

芯片技术在畜禽育种中的应用研究进展

马丽娜¹, 刘永进², 王 锦¹, 赵正伟¹, 马 青^{1*}

(1. 宁夏农林科学院动物科学研究所, 银川 750002; 2. 宁夏农业设计勘察院, 银川 750002)

摘要: 中国畜禽品种资源丰富, 且有许多优良性状基因, 但这些优良性状基因并没有被充分利用, 因此, 在基因水平上开展遗传资源的开发和利用是畜禽经济性状改良的重要方向。目前, 虽然传统系谱选择方法在育种工作中发挥了重要作用, 但存在准确率低、育种周期长等缺点。随着分子生物学技术的快速发展, 近年来先进的基因组测序和基因分型技术大大促进了畜禽育种方法的革新。从低通量、耗时的限制性片段多态标记(RFLP)到如今高通量、高密度的单核苷酸多态性(SNP)标记, 基因检测效率有了大幅度提高。基因芯片技术在分子标记辅助选择和全基因组选择育种研究中逐渐得到广泛应用, 成为畜禽育种的新技术手段和新热点。主要介绍了高、低密度 SNP 芯片技术在畜禽育种中的研究及应用, 并简述了其技术优势、存在问题及挑战、应用展望, 旨在表明基因芯片技术必将会成为畜禽分子育种工作中一项重要的基础技术, 在畜禽种业快速发展过程中起到重要的推动作用, 以期为基因芯片技术在畜禽育种中得到进一步应用提供理论参考, 推进中国畜禽育种遗传进展, 提升中国畜禽种业的科技竞争力。

关键词: 芯片技术; SNP; 高通量测序; 动物育种

中图分类号: S813.3

文献标识码: A

Doi: 10.16431/j.cnki.1671-7236.2020.01.012

开放科学(资源服务)标识码(OSID):



Research Advance on Application of Chip Technology in Livestock and Poultry Breeding

MA Lina¹, LIU Yongjin², WANG Jin¹, ZHAO Zhengwei¹, MA Qing^{1*}

(1. *Research Center of Grass and Livestock, Ningxia Academy of Agriculture and Forestry Sciences, Yinchuan 750002, China*; 2. *Agricultural Institute of Survey and Design of Ningxia, Yinchuan 750002, China*)

Abstract: There are abundant resources of livestock and poultry breeds in China, and there are many good trait genes. Unfortunately, these fine trait genes have not been fully utilized. Therefore, the development and utilization of genetics and resources at the genetic level is an important direction for improving economic traits of domestic animal genetics resources. The traditional pedigree selection method played a vital role in breeding while it also has shortcomings such as low accuracy and long breeding cycle. With the rapid development of molecular biology, advanced genome sequencing and gene typing technology have pushed the innovation of livestock and poultry breeding methods. The efficiency of genetic detection has been dramatically improved from the low-throughput, time-consuming restriction fragment polymorphism (RFLP) markers in the past to the single nucleotide polymorphism (SNP) with high-throughput and high density. Genotyping based on chip was widely used in research and breeding practice, becoming a new technology and hotspot in livestock and poultry breeding. In this review, the research and application of SNP chip

收稿日期: 2019-04-10

基金项目: 宁夏回族自治区农业育种专项(2013NYYZ04); 宁夏重点研发项目(2018BBF02016); 宁夏农林科学院项目(2017-NXKY05); 宁夏农林科学院先导项目(NKYQ-18-14)

作者简介: 马丽娜(1985-), 女, 宁夏银川人, 硕士, 研究方向: 动物遗传育种, E-mail: malina_2007nian@163.com

* 通信作者: 马 青(1973-), 男, 宁夏银川人, 硕士, 研究员, 研究方向: 动物营养与育种, E-mail: maqing1973@126.com

technology in livestock and poultry breeding, its technical advantages, existing problems, challenges and application prospects are summarized. The paper showed that gene chip technology would become an important basic technology in molecular breeding of livestock and poultry, and play an important role in the rapid development of livestock and poultry breeding industry, promoting the genetics progress of livestock and poultry breeding in China, and enhancing the scientific and technological competitiveness of livestock and poultry breeding industry in China.

Key words: chip technology; SNP; high-throughput sequencing; animal breeding

基因芯片, 又称 DNA 微阵列 (microarray), 属于生物芯片的一种, 是通过平面微细加工技术在固体芯片表面构建的微流体分析单元和系统。20 世纪 90 年代初, Affymetric 公司结合照相平板印刷技术、半导体技术、激光共聚焦扫描、寡聚核苷酸合成、荧光标记和计算机分析等技术在硅片上合成高密度的寡核苷酸点阵, 成功研制出世界上第一块基因芯片^[1-2]。随后, 各国的多家公司相继进行基因芯片的商业化开发, 如 Amersham Pharmacia Biotech 公司和 Molecular Dynamic 公司合作将特有的 Cy3 和 Cy5 荧光染料引入芯片检测; Illumina 公司成功开发了光纤微珠技术。如今, 基因芯片技术已在多个领域被广泛应用, 包括对基因表达水平的检测、基因诊断、药物筛选、医疗和测序、生物信息学研究、食品卫生监督、司法鉴定、国防等^[3-4]。农业在国民经济的发展中起举足轻重的作用, 遗传学已经由个体水平发展到细胞水平, 再到分子水平, 其不断发展为农业生产和人类健康带来科学理论指导, 基因芯片技术在畜禽遗传育种中的应用 (如动、植物优良性状基因的鉴定及定位等精准分子育种) 使得这一重要作用更好地发挥, 也逐渐得到认可和广泛应用。

