRS)法,该方法的原理是将不同个体的基因组 DNA 用一种限制性内切酶消化,消化后的片段经过凝胶电泳分离,切割某一个长度范围内的片段做成文库,然后对文库进行测序,测序结果经过序列比对发现其中的 SNP。Altshuler 利用 4 种内切酶构建了 20 个文库,从中发现了 47 172 个人类 SNP 位点。而 Sanger 中心、华盛顿大学和 SNP 联合会(The SNP Consortium)在两年的时间内联合发现了 300 000 个 SNP 位点^[21]。此后研究人员利用该方法分别在猪、火鸡、大豆等畜禽和经济作物中寻找 SNP^[22-24]。但是目前还没有该方法在水产动物中应用的报道。

此外,由于 EST 测序、基因组测序、BAC 测序等产生大量序列信息,一些数据库如 NCBI 等中也积累大量的序列信息,这些数据来自不同的个体,这就为使用生物信息学发掘 SNP 提供了一个很好的资源,该方法在水产动物如大西洋鳕(*Gadus morhua*)^[25]、东方牡蛎(*Crassostrea virginica*)^[26]、大西洋鲑(*Salmo salar*)^[27]、凡纳滨对虾(*Litopenaeus vannamei*)^[28]等中已有报道。

3.2 参考群体 SNP 分型

进行 SNP 分型的方法很多,目前应用较广的主要有聚合酶链式反应-限制性片段长度多态性法(PCR-RFLP)^[29]、变性高效液相色谱法(denaturing high-performance liquid chromatography, DHPLC)^[30-31]、高分辨率溶解曲线法(high resolution melting)、TaqMan 探针技术^[32]、变性梯度凝胶电泳(denaturing gradient gelelectrophoresis, DGGE)、单链构象多态性(single-strand conformational

polymorphism, SSCP)^[33]、基因芯片(gene chip)^[34]和质谱法等;随着测序通量的提高和价格的降低,直接测序的方法也成为一种可行的 SNP 分型方法。由于基因芯片具有极高的检测效率,因此在国外全基因组选择育种大多采用该法。如 Illumina 公司的 beadchip 全基因组 SNP 第二代芯片每张同时分型>250 K SNP 位点,澳大利亚、新西兰、美国及荷兰的 4 家公司在对牛进行全基因组选择时均采用了这种芯片^[35]。该芯片可以同时対如此多的位点进行分型,因此单个位点分型的成本大大降低。为了提高后续预测的准确性,在进行分型时要筛选高质量的芯片和 SNP 位点进行后续预测方程的计算。Hayes 提出要首先对获得的 SNP 位点的数据进行筛选,筛选后的数据再应用于下一步的染色体片段育种值估计和预测方程建立^[35]。

3.3 计算染色体片段效应,建立预测方程

用所筛选的 SNP 对基因组进行标记的方法有单标记法、单倍型法和同源一致性(identity by descent, IBD)等方法。Calus 通过比较 4 种方法 SNP-1(单标记)、SNP-2(2 个 SNP 位点决定的单倍型)、HAP-IBD2(2 个 SNP 位点决定的单倍型加入同源一致性的概率信息)、HAP-IBD10(10 个 SNP 位点决定的单倍型加入同源一致性的概率信息)在预测无表型个体的 GEBV 的准确性研究中,发现对于高遗传力的性状, HAP-IBD10 在高密度的分子标记和低密度分子标记条件下均得到最高的预测准确率^[17]。但是 HAP-IBD 方法需要进行同源一致性检测,这也是一项非常繁琐的工作,目前还没有该方法实际应用的报道。当分子标记的密度较高时,应用单标记和单倍型的方法即可得到较高的准确率,因此单标记和单倍型方法的使用比较广泛。澳大利亚、美国、新西兰及荷兰等国对牛进行的全基因组选择便是采用的单标记或单倍型的方法。

通过单倍型和单标记进行 SNP 效应估计和预测方程建立的方法很多, Harris 等在利用 44146 个高质量的 SNP 位点对 4500 头公牛参考群体进行全基因组分析时采用了多种方法,包括 BLUP(best linear unbiased prediction)法、Bayesian

