

基于测序技术的畜禽基因组学研究进展

梁素芸, 周正奎, 侯水生

中国农业科学院北京畜牧兽医研究所, 农业部动物遗传育种与繁殖(家禽)重点实验室, 北京 100193

摘要: 人类通过数千年的驯化和近代以来有计划性的育种, 形成了当今多样化的畜禽品种, 从而提供丰富的动物源性蛋白满足人类需求。在过去的 100 年里, 数量遗传学应用于动物育种领域引发了畜禽育种技术的革命, 但畜禽机体遗传发育体系相当复杂, 一些性状仍然难以通过基于系谱的育种值进行高效选育, 遗传潜能尚未充分发掘。人类基因组计划带来的理念和技术极大促进了畜禽基因组学的发展, 使得人们可以从全基因组水平精准定位功能变异, 挖掘功能元件的生物学意义, 为畜禽分子设计育种提供重要的理论基础。本文对近 10 年来猪(*Sus scrofa*)、牛(*Bos taurus*)、牦牛(*Bos grunniens*)、山羊(*Capra hircus*)、绵羊(*Ovis aries*)、鸡(*Gallus gallus*)、鸭(*Anas platyrhynchos*)和鹅(*Anser cygnoides*)等主要畜禽的基因组学研究进展进行综述, 分别从参考基因组构建和群体基因组学分析两个方面进行论述, 并对畜禽基因组未来的研究工作进行了展望。

关键词: 畜禽; 参考基因组; 重测序

The research progress of farm animal genomics based on sequencing technologies

Suyun Liang, Zhengkui Zhou, Shuisheng Hou

Key Laboratory of Animal Genetics Breeding and Reproduction (Poultry), Ministry of Agriculture, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 100193, China

Abstract: Various farm animal breeds have been domesticated and bred for thousands years, and they provide adequate animal-derived proteins to meet the human nutrition requirement. Although quantitative genetics was applied in animal breeding, which launched a technological revolution in the past century, a number of complex traits remain difficult to be selected based on pedigree derived breeding, due to complicated animal genetics and development mechanisms. Farm animal's genetic potential hasn't yet to be fully exploited. The concept and technology from the Human Genome Project have greatly promoted farm animal genomic researches. It is possible to fine map the causal variations at the whole genome level and then exploit their biological functions, thus providing the theoretical basis for molecular designed breeding. In this review, we summarize the genomics research progress of main

收稿日期: 2016-09-26; 修回日期: 2017-02-13

基金项目: 国家自然科学基金面上项目(编号: 31672410), 国家科技支撑计划项目(编号: 2015BAD03B06)和国家水禽产业技术体系项目(编号: CARS-43-01)资助[Supported by the National Natural Science Foundation of China (No. 31672410), National Scientific Supporting Projects of China (No. 2015BAD03B06) and China Agriculture Research System of Waterfowl (No. CARS-43-01)]

作者简介: 梁素芸, 硕士研究生, 专业: 水禽遗传育种。

通讯作者: 周正奎, 博士, 副研究员, 研究方向: 水禽基因组学与分子育种。E-mail: zhouzhengkui@caas.cn

侯水生, 博士, 研究员, 研究方向: 水禽新品种选育与营养调控。E-mail: houss@263.net

DOI: 10.16288/j.yczs.16-328

网络出版时间: 2017/3/15 13:04:53

URI: <http://kns.cnki.net/kcms/detail/11.1913.R.20170315.1304.008.html>

farm animals during the past decade, including pigs, cattle, yaks, goats, sheep, chickens, ducks and geese. We focus on the reference genome sequencing and follow-up population-level genomic studies based on high throughput resequencing technologies, and meanwhile envision the future work of farm animal genomics.