与传统系谱育种方法不同, 分子育种可通过 DNA 检测进行育种指导, 实现快速、准确育种。分子育种从标记辅助选择育种 (marker assisted selection, MAS) 发展到芯片或测序技术指导下的全基因组选择育种 (genomic prediction, GP 或 genomic selection, GS), 使得动、植物育种工作变得更加便捷、精准, 实现了在基因水平挑选、培育具有优良性状的

品种, 在多个主要农作物、家畜育种研究工作中都有应用, 极大地缩短育种周期, 加快育种进程^[5]。但同时生物芯片仍然处于发展阶段, 在技术开发及应用推广中仍然存在一定的问题和挑战。作者从畜禽基因芯片、高/低密度芯片在育种中的应用、芯片技术的优势及面临的挑战等多方面, 总结了生物芯片技术在畜禽育种中的相关研究, 以期生物芯片技术应用于未来畜禽育种工作提供理论依据。

1 畜禽基因育种芯片

分子育种中最关键的技术是基因分型, 从传统的以 Southern 杂交检测技术为基础的第一代分子标记 RFLP, 到第二代基于 PCR 检测技术的 RAPD、AFLP 和 SSR 标记等, 基因分型技术迅速发展, 如今基因分型技术已发展为以基因芯片检测技术和高通量测序 (next generation sequencing, NGS) 为主, 以及基于芯片检测和测序结果而开发的竞争性等位基因特异性 PCR (kompetitive allele specific PCR, KASP) 等检测技术^[6-7]。相比检测成本昂贵、数据分析复杂的 NGS, 基因芯片拥有成本低、准确、快速、便捷等优点, 在农业育种 (如畜禽等育种) 研究工作中得到了广泛使用。

目前已成功开发多款商业化动物基因芯片 (表 1), 如 Tosser-Klopp 等^[8] 研制的羊 52K 芯片、宣苏哲^[9] 研制的中国首款鸡 55K SNP 芯片等, 这些动物基因分型芯片在动物遗传机制研究及育种工作中发挥了重要作用, 为国内外畜禽品种改良、育种与繁殖提供了科学技术手段。

表 1 商业化的畜禽基因育种芯片

Table 1 Commercial animal gene breeding chips

| 物种 Species | 产品 Products | 位点数 Sites | 平台 Platform |
|------------|---|-----------|-------------|
| 牛 Cattle | Bovine SNP50 | 50K | Illumina |
| | Bovine HD | 777K | Illumina |
| | GGP-LDV3 | 30K | GeneSeek |
| | GGP-LDV2 | 7K | GeneSeek |
| | GGP-UHD | 140K | GeneSeek |
| | Genomic Profiler for Dairy Cattle LD | 7K | Illumina |
| | GGP-HDT | 77K | GeneSeek |
| | GGP-LDV4 | 40K | GeneSeek |
| | Axiom [®] Genome-Wide BOS1 Bovine Array | 648K | Affymetrix |
| | Axiom [®] Bovine Genotyping Array | 50K | Affymetrix |
| 猪 Pig | Porcine SNP60 | 60K | Illumina |
| | Axiom [®] Porcine Genotyping Array | 658K | Affymetrix |
| | Porcine LD | 10K | Illumina |
| | Genomic Profiler Porcine HD | 80K | Illumina |
| 绵羊 Sheep | Ovine SNP50 | 50K | Illumina |
| | Sheep HD | 680K | Illumina |
| | Ovine LD | 5K | Illumina |
| 山羊 Goat | Goat SNP50 | 50K | Illumina |
| 马 Horse | Equine SNP50 | 50K | Illumina |
| | Equine SNP70 | 70K | Illumina |
| | Axiom [®] Equine | 670K | Affymetrix |
| 水牛 Buffalo | Axiom [®] Buffalo Genotyping Array | 90K | Affymetrix |
| 犬 Dog | Canine HD | 170K | Illumina |
| | Canine SNP20 | 20K | Illumina |
| | Affymetrix Canine SNP | 127K | Affymetrix |
| | Axiom [®] Canine Genotyping Array Set A | 460K | Affymetrix |
| | Axiom [®] Canine Genotyping Array Set B | 670K | Affymetrix |
| 猫 Cat | Feline SNP60 | 62K | Illumina |
| 鸡 Chicken | Axiom [®] Genome-Wide Chicken Genotyping Array | 580K | Affymetrix |
| | Chicken SNP50 | 50K | Illumina |
| 三文鱼 Salmon | Axiom [®] Salmon Genotyping Array | 130K | Affymetrix |

2 芯片在畜禽育种中的应用

全基因组选择是一种在全基因组范围内的标记辅助选择方法,利用覆盖基因组的遗传标记信息对个体进行遗传评估,在育种过程中可以提高准确性、缩短世代间隔、降低近交系数上升速度、加快育种进程,并能进行早期选种。畜禽的多数重要经济性状都属于数量性状,基因芯片技术和计算机技术的发展使得全基因组 SNP 数据能够被快速获得,通过使这些 SNP 标记与数量性状位点(quantitative trait

locus, QTL)形成连锁不平衡(linkage disequilibrium, LD),从而实现检测及挖掘重要性状位点^[10-11]。

继 MAS 标记辅助选择育种的深入研究和广泛应用,以芯片辅助为主的全基因组选择 GS 技术逐渐成为动、植物分子育种研究和工作中的一项重要指导技术,实现了从大量基因型到表型的预测,在群体进化研究、优良性状选育、缩短育种周期及优良基因引进等工作中发挥了重要作用。芯片也随着应用目的不同逐渐划分为以研究价值为主的高密度基因