(贝叶斯)回归法、Bayes A、Bayes B、最小角回归法(least angle regression)等。为了提高预测的准确性,在分析中考虑了通过家系信息获得的“多基因效应”^[36]。通过比较这几种方法,Harris 发现 Bayesian 方法预测得到的准确性高于 BLUP 法,回归法得到的准确率最低^[37]。Van Raden 等^[11]也利用 3 576 个参考群体进行了全基因组分析,同样得到 Bayesian 法的预测准确性高于 BLUP 法的结论。De Roos 等进行的公牛全基因组分析采用了 Meuwissen 等提出的 Bayesian multiple QTL model 的方法,该方法中同样也考虑了多基因效应,使得预测的准确性更高。可见在建立预测方程时采用 Bayesian 法并加入“多基因效应”是一种较好的方法。

3.4 选择群体 SNP 体分型,计算 GEBV

使用芯片对选择群体分型,利用获得的 SNP 分型和预测方程计算 GEBV,根据 GEBV,筛选出所需要的个体进行培育或交配。通过选择,可以大大减少所要饲养动物的个体数目,同时由于 GEBV 的估计可以在幼体时期进行,这样选出的种畜进入了繁殖期即可使用,不需要等待后裔测定的结果,从而避免错过最佳的繁殖时期^[14]。

4 全基因组选择在水产动物育种中的应用前景分析

目前,全基因组选择的方法已经在美国、澳大利亚、新西兰、荷兰等国家的奶牛产业中得到较为广泛的应用,在肉鸡、猪的育种中也已经开始应用^[38-39],但是在水产养殖生物的育种中尚没有任何报道。其主要原因是由于水产生物分子标记、连锁图谱构建、QTL 定位等研究的起步较晚,研究相对落后。但是近年来,SNP 的研究技术取得很大进步,在水产动物中,SNP 标记的开发和应用已经成为一个新的研究热点,基于 SNP 标记的连锁图谱也已经在鱼虾贝等的多个物种中公布,因此在水产动物中进行全基因组选择育种已经具备了一定的条件。下面将就全基因组选择育种在水产动物育种领域的前景进行探讨。

首先,要进行全基因组育种,需要覆盖整个

基因组的分子标记,即要有精密的遗传图谱。目前水产动物遗传图谱的标记密度普遍较低,因此水产动物全基因组育种的第一步应该是增加图谱上的标记密度。一些重要的水产养殖动物,鱼类如虹鳟(*Oncorhynchus mykiss*)^[40-41]、罗非鱼(*Oreochromis spp.*)^[42]、虾类如斑节对虾(*Penaeus monodon*)^[43]凡纳滨对虾(*L. vannamei*)^[44-45]等的连锁图谱通过不断更新标记的种类和数目,已经达到了初步进行 QTL 精细分析的密度要求。但是进行全基因组选择需要更加精密和更广覆盖度的图谱,即需要具有出现频率高、分布广的标记构建遗传图谱。近年来,SNP 这一新型的分子标记引起水产生物研究者的关注^[25, 46-47]。Gorbach 报道了一种利用获得的凡纳滨对虾(*L. vannamei*)EST 序列寻找 SNP,预测到可能的 504 个 SNP 位点,并经检验准确性达 44%^[28]。Zhang 通过对东方牡蛎(*Crassostrea virginica*)EST 序列进行分析,同时对 6.8kb 的基因组进行测序共预测了 399 个可能的 SNP 位点^[26]。在遗传图谱的构建方面,Du 等利用 SNP 和表达序列标签(EST)构建了 *L. vannamei* 的第一张 SNP 连锁图,其中 418 个 SNP 标记分别归属于 45 个性连锁群,对于雌虾的覆盖度是 2 017 cM,雄虾的覆盖度为 2 130 cM^[48]。此外,Hubert 等利用 924 个 SNP 构建了大西洋鳕(*G. morhua*)初级遗传连锁图^[25],Moen 等利用 138 个微卫星和 304 个 SNP 构建了大西洋鲑(*S. salar*)的连锁图^[27]。这些 SNP 连锁图的出现,为全基因组选择育种奠定了很好的基础。相信随着研究的深入,更多海洋水产动物的 SNP 图谱会被报道,连锁图谱的精密程度也会进一步提高。