Keywords: farm animal; reference genome; resequencing

1990年,美国国家人类基因组研究中心联合国际合作机构发起人类基因组计划(Human Genome Project, HGP)^[1],这个计划耗时13年绘制出人类基因组30亿对碱基,被认为是最具历史意义的科学成就之一。它不但极大地加速了生物医学研究的进程,还引领了全新的生命科学研究模式。作为生物领域首个大规模项目,人类基因组计划为众多基于联盟协作的研究开创了新的局面。美国又先后启动了“国际人类基因组单体型图计划”(The International HapMap Project, HapMap)^[2]、“DNA元件百科全书”计划(Encyclopedia of DNA Elements, ENCODE)^[3];千人基因组计划(1000 Genome Project)^[4],以及近期受到广泛关注的“精准医疗计划”(The Precision Medicine Initiative)。这些计划的实施与成果为人类的健康带来了诸多福音。

生命科学研究领域受到人类基因组计划的启示与引领,目前已有数百种动植物完成了参考基因组序列测定。而鸡、牛、猪等畜禽作为人类重要的食物来源和疾病研究模型,其参考基因组的测定将会加速家养动物的培育进程,为全基因组范围内探索遗传变异和表型多样性的关系提供详尽信息。近期,多国科学家联合倡导并启动了“动物基因组功能注释计划”(The Functional Annotation of Animal Ge-

nomes Project, FAANG)^[5],该计划致力于通过国际合作来加快绘制整个家养动物基因组元件的综合性功能图谱。这些基因组领域的基础性工作是基于分子育种策略培育畜禽品种的理论基础,具有重要的意义。

目前,多个畜禽的基因组计划陆续完成,第一个完成基因组序列测定的是红原鸡(*Gallus gallus*)^[6],其后海福特牛(*Bos taurus*)^[7]、杜洛克猪(*Sus scrofa*)^[8]、云南黑山羊(*Capra hircus*)^[9]、特克塞尔羊(*Ovis aries*)^[10]、青海牦牛(*Bos grunniens*)、北京鸭(*Anas platyrhynchos*)^[11]、浙东白鹅(*Anser cygnoides*)^[12]等也相继完成基因组测序(表1)。同时,在这些参考基因组的基础上,利用群体全基因组重测序或转录组测序技术又陆续完成了一系列重要遗传学问题的解析工作(表2)。本文以物种分类,按参考基因组构建和群体基因组分析与应用两个层次进行总结和论述。

1 鸡

鸡是人类饲养最普遍的家禽,全球鸡肉消费量占肉类的比重最高,更重要的是鸡基因组大小仅为家畜的1/3,因此成为畜禽中第一个测定全基因组序列的物种。鸡全基因组草图的公布,极大地推动了其他经济动物和模式生物的研究工作。

表1 畜禽参考基因组基本信息

Table 1 The information of farm animal reference genomes

物种	测序品种	发表时间	基因数目	基因组大小(Gb)	测序方法与深度	Contig/Scaffold N50
鸡	红原鸡 ^[6]	2004.12	20 000~23 000	1.05	BAC, 6.6×	36 kb/7067 kb
牛	海福特牛 ^[7]	2009.04	22 000	2.87	BAC, 7×	48.7 kb/1.9 Mb
	青海牦牛 ^[13]	2012.07	22 282	2.66	Shotgun, 65×	20.4 kb/1.4 Mb
猪	杜洛克猪 ^[8]	2012.11	21 640	2.60	BAC, Shotgun, 44×	80.7 kb/637 kb
	五指山猪 ^[14]	2012.11	20 326	2.60	Shotgun, 78×	23.5 kb/5.4 Mb
山羊	云南黑山羊 ^[9]	2012.12	22 175	2.66	Shotgun, 65.6×	18.7 kb/2.21 Mb
绵羊	特克塞尔羊 ^[10]	2014.06	20 908	2.61	Shotgun, 75×	39.9 kb/2.23 Mb
鸭	北京鸭 ^[11]	2013.06	19 144	1.10	Shotgun, 64×	26 kb/1.2 Mb
鹅	浙东白鹅 ^[12]	2015.05	16 150	1.12	Shotgun, 130×	27.5 kb/5.2 Mb