芯片和以应用价值为主的低密度功能基因芯片。高密度基因芯片的 SNP 位点数占物种整个基因组的碱基数比例较大,价格较高,多倾向于研究型应用,如 QTL、GWAS、基因定位,群体进化分析,育种值评估,通过基因型填充达到测序密度。低密度基因芯片的 SNP 位点数占物种整个基因组的碱基数比例较小,价格稍低,多倾向于功能性应用,如育种公司、农场等大量样本的育种值评估,快速评估重要性状分布,可通过基因型填充提高芯片密度。

2.1 高密度基因育种芯片

目前,高密度基因芯片主要用于动物遗传科学研究工作中,应用方向多集中在群体进化分析、优良性状基因鉴定和定位、育种值估计等领域。

2.1.1 分析群体进化

畜禽遗传多样性是遗传育种研究的基础,可以利用群体间及个体间的遗传变异来研究畜禽的发育和生理机制,分析畜禽进化、驯化、品种形成过程。利用基因芯片技术对基因型分型,开展群体进化分析和遗传背景分析,为了解品种遗传背景、进化历史、分析和预测优良性状基因的遗传来源和进化方向等提供了可行途径。Cañas-Álvarez 等^[12]利用 Illumina Bovine HD Bead Chip 777K 芯片将 7 种西班牙种牛分为两大类群,所研究的西班牙肉牛具有较大的品种内多样性和较低的分化程度,为后期牛品种选育提供了理论基础,同时也加速了后期选择优良品种的速度。da Silva 等^[13]利用 Affymetrix 650K SNP 芯片对 2 175 只鸟进行基因分型,检测出了 41 029 个 CNV 位点,并通过实时荧光定量验证了 93.75% 的 CNVs。CNVs 的详细特征可以揭示选择的进化足迹,并提供其对野生种群表型变异的贡献。这些研究均显示出 SNP 芯片在畜禽遗传进化分析中的效力,对畜禽的遗传育种工作具有科学指导意义。

2.1.2 挖掘优良/重要性状基因/位点

利用基因芯片获得全基因组基因型数据,结合表型对重要性状进行全基因组关联分析,发掘和定位优良性状基因/位点,并通过精细定位对相应基因进行检测性标记的开发,是一种被广泛认为可用于标记辅助选择育种的重要方法。袁泽湖等^[14]利用 SNP 芯片分析了中国 11 个地方绵羊群体的遗传结构,结果显示,乌珠穆沁羊与除罗布羊以外的蒙古系绵羊均有直接遗传关系,与罗布羊有间接的遗传关系。兰蓉等^[15]利用 Illumina 公司 Iselect Goat60K 芯片对不同品种的山羊进行遗传结构评估,结果发现,云上黑山羊新品种在基因组水平上可与其育种素材明显区

分,在基因组层面已具备独立品种特点,已达到了云上黑山羊的育种期望目标。Martin 等^[16]应用山羊 SNP50 芯片对法国高山奶山羊和萨能奶山羊多乳头进行全基因组关联分析(GWAS),结果发现,在 10 号染色体上有 17 个达到染色体水平显著的 SNP 位点。兰蓉等^[17]利用山羊 SNP 芯片找到了 2 个位于 2 号染色体上的山羊产羔性状候选基因。Yi 等^[18]利用鸡 600K 芯片对 12 个不同品种的 96 只鸡进行全基因组检测,发现了 231 个拷贝数变异区域(copy number variation,CNV),其中 102 个 CNVs 来自 128 个与鸡免疫反应相关的基因。刘澳星等^[19]利用 Illumina Bovine SNP50 BeadChip 对奶牛进行全基因组关联分析,获得了多个与繁殖障碍疾病相关的基因。Dahlgren 等^[20]利用 170K 芯片发现了与犬周期性爪缺失病相关的基因区域。Zhu 等^[21]利用 55K Affymetrix Axiom 鸡基因分型芯片鉴定出了 5 个与鸡抗病性能(异嗜细胞与淋巴细胞的比率被研究学者认为是抗病性能指标)显著相关的 SNPs 位点,CARD11、BRIX1、BANP 是与鸡抗病性能相关的 3 个候选基因,为揭示抗病的遗传基础和今后的抗病标记辅助选择奠定了基础。以上研究充分表明,基因芯片技术近年来在发掘优良性状基因/位点等研究中发挥了重要作用。

2.1.3 估算育种值

以 SSR 等为主的 MAS 标记辅助选择育种技术大大地缩短了传统系谱育种所需的育种周期。随着芯片技术在基因型分型功能上的大力改进,使得通过基因型预测表型(即估算育种值)来指导育种的基因育种(genomic breeding)成为可能。Junqueira 等^[22]利用 777K 和 50K 芯片分析了不同牛种对虱的抗性,认为高质量遗传关系不会影响对遗传变异的评估,相对于传统的系谱分析,利用芯片技术会使育种值更精准。田佳等^[23]利用 Bovine SNP 54K 或 150K 芯片并借助中国荷斯坦牛全基因组选择技术平台对 SNP 效应进行评估,计算青年母牛基因组育种值(GEBV),以便筛选宁夏优质高产奶牛早期选育标准。通过基因芯片在全基因组范围内进行育种值的估计,实现了从单一标记的 MAS 方法进入到多个基因同时评估的精准育种方法。