其次,随着测序技术的发展和测序成本的降低,已经陆续在海洋水产动物开展了全基因组测序,目前牡蛎(*C. gigas*)^[49]和半滑舌鳎(*Cynoglossus semilaevis*)^[50]和大黄鱼(*Pseudosciaena crocea*)初步的基因组测序和安装工作已经完成,大西洋鲑(*S. salar*)^[51]、鲈(*Ictalurus punctatus*)^[52]、鲤(*Cyprinus carpio*)及凡纳滨对虾(*L. vannamei*)^[53]等的基因组计划也正在进行中。全基因组序列的获得可以为全基因组选择育种的进行提供重要的基

础。首先,全基因组序列的获得为使用 RRS 法筛选 SNP 提供了参考序列,这样大大提高了 SNP 发现的效率和准确性^[54]。另外,利用获得的全基因组序列,可以确定筛选得到的 SNP 所代表的基因或染色体片段,为全基因组选择过程中确定 SNP 标记效应提供帮助。再者,全基因组序列可以提供一张精密的物理图,通过该图可以获得不同 SNP 标记之间的相对距离,为使用单倍型策略进行全基因组育种值估计提供依据。可以预言,全基因组测序可以有力地推进全基因组选择育种的开展。

除了精密的 SNP 标记外,进行全基因组选择另一项重要前提是要有一个性状稳定的家系。虽然水产生物人工选育的时间较短,但是由于它们大多繁殖迅速,产生的后代众多,因此到目前为止,已经陆续建立了许多的鱼、虾、贝类的家系^[55-57],通过对家系进行分析,可以得到若干表型性状数据,用于进行相应的关联分析。

因此,全基因组选择育种的策略在水产动物育种中的应用技术上是可行的。目前面临的主要问题是进行全基因组范围遗传标记分型的价格过高,导致实施全基因组选择的投入和产出比较低。为了提高投入产出比,现阶段可以采取使用 Tag SNP 进行选择^[58]或使用中等密度的 SNP 图谱进行辅助选育^[59],这样可以相应地降低分型检测的成本,提高效率。随着基因分型技术的发展和成本的降低,可以逐渐地提高 SNP 位点的密度,提高精确度,为水产动物新品种的快速培育提供重要指导。

参考文献:

- [1] 吴常信. 为创建一门新的边缘学科:《分子数量遗传学》而努力[J]. 中国畜牧, 1993, 29(4): 54-55.
- [2] 张慧, 王守志, 李辉. 畜禽全基因组选择[J]. 东北农业大学学报, 2010, 41(3): 145-149.
- [3] Grisart B, Coppieters W, Farnir F, et al. Positional Candidate Cloning of a QTL in Dairy Cattle: Identification of a Missense Mutation in the Bovine DGAT1 Gene with Major Effect on Milk Yield and Composition[J]. Genome Research, 2002, 12(2): 222-231.
- [4] 孟宪红. 中国明对虾抗白斑综合症病毒分子标记的筛选[J]. 中国水产科学, 2005, 12(1): 14-19.
- [5] 刘云国. 牙鲆抗鳃弧菌病 AFLP 分子标记筛选[J]. 中国水产科学, 2007, 14(1): 155-159.
- [6] 李琪. 海洋贝类微卫星 DNA 标记的开发及其在遗传学研究中的应用[J]. 中国水产科学, 2006, 13(3): 502-509.
- [7] Thorgaard G H, Nichols K M, Phillips R B. Comparative gene and QTL mapping in aquaculture species[J]. Israeli Journal of Aquaculture-Bamidgeh, 2006, 58(4): 341-346.
- [8] Leder E H. The candidate gene, clock, localizes to a strong spawning time quantitative trait locus region in rainbow trout[J]. J Hered, 2005, 97(1): 74-80.
- [9] Lyons R E, Dierens L M, Tan S H, et al. Characterization of AFLP markers associated with growth in the kuruma prawn, *Marsupenaeus japonicus*, and identification of a candidate gene[J]. Mar Biotechnol, 2007, 9(6): 712-721.
- [10] Meuwissen T H E, Goddard M E. Mapping multiple QTL using linkage disequilibrium and linkage analysis information and multitrait data[J]. Gen Select Evol, 2004, 36(3): 261-279.
- [11] VanRaden P M, Van Tassell C P, Wiggans G R, et al. Invited review: Reliability of genomic predictions for North American Holstein bulls[J]. J Dairy Sci, 2009, 92(1): 16-24.
- [12] Meuwissen T H, Hayes B J, Goddard M E. Prediction of total genetic value using genome-wide dense marker maps[J]. Genetics, 2001, 157(4): 1819-1829.
- [13] Meuwissen T. Genomic selection: marker assisted selection on a genome wide scale[J]. J Anim Breed Gen, 2007, 124(6): 321-322.
- [14] Seidel G E. Brief introduction to whole-genome selection in cattle using single nucleotide polymorphisms[J]. Reprod Fert Dev, 2010, 22(1): 138-144.
- [15] De Roos A P W, Schrooten C, Mullaart E, et al. Breeding value estimation for fat percentage using dense markers on Bos taurus autosome 14[J]. J Dairy Sci, 2007, 90(10): 4821-4829.
- [16] Schaeffer L R. Strategy for applying genome-wide selection in dairy cattle[J]. J Anim Breed Gen, 2006, 123(4): 218-223.
- [17] Calus M P L, Meuwissen T H E, de Roos A P W, et al. Accuracy of genomic selection using different methods to define haplotypes[J]. Genetics, 2008, 178(1): 553-561.
- [18] Piyasatian N, Fernando R L, Dekkers J C M. Genomic selection for marker-assisted improvement in line crosses[J]. Theoret Appl Gen, 2007, 115(5): 665-674.
- [19] International Chicken Polymorphism Map Consortium. A