表 2 主要畜禽群体基因组学研究进展

Table 2 The research progress of main farm animal population genomics

物种	材料(数量)	主要结论	参考文献
鸡	艾维因肉鸡(1)、来航蛋鸡(1)、丝羽乌骨鸡(1)#	绘制了鸡高密度全基因组遗传变异图谱。	[15]
	蛋鸡: 白来航 A 系(11)、白来航 B 系(8)、OS 肥大系(10)、罗得岛红鸡(8); 肉鸡: Ross308 商业品系(10)、Hubbard ISA 商业品系(10)、高、低生长系(11); 野鸡: 红原鸡(8)*	<i>TSHR</i> 是典型的驯化基因, 突破了鸡繁殖力的瓶颈。	[16]
	林芝藏鸡(1, <i>de novo</i>); 藏鸡(10)、红原鸡(5)、云南本地鸡(8)、西双版纳斗鸡(8)、罗曼蛋鸡(1)	藏鸡有两个支系, 钙信号转导途径相关基因可能参与低氧的适应性进化过程。	[17]
	红原鸡(5)、云南本地鸡(8)	<i>VIT</i> 是重要的视觉功能基因之一, 家鸡视力下降是由于该基因在驯化过程中受到正选择而非选择放松。	[18]
牛	荷斯坦牛(1)、安格斯牛(1)、婆罗门牛(1)、利木赞牛(1)、娟姗牛(1)、挪威红牛(1)#	构建了牛 SNP 数据库, 应用于 19 个品种共 497 头牛的遗传结构分析, 发现牛的有效群体含量在近期急剧下降。	[19]
	黑白花荷斯坦牛(125)、红白花荷斯坦牛(4)、弗莱维赫牛(43)、安格斯牛(47)、娟姗牛(15)	定位了牛胚胎致死、被毛卷曲和乳脂率等性状的基因和突变位点。	[20]
	弗莱维赫牛(43)	鉴定了 106 个受选择的区域, 包含与毛色、性情和感官相关的驯化基因, 毛色是主要驯化性状之一。	[21]
牦牛	野牦牛(13)、家养牦牛(48)、天祝白牦牛(11)	鉴定了与行为和温顺习性等相关的 209 个驯化基因, 推测牦牛自 7300 年前起由游牧民驯化。	[22]
猪	大白猪(14)、长白猪(5)、汉普夏猪(2)、皮特兰猪(5)、杜洛克猪(4)、眉山猪(4)、香猪(2)、姜曲海猪(1)、欧洲野猪(6)、亚洲野猪(5)、外群(7)**	<i>NR6A1</i> 、 <i>PLAG</i> 和 <i>LCORL</i> 基因座变异引起椎骨数的增加; <i>KIT</i> 基因复杂结构变异形成丰富的毛色表型。	[23]
	稻城藏猪(1, <i>de novo</i>); 甘孜藏猪(5)、迪庆藏猪(5)、林芝藏猪(5)、日喀则藏猪(5)、甘南藏猪(5)、阿坝藏猪(5)、金华猪(3)、内江猪(3)、盆州猪(3)、乌金猪(3)、雅安猪(3)、野猪(3)	揭示了藏猪高海拔适应性机制, 受选择基因主要参与缺氧、嗅觉、能量代谢和药物反应等过程; 揭示了驯化过程中家猪增加唾液分泌的遗传基础。	[24]
	在文献[23]基础上增加: 丹麦野猪(10)、法国野猪(1)、瑞士野猪(1)、意大利野猪(3)、希腊野猪(2)、日本野猪(1)、中国野猪(7)、苏门答腊野猪(2)	欧亚猪种存在广泛的基因渗入, 渗入的单倍型主要与肉质、生长以及繁殖性能有关。 <i>AHR</i> 基因是典型的亚洲猪种渗入基因, 增加了欧洲猪的产仔数。	[25]
	欧洲猪品种 26 个; 欧洲野猪品种 10 个; 亚洲猪品种 7 个; 亚洲野猪品种 5 个; 外群 1 个**、***	欧亚猪种存在广泛的基因流动, 驯化与杂交是持续交替进行的, 形成基因组上的“驯化孤岛”。	[26]
	巴马香猪(6)、二花脸猪(6)、河套猪(6)、莱芜猪(6)、陆川猪(6)、民猪(6)、五指山猪(6); 高原: 甘肃猪(4)、四川猪(6)、藏猪(6)、云南猪(6)、野猪(6)	新鉴定了 2100 万 SNP 位点; X 染色体上南北猪种存在两种低重组单倍型长达 14 Mb, 是环境适应性自然选择的强有力证据。	[27]
	青藏高原: 那曲藏羊(4)、昌都藏羊(4)、日喀则藏羊(4)、林芝藏羊(4)、贵德黑羊(4)、甘孜羊(4)、岷县黑羊(5); 云贵高原: 腾冲羊(5)、迪庆羊(5)、石屏羊(5)、威宁羊(5); 平原地区: 罗布羊(4)、巴尔楚克羊(5)、阿勒泰羊(1)、巴音布鲁克羊(1)、哈萨克羊(1)、和田羊(1)、塔什库尔干羊(1)、乌珠穆沁羊(1)、乌冉克羊(1)、苏尼特羊(1)、呼伦贝尔羊(1)、洼地绵羊(5)、湖羊(5); 野羊: 欧洲摩弗伦羊(1)、盘羊(1)、北山羊(1)	揭示了绵羊对高原、干旱和湿润等极端环境的适应性进化机制。描绘了中国地方绵羊的群体遗传结构和种群历史动态。	[28]