2.1.4 基因型填充

通过结合基因芯片检测数据和部分参考群体的测序数据,对基因芯片进行基因型填充(genotype imputation)以填补缺失数据,是目前除测序技术之外获得更高密度数据的另一种途径。此外,基因芯片的检测结果也可作为对测序结果的校验^[24]。Ye 等^[25]利用鸡 600K 芯片检测

450 个样本,并利用 24 个重要样本的全基因组重测序结果对挑选的 60K 位点数据和 600K 芯片进行填充,获得全基因组高密度数据后进行了鸡的 GS 相关研究。van den Berg 等^[26]利用全基因组关联分析(genome-wide association studies, GWAS)结合芯片检测和测序获得的数据,与性状数据进行关联分析,挖掘了多个优良性状基因。但一些研究显示,当芯片标记密度达到一定程度后,如从 50K 提高至 777K 或者更高,所获得育种值的准确度提高并不显著,仅提高了约 1%^[27]。Meuwissen 等^[28]通过比较发现,当芯片标记密度达到一定水平后,增加标记密度虽然会使得育种值有较小幅度的提高,但最终获得的与表型相关联的标记与低密度芯片所获得的关联标记并没有差异,过高密度的基因型数据反而是一种浪费。因此,芯片密度在芯片定制和芯片选择过程中也是重要的考虑因素之一。基因填充可以使得测序数据或高密度芯片分析数据的可用性不断提高,可以评估更多的单核苷酸多态性、插入缺失位点(Indel)和基因拷贝数变异(CNVs),未来基因填充将会在越来越多的全基因组关联研究中发挥重要作用^[29]。

2.2 低密度功能基因育种芯片

高密度基因芯片因位点丰富等特点主要用于科学研究。随着科研工作的深入和科研成果转化的促进,用于辅助育种工作的低密度功能芯片逐渐得到开发和商品化。低密度功能芯片主要是对与表型具有较强相关的 SNP 进行挑选并定制而成的功能型芯片。低密度功能芯片的开发使得在未来农业育种工作中实验室科研成果向大田及农场等育种实践场所的转化成为可能。目前已有部分科研单位和一些国际化的育种公司开始使用低密度功能芯片,在精准育种、检测优良性状位点、快速评估育种材料、加快育种进程、促进动物基因资源的交流等多方面实现应用。

2.2.1 降低检测成本 当通过高密度芯片发现及定位优良基因或位点后,在实际育种工作中所需要使用的芯片则更倾向于功能芯片,即位点密度无需过高,但能含有所需优良基因/位点,并能检测、鉴定优良性状基因的芯片,从而对育种材料进行筛选,确定是否可以作为亲本材料。Hulsege 等^[30]利用牛 50K 和 777K 芯片对 4 个品种牛进行基因型检测,二者分别获得 30 447 和 452 525 个 SNPs 位点,但采用 Delta、Wright's Fst、Weir 和 Cockerham's Fst 以及综合 3 种方法进行比较后,发现能够用于区分 4 个品种的 SNP 位点差异不大,因此认为在育

种指导工作中 777K 相对于 50K 芯片优势较小。Gunia 等^[31]采用多种方法利用 French Charolais 肉牛作为试验材料对 50K 和 777K 芯片进行对比,认为利用 BayesC 方法分析 DEBV(deregressed estimated breeding values)是利用芯片标记育种最为准确的方法,同时也得出 777K 和 55K 在育种指导工作中所起到的作用差异不大,所需检测的位点在两款芯片中均可体现。在育种工作中低密度功能芯片性价比较高,既可以保证育种值的准确性,又可以降低检测成本,使得芯片技术在育种公司或农场中的使用成为可能。目前已报道有多家公司,如荷兰的 Hypor 公司、丹麦的丹育公司、加拿大的 TOPIGS 公司及中国广东温氏食品集团等,通过基于芯片技术的全基因组选择进行种猪遗传改良。

2.2.2 提高育种准确度 研究显示,当芯片密度达到一定程度后,由于避免了大量“冗余位点”的干扰,低密度芯片对育种值的估计准确度等同于或优于高密度芯片。Bolormaa 等^[32]利用羊 12K 低密度芯片对 50K 芯片进行加密,对 15 个性状在 50K 芯片加密前/后获得的育种值进行比较,发现在检测各性状时,虽然加密后的芯片获得的精度更高,但对育种工作起指导作用的 GEBVs 指标在加密后并没有明显的改变。Ogawa 等^[33]利用牛 50K 芯片将 3K 和 10K 位点的基因型数据进行填充,通过两种不同统计方法检测,认为当等位基因频率较高或在染色体末端时,检测准确度的一致率较低;当 3K 或 10K 位点均匀分布填充时,检测准确度会更高,可提高至 97%;当根据连锁不平衡情况来选择 SNP 位点时,基因型填充前所获得的计算精度反而较加密后更准确。目前科研人员多认为采用高/低密度芯片结合使用的方法,通过高密度芯片筛选优良位点,然后采用定制低密度功能芯片指导育种,将会是利用 GS 方法来指导育种工作的一个经济、高效的可行途径。

2.2.3 缩短育种周期 随着大量优良性状基因/位点被发现及定位,研究者们面临的最大挑战是如何在某个遗传背景中聚集多个优良基因/位点,即多基因的聚集或累加。基因聚合需要研究者在短时间内利用有限的经济基础对多个优良基因进行鉴定、累加及追踪等。低密度功能芯片的开发及使用可以同时检测多个优良基因/位点,对遗传材料进行快速评估,缩短了基因鉴定和育种周期,使得基因聚合成为可能。Auvuray 等^[34]利用羊 50K 芯片结合 8 个性状对新西兰的羊群体结构进行分析,认为结合 GRM(genomic relationship matrices)和传统的系谱