- Genetic variation map for chicken with 2.8 million singlenucleotide polymorphism[J]. *Nature* 2004, 432: 717–722.
- [20] The Bovine HapMap Consortium. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds[J]. *Science*, 2009, 324: 528–532.
- [21] Altshuler D, Pollara V J, Cowles C R, et al. An SNP map of the human genome generated by reduced representation shotgun sequencing[J]. *Nature*, 2000, 407(6803): 513–516.
- [22] Kerstens H H D, Crooijmans R P M A, Veenendaal A et al. Large scale single nucleotide polymorphism discovery in unsequenced genomes using second generation high throughput sequencing technology: applied to turkey[J]. *Bmc Genomics*, 2009, 10(1): 479.
- [23] Wiedmann R T, Smith T P, and Nonneman D J. SNP discovery in swine by reduced representation and high throughput pyrosequencing[J]. *Bmc Genetics*, 2008, 9: 81.
- [24] Hyten D L, Song Q, Fickus E W, et al. High-throughput SNP discovery and assay development in common bean[J]. *Bmc Genomics*, 2010, 11: 475.
- [25] Hubert S, Higgins B, Borza T, et al. Development of a SNP resource and a genetic linkage map for Atlantic cod (*Gadus morhua*)[J]. *BMC Genomics*, 2010, 11(1): 191.
- [26] Zhang L S, Guo X M. Development and validation of single nucleotide polymorphism markers in the eastern oyster *Crassostrea virginica* Gmelin by mining ESTs and resequencing[J]. *Aquaculture*, 2010, 302(1-2): 124–129.
- [27] 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(1): 223.
- [28] Gorbach D M, Hu Z L, Du Z Q, et al. SNP discovery in *Litopenaeus vannamei* with a new computational pipeline[J]. *Anim Genet*, 2009, 40(1): 106–109.
- [29] Tao W J, Boulding E G. Associations between single nucleotide polymorphisms in candidate genes and growth rate in Arctic charr (*Salvelinus alpinus* L.)[J]. *Heredity*, 2003, 91(1): 60–69.
- [30] Abbas A, Lepelley M, Lechevrel M, et al. Assessment of DHPLC usefulness in the genotyping of GSTP1 exon 5 SNP: comparison to the PCR-RFLP method[J]. *J Biochem Biophysical Meth*, 2004, 59(2): 121–126.
- [31] Yoshida N, Nishimaki Y, Sugiyama M, et al. SNP genotyping in the beta(2)-adrenergic receptor by electronic microchip assay, DHPLC, and direct sequencing[J]. *J Hum Genetics*, 2002, 47(9): 500–503.
- [32] Martino A, Mancuso T, Rossi A M. Application of High-Resolution Melting to Large-Scale, High-Throughput SNP Genotyping: A Comparison with the TaqMan (R) Method[J]. *J Biomol Screen*, 2010, 15(6): 623–629.
- [33] Klinbunga S, Preechaphol R, Thumrunthanakit S, et al. Genetic diversity of the giant tiger shrimp (*Penaeus monodon*) in Thailand revealed by PCR-SSCP of polymorphic EST-derived markers[J]. *Biochem Gen*, 2006, 44(5-6): 222–236.
- [34] Schmitt A O, Bortfeldt R H, Brockmann G A. Tracking chromosomal positions of oligomers - a case study with Illumina's BovineSNP50 beadchip[J]. *Bmc Genomics*, 2010, 11: 411–415.
- [35] Hayes B J, Bowman P J, Chamberlain A J, et al. Invited review: Genomic selection in dairy cattle: Progress and challenges[J]. *J Dairy Sci*, 2009, 92(2): 433–443.
- [36] Xu S Z. Estimating polygenic effects using markers of the entire genome[J]. *Genetics*, 2003, 163(2): 789–801.
- [37] Harris B L, Johnson D L, Spelman R J. Genomic selection in New Zealand and the implications for national genetic evaluation[R]. *Proc Interbull Meeting Niagara Falls, Canada*, 2008.
- [38] Honkatukia M, Reese K, Preisinger R, et al. Fishy taint in chicken eggs is associated with a substitution within a conserved motif of the FMO3 gene[J]. *Genomics*, 2005, 86(2): 225–232.
- [39] Muira W M, Wong G K S, Zhang Y, et al. Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds[J]. *Proc National Academy of Sciences of the United States of America*, 2008, 105(45): 17312–17317.
- [40] Nichols K M, Young W P, Danzmann R G, et al. A consolidated linkage map for rainbow trout (*Oncorhynchus mykiss*)[J]. *Anim Gen*, 2003, 34(2): 102–115.
- [41] Guyomard R, Mauger S, Tabet-Canale K, et al. A Type I and Type II microsatellite linkage map of Rainbow trout (*Oncorhynchus mykiss*) with presumptive coverage of all chromosome arms[J]. *BMC Genomics*, 2006, 7: 302.
- [42] Lee B Y, Lee W J, Streelman J T, et al. A second-generation genetic linkage map of tilapia (*Oreochromis* spp.)[J]. *Genetics*, 2005, 170(1): 237–244.
- [43] Staelens J, Rombaut D, Vercauteren I, et al. High-density linkage maps and sex-linked markers for the black tiger shrimp (*Penaeus monodon*)[J]. *Genetics*, 2008, 179(2): 917–925.
- [44] Alcivar-Warren A, Meehan-Meola D, Park S W, et al. ShrimpMap: A low-density, microsatellite-based linkage map of the pacific whiteleg shrimp, *Litopenaeus vannamei*:

- Identification of sex-linked markers in linkage group 4[J]. J Shellfish Res, 2007, 26(4): 1259–1277.
- [45] Zhang L S, Yang C J, Zhang Y, et al. A genetic linkage map of Pacific white shrimp (*Litopenaeus vannamei*): sex-linked microsatellite markers and high recombination rates[J]. Genetica, 2007, 131(1): 37–49.
- [46] Moen T, Hayes B, Nilsen F, et al. Identification and characterisation of novel SNP markers in Atlantic cod: Evidence for directional selection[J]. BMC Genetics, 2008, 9(1): 18.
- [47] Ciobanu D C, Bastiaansen J W M, Magrin J, et al. A major SNP resource for dissection of phenotypic and genetic variation in Pacific white shrimp (*Litopenaeus vannamei*)[J]. Anim Gen, 2010, 41(1): 39–47.
- [48] Du Z Q, Ciobanu D C, Onteru S K, et al. A gene-based SNP linkage map for Pacific white shrimp, *Litopenaeus vannamei* [J]. Anim Genet, 2010, 41(3): 286–94.
- [49] Guofan Zhang, Ximing Guo, Li Li, et al. The Oyster Genome Project: An Update[R]. Ninth International Marine Biotechnology Conference, Qingdao, China, 2010: 371.
- [50] Chen S. Progress on genome sequencing project in half-smooth tongue sole (*Cynoglossus semilaevis*)[R]. The fifth international conference on genomics, Shenzhen, China, 2010: 50–51.
- [51] William S Davidson, Ben F Koop, ICSASG International Collaboration. Sequencing The Atlantic Salmon (*Salmo salar*) Genome The Old Fashioned Way[R]. Plant & Animal Genomes XIX Conference, 2011. San Diego, CA, USA, 033.
- [52] Liu J, Strategies For Efficient Assembly And Annotation Of The Catfish Whole Genome Sequence[R]. Plant & Animal Genomes XIX Conference, 2011. San Diego, CA, USA, 049.
- [53] Zhang X, Zhang T, Zhao C, et al. Research Progress in Sequencing of *Litopenaeus vannamei*[R]. Seventh International Crustacean Congress, Qingdao, China, 2010: 361.
- [54] Van Tassell C P, Smith T P L, Matukumalli L K, et al. SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries[J]. Nature Methods, 2008, 5(3): 247–252.
- [55] 陈锚. 凡纳滨对虾的选育与家系的建立[J]. 海洋科学, 2008, 32(11): 5–8.
- [56] 陈文华. 五个家系吉富罗非鱼的遗传多样性分析[J]. 生物技术通报, 2009(8): 83–87.
- [57] 张国范. 皱纹盘鲍中国群体和日本群体的自交与杂交 F_1 的 RAPD 标记[J]. 海洋与湖沼, 2002, 33(5): 484–491.
- [58] Wang W B, Jiang T. A New Model of Multi-Marker Correlation for Genome-Wide Tag SNP Selection[J]. Genome Informatics, 2008, 21: 27–41.
- [59] Habier D, Fernando R L, Dekkers J C M. Genomic Selection Using Low-Density Marker Panels[J]. Genetics, 2009, 182(1): 343–353.

Strategy of whole genomic selection breeding and its application prospect in aquaculture

YU Yang^{1,2}, ZHANG Xiaojun¹, LI Fuhua¹, XIANG Jianhai¹

1. Key Laboratory of Experimental Marine Biology, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China;

2. Graduate School, Chinese Academy of Sciences, Beijing 100049

Abstract: Whole genomic selection breeding was introduced firstly by Meuwissen in 2001. It has attracted a lot of attentions in animal breeding. The research groups in Australia, New Zealand, Netherland and United States successfully utilized this method in the selection of Holstein cattle. The utilization of this method in chicken and swine was also reported. However, in aquaculture animals, no related report was found. In this paper, conception of whole genomic selection breeding was introduced firstly, secondly the advantage of this method was clarified, and the strategy of this method was focused, the techniques applied in whole genomic selection breeding were also summarized. Finally, the application of whole genomic selection in aquaculture animal was discussed; suggestions were also given for the use of whole genomic selection breeding in aquaculture.

Key words: whole genomic selection; aquaculture animal breeding; SNP; QTL; GEBV

Corresponding author: ZHANG Xiaojun. E-mail: xjzhang@ms.qdio.ac.cn