注: 该表除 Sanger 法测序构建 SNP 数据库的研究(#)外, 所列研究群体 > 10 个体, 测序深度 > 5×。

*表示均为混样测序; **表示因篇幅原因未能逐一列出, 参见原文; ***表示混样测序样本未列出。

1.1 参考基因组

国际鸡基因组测序联盟(International Chicken Genome Sequencing Consortium)于 2004 年公布了红

原鸡的全基因组序列草图^[6], 他们选择红原鸡近交系(UCD001)的一只母鸡为材料, 分别构建了质粒、fosmid 和 BAC 末端文库, 基于一代测序技术(Sanger 法)采用全基因组鸟枪法绘制出基因组草图, 测序深

度为 $6.6\times$, Contig N50 为 36 kb, Scaffold N50 为 7067 kb, 组装后的基因组草图全长为 1.05 Gb, 其中的 933 Mb 能够根据遗传图谱信息分配到对应的染色体上。组装质量检测表明, 与测序个体同来源的 38 个 BAC 克隆序列, 98% 能够比对到该基因组上; NCBI 数据库中约 485 000 条鸡的 EST 序列有 96% 的能够比对到基因组上, 说明组装质量已达到较为理想的水平。基因注释采用了 Ensembl 系统和 Twinscan37、SGP-2 法相结合的方法, 共注释出 20 000~23 000 个基因。

1.2 基因组后续研究

鸡参考基因组序列的绘制为构建品种间遗传变异图谱奠定了基础, 红原鸡基因组序列公布的同时, 以中国学者为首的国际鸡多态性图谱联盟(International Chicken Polymorphism Map Consortium)^[15] 采用基于第一代测序技术的重测序策略, 对肉鸡、蛋鸡和乌骨鸡进行全基因组测序, 与红原鸡基因组序列进行比对后绘制了包含 280 万个单核苷酸多态性(SNP)变异位点的遗传变异图谱。研究人员通过比较原鸡与家鸡之间以及家鸡种属间 SNP 位点, 发现家鸡和野生红原鸡之间、鸡种之间以及鸡种之内的变异程度相当接近, 变异密度为 5 SNP/kb, 且变异位点大部分在 5000~10 000 年前就已经形成, 表明驯化并没有明显降低遗传多样性。当时发掘的全基因组 SNP 最重要的作用是丰富了数量性状基因座(QTL)定位的遗传标记, 相比于之前有限的变异位点, 这些覆盖全基因组高密度 SNP 提升了复杂性状在基因组上的定位工作的便利性, 研究者可以构建更加精确的单倍型进行功能基因精细定位。例如 Leif Andersson 团队精确定位了一系列鸡冠表型调控基因, 并且明确了这些变异的调控机制: *SOX5* 基因的 1 个内含子拷贝数变异形成豆冠^[29]; 7 号染色体一段 7.4 Mb 序列反转引起 *MNR2* 同源结构域蛋白基因异位表达形成玫瑰冠^[30]; *EOMES* 基因上游 200 kb 调控区 20 kb 片段串联重复则形成双冠^[31]。该团队还定位了鸡色素沉积相关的基因, 如黄皮肤基因 *BCDO2*^[32], 该基因顺式元件或组织特异调控元件突变均导致类胡萝卜素在皮肤沉积; 乌骨鸡独有的黑色素过度沉积性状是由于 20 号染色体 10.7~11.5 Mb 处两个相距 400 kb 的区段分别发生反向重复, 导致