分子关系矩阵 (numerator relationship matrix, A) 的方法能最准确地预测育种值。Lopes 等^[35] 利用多款芯片对大量牛样本同时检测 10 个重要性状, 经分析获得育种模型, 并在 1 106 个体中进行验证, 认为虽然育种值比实际值偏小, 但实现了利用低密度芯片对多个性状同时检测, 从而实现对优良性状基因/位点进行预测和对育种工作进行指导。

2.2.4 评价动物基因资源情况 通过远缘杂交促进多个优良性状的聚合是已得到多数科研工作者广泛认可的一种育种方法。利用低密度功能芯片对动物基因资源在全基因组范围进行评估, 对促进远缘基因的鉴定和交流、国内外科研工作的合作、国内外资源的交流、优良品种的培育等都具有积极作用。利用基因芯片技术在全基因组水平评估种质资源及基因资源, 有利于全面了解物种资源情况, 利于不同科研机构、国家/地区资源的交流及共享, 对未来更好地进行优良品种培育具有积极作用。

表 2 基因芯片与测序基因分型比较

Table 2 Comparison of gene chip and sequencing genotype

| 属性 Attribute | 基因芯片 Gene chip | 测序基因分型 Sequencing genotype |
|---------------------------|--|---|
| DNA 质量 Quality of DNA | 对质量和数据有严格要求, 200~3 000 ng | 100~200 ng |
| 标记选择 Tag selection | 无确定偏差 | 通过多个品系的重测序, 可减少确定偏差; 能够靶定特定染色体 |
| 应用范围 Range of application | 有参考基因组或无参考基因组都可; 不能靶定特定染色体; 不能控制不同区域的 read 深度; 数据可能因测序深度而丢失 | 需要有参考基因组; 可用于 NGS 验证/GWAS/GS/QTL; 不能用于新 SNP 的发现; 芯片位点数据不会丢失 |
| 基因组复杂度 Genome complexity | 增加遗传差异可能会影响酶切位点, 并增加数据丢失的频率; 重复可能增加组装的难度 | 没有酶切位点的影响; 能够检测大型基因组或高杂合型基因组的标记 |
| 多倍体检测 Polyploid detection | 多倍体、高杂合型和 CNV 需要深度测序; 纯合 6 倍, 杂合 13.5 倍 ($P \geq 0.9975$); 组装可能是嵌合 | 可分析异源多倍体数据 |
| 周期 Cycle | 20~24 周 | 4~5 d |
| 数据分析 Data analysis | 较难 | 容易 |

3.1.1 重复性高, 数据质量高 研究显示, 在实际育种工作中, 通过测序获得的高密度数据并未提高育种的准确度, 这主要是因为新增加的这些位点因测序深度、数据筛选等环节并未能捕获到频率较低的等位基因变异, 或因过度看重测序群体而导致部分 SNP 位点遗漏。由于芯片所含位点多为群体内具有多态性的可检测位点, 通常在待测群体中具有较高的检出率, 而通过测序所获得的数据中, 非多态性的位点所占比例可能会大于多态性位点, 即数据冗余。当鉴定和选育特定优良性状位点时, 利用特定标记或基因芯片即可, 无需检测其他位点, 此时

3 芯片技术的优势及面临的挑战

近年来随着科学技术的快速发展, 科研及育种工作已进入高通量检测时代, 除了利用基因芯片获得基因分型, 还可以利用测序技术。与测序相比, 芯片技术拥有它特有的优势, 但同时, 农业芯片的应用与推广也面临很多挑战。

3.1 芯片技术在畜禽育种中的优势

测序技术从最初的 Sanger 测序, 到现在的下一代测序或二代测序 (next generation sequencing, NGS), 检测成本大幅度降低, 检测周期大幅度缩短, 使其逐渐在科研领域受到青睐^[36]。NGS 技术可以用来对任何一种生物的 DNA 进行测序, 也可通过多种方法应用于小的、靶向的区域, 提供所需信息或大量未知但具有重要价值的信息。与现阶段的二代测序技术相比, 基因芯片技术在育种研究和应用中逐渐得到广泛认可, 并具有一定的优势 (表 2)。

如果选用测序则会使结果存在很大比例的“无价值或冗余数据”^[37]。

3.1.2 检测快速、高通量、检测方便

在芯片检测中, 检测后系统显示出的基因数据或基因分型清晰明了, 无需数据去污环节, 显示出其快速简单的优势, 而测序数据处理中的去污过程对最终数据结果及分析也将有重要影响。Jami 等^[38] 利用 Illumina Bovine HD BeadChip 芯片和 ddRAD 测序 (double digest restriction site-associated DNA) 两种方法对野生曲角羚羊进行群体谱系分析, 结果显示, 两种方法均能获得较好结果, 但测序显示的 SNP 差异数比

芯片检测所获得的差异要少,即部分 SNP 位点通过测序未能检测到,因而在数据处理和分析过程中测序比芯片检测更需要谨慎,尤其是采用评估 AS(allele sharing)作为数据分析指标时,同时可以对大量样本进行高通量检测。