区段中的色素调控基因 *EDN3* 表达失控, 黑色素在体内过度沉积^[33]。中国学者在鸡重要性状研究方面也取得了多项重要进展, 精确定位并解析了绿壳蛋和丝羽等性状的调控基因和遗传机制。如 Wang 等^[34] 发现 EVA-HP 转座子插入 *SLCO1B3* 基因的 5'侧翼调控区导致了在胆壳腺大量分泌胆绿素, 沉积于蛋壳形成绿壳蛋; Feng 等^[35] 发现 *PDSS2* 基因一个顺式元件突变导致鸡丝羽表型的出现; Jin 等^[36] 鉴定了 7 号染色体内含 *IHH* 基因的 11.89 kb 大片段缺失变异, 纯合和杂合型分别导致鸡胚胎期致死和匍匐性状(胫短身矮、翅膀短小)。

随着测序技术的进步, Rubin 等^[16] 采用二代测序技术, 以红原鸡、罗得岛红鸡(Rhode Island Red)、肉鸡和蛋鸡的多个经典品系共 9 个混合样本进行全基因组重测序, 对人工选择下鸡表型分化的基因组变异机制进行了研究, 得到了若干个受选择片段, 在这些片段上, 与提高存活能力相关的有利基因变异相对于其他等位基因在频率上有所增加, 其中促甲状腺激素受体(*TSHR*)基因在 6000 年前开始的人工驯化过程中受到强烈选择, 推断 *TSHR* 基因在脊椎动物的代谢调节和繁殖的光周期调控过程中起重要作用。该研究还对人工选择研究的经典素材——鸡高低生长系^[37]进行了研究, 通过比较发现一个与体重相关的 QTL 区域, 研究者发现并推测在此区域内 *SH3RF2* 基因编码序列的大片段缺失是造成鸡异速生长的主效位点, 同时还发现肉鸡群体中鉴定的受选择性区间内基因多与生长、食欲和代谢调控途径相关。对于上述结论, Flink 等^[38] 根据来源于不同时期鸡古 DNA 与现代家养鸡的比较, 提出了修正性证据, 研究发现此前鉴定的重要驯化基因 *TSHR* 基因频率在群体中由平衡状态到现代鸡种接近固定的过程是在最近 500 年内发生的, 因此他们认为利用现存群体基因组数据来准确推演历史上种群变化存在一定风险。Wang 等^[18] 基于家鸡和红原鸡群体基因组数据和视网膜、脑部等转录组数据, 发现大量视觉相关基因在家鸡驯化过程中受到正选择作用, 而非选择压力放松, 并在模式动物上验证了正选择基因 *VIT* 在视觉上的重要作用。

藏鸡是研究禽类高原适应性的理想模型。Wang 等^[17] 对藏鸡群体重测序研究发现可以大致分为两个

支系,说明藏鸡可能有多个起源;通过比较藏鸡和平原鸡,发现与钙信号转导途径相关的候选基因受到选择,它们可能参与低氧适应性的过程,并且这些基因经历了定向选择。Zhang 等^[39]同样比较了藏鸡和平原地区鸡的基因组特征,发现受选择的基因主要集中在心血管系统发育、DNA 修复、辐射、炎症和免疫反应功能,证明藏鸡经历了应对低氧高辐射的适应性进化过程。