3.1.3 可进行位点精选、定制芯片 在育种工作中,芯片的 SNP 位点选择很重要,如何选择位点来提高与目标优良性状的关联,将影响利用芯片进行基因育种的准确度。若位点选择准确,且具有代表性,则在实际育种工作中,芯片检测技术将以检测便捷快速、数据处理简单、成本较低等特点优于现阶段的二代测序^[37]。虽然芯片和测序两种方法都可以用于群体结构分析和分子育种,但在工作中具体选用哪种方法需要根据被研究对象而定。如果被研究对象已有相关的芯片产品,则推荐使用芯片检测方法,因为其检测方法更简单、数据处理平台更完善;如果被研究对象没有相应芯片产品,则可以使用测序技术,但后续数据的处理和分析对技术人员的能力要求较高。此外,如果检测样本数量足够(如几百个),则可以先对部分样本进行测序,根据测序数据定制芯片,然后对剩余样本进行芯片检测,不过定制芯片一般需要周期较长,但一旦芯片定制完成,则会便于后续其他样本的检测及分析工作。

3.2 畜禽育种芯片在开发及推广中的挑战

目前用于畜禽育种领域的国内外基因芯片有多款,但是在国内外芯片的开发中,科研单位或芯片公司面临国内外参考材料差异较大、标记聚堆、部分特定位点无法设计探针、芯片位点间 Gap 过大、低密度功能芯片位点难以选择等问题。为了解决上述问题,近几年中国已开发多款基于中国品种的定制芯片,如猪 51K 中芯一号、90K 肉鸡京芯一号、55K 蛋鸡凤芯一号等 SNP 育种芯片,这些基因芯片极大地促进了群体结构分析、品种培育等农业育种工作,大大提高了中国农业育种的准确度和效率。

在芯片的实际应用和推广过程中也面临着一些问题。首先,如何选择最佳育种值的计算方法、育种模型的选择与构建等将直接影响育种值估计的准确性。林彭^[39]对全基因组选择中常用的几种方法(Bayes A、Bayes C π 、Bayes LASSO、Bayes Ridge、SVM)进行了比较,认为 Bayes A 和 Bayes C π 具有较好的预测相关性。李宏伟等^[40]对最小二乘法、GBLUP 法和贝叶斯法等多个计算方法进行了详细的比较和阐述,认为 Bayes A 和 Bayes B 具有较高的准确性。第二,芯片的参考群体差异对育种值估

计的影响较大。因此针对某一国家或区域收集样本定制芯片,并将其利用在该国家或区域的育种工作中仍有待开发。此外,大规模进行育种工作的公司和农场仍然面临芯片及检测仪器价格高、检测及数据分析技术落后等问题,芯片的可重复利用性也有待研究,因此,目前公司或农场更倾向于使用价格低、功能明显的低密度功能芯片。

4 展望

随着分子生物学技术的发展及 SNP 芯片数据的研究不断深入,芯片和测序技术逐渐在育种选择中得到应用,逐渐代替了传统最佳线性无偏估计(best linear unbiased prediction, BLUP)和标记辅助选择(MAS)选育方法,SNP 芯片针对性强、芯片密度和检测效率高,分析方法科学到位,进而实现与性状相关的 SNP 位点的精准定位。利用基因芯片可以进行基因分型、群体分析、进化分析等,可促进动物品种遗传基础的探索及亚种群间差异的比较,缩短育种周期,提高育种效率。同时,在全基因组水平上的基因分型和单倍型的分析与挖掘,为动物分子育种积累素材,使得多个优良基因的选择、GWAS 及 GS 等全基因组范围的选择在农业尤其是动物育种工作中成为现实,促进了优良基因聚合和品种改良,对农业育种工作具有重要的指导意义。

基因育种芯片的开发主要包括用于前期优良性状挖掘的高密度芯片和用于后期商品化的低密度芯片。随着畜禽育种芯片技术的发展,用于研究基因信号通路、基因快速克隆等的基因表达芯片^[41]也有待于开发和创新,如鉴定与动物疾病相关的基因甲基化芯片等^[42],虽然已有部分商业化产品,如 Chicken Gene 1.0 ST Array、850K Infinium methylation EPIC Beadchip 等,但涉及物种数量小,极大部分物种的相应产品仍有待开发。基因表达芯片及甲基化芯片将极大地推动基因领域的深入研究,促进优良性状基因的有效利用。此外,芯片在印迹基因的研究中也拥有广阔的应用前景,如周泉勇^[43]利用猪 Affymetrix 芯片对猪妊娠后期胎盘转录谱进行比较,最终发现并通过实时荧光定量 PCR 等方法验证获得了 5 个候选印迹基因。因此,未来生物芯片的制造很可能会出现与现有芯片产品不同的新方向、新定位。

21 世纪是生命科学与信息科学的世纪,芯片技术是将这两种科学融为一体的技术。基因芯片技术在畜禽育种研究和育种应用中已取得了丰硕的成果,还将会不断完善,相信在不久的将来,芯片技术

可能会像目前的 PCR 等实验室常规技术一样,成为畜禽育种中的基本工具,推动畜牧业的发展。

参考文献(References):