鸡是鸟纲中与人类关系最为密切的物种,研究鸡在鸟纲中的进化地位具有重要意义。国际鸟类基因组研究联盟(Avian Genome Consortium)^[40]对 48 只鸟类物种进行了基因组测序、组装和全基因组比较分析,这些物种包括鸡、鸭、火鸡、乌鸦、隼、鸢、企鵝、朱鹮、啄木鸟、鹰等,这项研究基于全基因组数据构建了有史以来最高可信度鸟类分子演化树,同属鸡雁小纲的鸡、火鸡和鸭三者 在鸟纲中亲缘关系最近。在基因核型结构的研究中,Romanov 等^[41]分析了家鸡、火鸡、北京鸭、斑马雀(*Taeniopygia guttata*)和虎皮鸢(*Melopsittacus undulates*)的全基因组,发现家鸡具有与鸟类共同祖先最相似的染色体结构。而对于鸟类独特的性染色体 Z、W 也进行细致的研究,发现不同鸟类的性染色体处于不同的演化阶段中,相比于其他鸟类,家鸡和斑马雀的性染色体 W 则只包含少数的功能基因,这也可能是鸡公母两性在外表和性能上差异不显著的原因之一^[42]。

2 普通牛和牦牛

牛是家养大型反刍动物,从早期的役用,到当今的乳用和肉用,与人类的关系一直非常密切。随着农业的发展和消费需求的变化,牛的选育已向专门化方向发展,牛基因组的破译不仅有助人们更深入了解牛的驯化过程和逆境适应的机理,对于提高牛肉、牛奶的品质具有更重要意义。

2.1 参考基因组

牛基因组测序联盟(The Bovine Genome Sequencing and Analysis Consortium)^[7]于 2009 年公布了海福特牛的全基因组序列,他们选择海福特 L1 系(近交系数为 0.3)的一头母牛为材料,采用 BAC 策略结合一代测序技术,测序深度为 7×,其中 Contig N50

为 48.7 kb, Scaffold N50 为 1.9 Mb, 组装后基因组全长 2.87 Gb。组装质量检测表明,全基因组序列的 90%能够被装配到 29 个染色体和 X 染色体上,其中 Contig 包含了 95%的 EST 序列,GC 含量为 41.7%。基因注释采用 NCBI 系统、Ensembl 系统以及 Fgenesh 和 Fgenesh++、Geneid 和 SGP2、GLEAN consensus gene set 相结合的方法,共注释出 22 000 个基因。

牦牛(*Bos grunniens*)是高海拔地区所特有的牛属家畜,2012 年中国学者 Qiu 等^[13]测定了牦牛的全基因组序列,他们选择近交系数为 0.094 的一头雌性牦牛为材料,采用全基因组鸟枪法结合二代测序技术,测序深度为 65×,利用 SOAP *de novo* 软件进行组装,其中 Contig N50 为 20.4 kb, Scaffold N50 为 1.4 Mb。组装后基因组全长为 2.66 Gb,包括 22 282 个基因。用 fosmid 序列、RNA-seq 对组装质量进行检测,结果表明 scaffolds 中包含了 97%的 fosmid 序列,GC 含量为 41.7%。

2.2 基因组后续研究

牛基因组公布的同时,牛单体型图计划联盟(The Bovine HapMap Consortium)在同期杂志上发表了牛主要品种遗传结构和进化历史的研究报告^[19],该研究首先基于 6 个主要品种的重测序构建了牛 SNP 数据库,再定制 SNP 芯片对 19 个品种共 497 头牛的遗传变异分析发现,由于驯化、选择和育种过程中可能经历了瓶颈效应,尤其是公牛数量的急剧下降,牛的有效群体含量由原本非常大的祖先群体在近期快速下降,并分析鉴定了基因组中人工选择的区域。

旨在为芯片 SNP 填充提供充分的参考数据,进而促进牛基因组选择和全基因组关联分析研究,多国研究机构联合发起了“千牛基因组计划”(The 1000 Bull Genomes Project)。2014 年,澳大利亚科学家们率先发表了一项重要研究成果^[20]:他们测定了 234 头牛的全基因组序列,所选个体是荷斯坦牛、弗莱维赫牛、娟姗牛 3 个品种的“关键祖先”,涵盖各自品种内的遗传多样性信息,这项研究共得到 2830 多万个变异,在此基础上该研究通过全基因组关联分析等方法定位了与牛健康和生产性状密切相关的基因和致因变异位点,研究发现胚胎死亡是由 *SMC2*