- [1] FORDOR S P, READ J L, PIRNING M C. Light-directed spatially addressable parallel chemical synthesis[J]. *Science*, 1991, 251(4995): 762-773.
- [2] FORDOR S P, RAVA R P, HUANG X C. Multiplexed biochemical assays with biological chip[J]. *Nature*, 1993, 364: 555-556.
- [3] 张于光, 李迪强, 肖启明, 等. 基因芯片及其在环境微生物研究中的应用[J]. *微生物学报*, 2004, 44(3): 406-410.
ZHANG Y G, LI D Q, XIAO Q M, et al. Microarrays and their application to environmental microorganisms[J]. *Acta Microbiologica Sinica*, 2004, 44(3): 406-410. (in Chinese)
- [4] 邢婉丽, 程京. 生物技术芯片[M]. 北京: 清华大学出版社, 2004.
XING W L, CHENG J. Biological Chip Technology[M]. Beijing: Tsinghua University Press, 2004. (in Chinese)
- [5] 于文静. 我国诞生全球首款水稻全基因组育种芯片[J]. *北京农业*, 2012, 17: 5.
YU W J. The world's first rice genome breeding chip was born in China[J]. *Beijing Agriculture*, 2012, 17: 5. (in Chinese)
- [6] DAVEY J W, HOHENLOHE P A, ETTER P D, et al. Genome-wide genetic marker discovery and genotyping using next-generation sequencing[J]. *Nature Reviews Genetics*, 2011, 12(7): 499-510.
- [7] GUPTA P K, RUSTGI S, MIR R R. Array-based high-throughput DNA markers for crop improvement[J]. *Hereditas*, 2008, 101(1): 5-18.
- [8] TOSSER-KLOPP G, BARDOU P, BOUCHEZ O, et al. Design and characterization of a 52K SNP chip for goats[J]. *PLoS One*, 2014, 9(1): e86227.
- [9] 宣苏哲. 我国研制出首款鸡 55K SNP 芯片[J]. *家禽科学*, 2017, 19: 12.
XUAN S Z. China developed the first chicken 55K SNP chip[J]. *Poultry Science*, 2017, 19: 12. (in Chinese)
- [10] DE ROOS A P, HAYES B J, SPELMAN R J, et al. Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus cattle[J]. *Genetics*, 2008, 179(3): 1503-1512.
- [11] DE LOS C G, VAZQUEZ A I, FERNANDO R, et al. Prediction of complex human traits using the genomic best linear unbiased predictor[J]. *PLoS Genetics*, 2013, 9(7): e1003608.
- [12] CAÑAS-ÁLVAREZ J, GONZALEZ-RODRIGUEZ A, MUNILLA S, et al. Genetic diversity and divergence among Spanish beef cattle breeds assessed by a bovine high-density SNP chip[J]. *Journal of Animal Science*, 2015, 93(11): 5164-5174.
- [13] DA SILVA V H, LAEIN V N, BOSSE M, et al. CNVs are associated with genomic architecture in a songbird[J]. *BMC Genomics*, 2018, 19(1): 195.
- [14] 袁泽湖, 王慧华, 胡师金, 等. 利用 50K 芯片数据分析中国 11 个地方绵羊群体的遗传结构[J]. *畜牧兽医学报*, 2016, 47(5): 899-908.
YUAN Z H, WANG H H, HU S J, et al. Population structure analysis of Chinese indigenous sheep by 50K chip data[J]. *Acta Veterinaria et Zootechnica Sinica*, 2016, 47(5): 899-908. (in Chinese)
- [15] 兰蓉, 朱兰, 杨红远, 等. 基于 SNP 芯片的云上黑山羊遗传结构分析[J]. *中国畜牧兽医*, 2019, 46(2): 480-488.
LAN R, ZHU L, YANG H Y, et al. Analysis of genetic structure of Yunshang Black goat based on SNP chip[J]. *China Animal Husbandry & Veterinary Medicine*, 2019, 46(2): 480-488. (in Chinese)
- [16] MARTIN P, PALHIÈRE I, TOSSER K G, et al. Heritability and genome-wide association mapping for supernumerary teats in French Alpine and Saanen dairy goats[J]. *Journal of Dairy Science*, 2016, 99(11): 8891-8900.
- [17] 兰蓉, 朱兰, 姚新荣, 等. 山羊产羔数全基因组关联分析[J]. *畜牧兽医学报*, 2015, 46(4): 549-554.
LAN R, ZHU L, YAO X R, et al. Genome-wide association analysis of lambing number in goat[J]. *Acta Veterinaria et Zootechnica Sinica*, 2015, 46(4): 549-554. (in Chinese)
- [18] YI G, QU L, CHEN S, et al. Genome-wide copy number profiling using high-density SNP array in chickens[J]. *Animal Genetics*, 2015, 46(2): 148-157.
- [19] 刘澳星, 郭刚, 王雅春, 等. 中国荷斯坦牛初产日龄遗传评估及全基因组关联分析[J]. *畜牧兽医学报*, 2015, 46(3): 373-381.
LIU A X, GUO G, WANG Y C, et al. Genetic analysis and genome wide association studies for age at first calving in Chinese Holsteins[J]. *Acta Veterinaria et Zootechnica Sinica*, 2015, 46(3): 373-381. (in Chinese)
- [20] DAHLGREN S, ZIENER M L, LINGAAS F. A genome-wide association study identifies a region strongly associated with symmetrical onychomadesis on chromosome 12 in dogs[J]. *Animal Genetics*, 2016, 47(6): 708-716.
- [21] ZHU B, LI Q H, LIU R R, et al. Genome-wide association study of H/L traits in chicken[J]. *Animals(Basel)*, 2019, 9(5): E260.
- [22] JUNQUEIRA V S, CARDOSO F F, OLIVEIRA M M,

- et al. Use of molecular markers to improve relationship information in the genetic evaluation of beef cattle tick resistance under pedigree-based models[J]. *Journal of Animal Breeding and Genetics*, 2017, 134(1): 14-26.
- [23] 田 佳,李委奇,刘丽元,等.宁夏地区青年荷斯坦母牛基因组育种值分析[J]. *中国奶牛*, 2018, 2: 15-19.
TIAN J, LI W Q, LIU L Y, et al. Analysis on genome breeding value of young Holstein cows in Ningxia area[J]. *China Dairy Cattle*, 2018, 2: 15-19. (in Chinese)
- [24] MARCHINI J, HOWIE B. Genotype imputation for genome-wide association studies[J]. *Nature Reviews Genetics*, 2010, 11(7): 499-511.
- [25] YE S P, YUAN X L, LIN X R, et al. Imputation from SNP chip to sequence: A case study in a Chinese indigenous chicken population[J]. *Journal of Livestock and Biotechnology*, 2018, 2: 294-305.
- [26] VAN DEN BERG I, GULDBRANDTSEN B, HOZE C, et al. Across breed QTL detection and genomic prediction in French and Danish dairy cattle breeds[A]. 10th World Congress of Genetics Applied to Livestock Production[C]. 2014.
- [27] VANRADEN P M, O'CONNELL J R, WIGGANS G R, et al. Genomic evaluations with many more genotypes[J]. *Genetics Selection Evolution*, 2011, 43: 10.
- [28] MEUWISSEN T, GODDARD M. Accurate prediction of genetic values for complex traits by whole-genome re-sequencing[J]. *Genetics*, 2010, 185(2): 623-631.
- [29] NAJ A C. Genotype imputation in genome-wide association studies[J]. *Current Protocols in Human Genetics*, 2019, 102(1): e84.
- [30] HULSEGG B, CALUS M P, WINDIG J J, et al. Selection of SNP from 50K and 777K arrays to predict breed of origin in cattle[J]. *Journal of Animal Science*, 2013, 91(11): 5128-5134.
- [31] GUNIA M, SAINTILAN R, VENOT E, et al. Genomic prediction in French Charolais beef cattle using high-density single nucleotide polymorphism markers [J]. *Journal of Animal Science*, 2014, 92(8): 3258-3269.
- [32] BOLORMAA S, GORE K, HAYAYES B J, et al. Design of a low-density SNP chip for the main Australian sheep breeds and its effect on imputation and genomic prediction accuracy[J]. *Animal Genetics*, 2015, 46(5): 544-556.
- [33] OGAWA S, MATSUDA H, TANIGUCHI Y, et al. Accuracy of imputation of single nucleotide polymorphism marker genotypes from low-density panels in Japanese Black cattle[J]. *Journal of Animal Science*, 2015, 87(1): 3-12.
- [34] AUVUVRAY B, MCEWAN J C, NEWMAN S A, et al. Genomic prediction of breeding values in the New Zealand sheep industry using a 50K SNP chip[J]. *Journal of Animal Science*, 2014, 92(10): 4375-4389.
- [35] LOPES F B, WU X L, LI H, et al. Improving accuracy of genomic prediction in Brangus cattle by adding animals with imputed low-density SNP genotypes[J]. *Journal of Animal Breeding and Genetics*, 2018, 135(1): 14-27.
- [36] LIACA V. Sequencing Technologies and Their Use in Plant Biotechnology and Breeding[M]. DNA Sequencing: Methods and Applications. In Tech, 2012.
- [37] PEREZ-ENCISO M, RINCON J C, LEGARRA A, et al. Sequence-vs. chip-assisted genomic selection: Accurate biological information is advised[J]. *Genetics Selection Evolution*, 2015, 47: 43.
- [38] JAMI A I, ANDREA S P, ASAKO Y N, et al. Applying SNP-derived molecular coancestry estimates to captive breeding programs[J]. *Journal of Heredity*, 2016, 107(5): 403-412.
- [39] 林 彭. 几种基因组评估方法选择效率的比较研究[D]. 雅安: 四川农业大学, 2013.
LIN P. Comparison of selection efficiency in several genomic evaluating methods[D]. Ya'an: Sichuan Agricultural University, 2013. (in Chinese)
- [40] 李宏伟, 王瑞军, 王志英, 等. 家畜基因组选择研究进展[J]. *遗传*, 2017, 39(5): 377-387.
LI H W, WANG R J, WANG Z Y, et al. The research progress of genomic selection in livestock[J]. *Heredity*, 2017, 39(5): 377-387. (in Chinese)
- [41] 周立国. 水稻水分胁迫相关基因克隆及功能验证[D]. 武汉: 华中农业大学, 2010.
ZHOU L G. Cloning and functional characterization of genes related to water stress resistance of rice[D]. Wuhan: Huazhong Agricultural University, 2010. (in Chinese)
- [42] 冯 政, 刘贵生, 武 华, 等. 应用基因芯片技术筛选猪抗病相关基因[A]. 第十二次全国畜禽遗传标记研讨会论文集[C]. 2010.
FENG Z, LIU G S, WU H, et al. Gene chip technology was used to screen genes related to disease resistance in pigs[A]. Proceedings of the 12th National Symposium on Genetic Markers of Livestock and Poultry[C]. 2010. (in Chinese)
- [43] 周泉勇. 大白猪和二花脸猪妊娠后期胎盘转录谱比较及印记基因鉴定研究[D]. 武汉: 华中农业大学, 2009.
ZHOU Q Y. Transcriptional analysis of placentas from large white pigs and Erhualian pigs on later gestation and imprinting detection[D]. Wuhan: Huazhong Agricultural University, 2009. (in Chinese)

(责任编辑 晋大鹏)