基因突变导致;被毛卷曲是由 *KRT27* 基因的一个错义突变所致;对于乳脂率性状,除了定位已知的主效基因 *DGAT1* 外,还发现了多个具有较大效应的基因如 *AGPAT6* 等。这项研究使得今后早期选择健康个体成为可能,从而增加产肉和产奶的效率,体现了“千牛基因组计划”的重要现实意义。

遗传变异的挖掘为加速改良牛奶和肉类的遗传性能提供巨大的资源库,而充分利用这些资源便成为众多研究工作的内容之一。Larkin 等^[43]结合全基因组重测序技术,对奶牛育种史上占有重要地位的两头种公牛(Chief 和 Mark)进行全基因组测序后,结合芯片数据推断了 1149 个后裔的单倍型图谱,旨在利用这些信息鉴定奶牛中受选择的染色体片段,在 Chief 个体中鉴定到受强烈选择的 49 个染色体片段,明确了其中的 11 个候选基因,分别与产奶量、繁殖力以及抗病性状有关。Qanbari 等^[21]对 43 头弗来维赫牛进行全基因组重测序,找到了 106 个可能受选择的区域,包含了 *KIT*、*MITF*、*MC1R*、*NRG4*、*ERBB4*、*TMEM132D* 和 *TAS2R16* 等与毛色、性情和感观相关的驯化基因,并进一步通过 3062 个个体的外观、体尺和体细胞数性状全基因组关联分析发现毛色的显著区间与驯化信号高度重合,表明外观是驯化过程中的重要性状之一。

牛的进化一直受到关注,在牛漫长的进化过程中,曾进化为超过 200 个不同的族类、类型和品种,遍布世界各地^[19]。普通牛的野生祖先为已灭绝的原牛(*Bos primigenius*),但原牛的系统发育地位如何,现存牛品种与原牛是何种进化关系?古 DNA 技术和全基因组测序技术相结合可以帮助揭开进化历史的真相。Park 等^[44]从 6750 年前英国原牛的肱骨化石中提取 DNA 进行测序,重构了一只原牛的古 DNA 基因组,用于分析欧洲牛的进化历史。结合 81 头现代牛的全基因组测序数据以及其他 1225 头牛的 SNP 数据,对总共 73 个品种变异信息比较分析发现英国与爱尔兰的现代牛品种与原牛存在共有的遗传变异,表明两国现代牛的祖先曾与原牛有过基因交流,可能是早期的英国人曾经有意识的引入原牛基因来拓展牛的遗传基础,以原牛为外群的系统聚类结果也支持此前报道的欧洲牛起源于中东地区的说法。此外,研究人员还发现在牛的驯化过程中,神经、肌

肉发育和免疫相关基因受正向选择压力,表明在早期驯化中这些生物过程的改变是至关重要的。这项研究为现代牛的起源以及功能进化提供了全新信息。Mei 等^[45]对家牛的近缘物种——大额牛(*Bos frontalis*)个体进行重测序(13×),推测大额牛是雄性印度野牛(*Bos gaurus*)与家牛杂交的后代。

经过长期的自然和人工选择,牦牛已经获得了能够适应高原环境的生理特征和特有的生产性能。Qiu 等^[22]对 13 只野牦牛、48 只未经选育的驯化后牦牛和 11 只经过严格选育的天祝白牦牛群体进行重测序^[22],构建了包含 1400 万个位点的全基因组群体遗传变异图谱,野生牦牛和家养牦牛的人工选择信号分析发现牦牛基因组中大约有 209 个基因受到了显著的驯化选择,这些基因主要与温顺行为和经济性状等相关。通过重演牦牛驯化的群体动态和群体分化历史,推测牦牛由游牧民从 7300 年前开始驯化,牦牛数量在 3600 年前时增长到了驯化之初的 6 倍,表明通过牦牛驯化对早期人类永久性地征服青藏高原地区和随后的社会稳定发展起到了至关重要的作用。无角牦牛更适于现代化的集中饲养,Liang 等^[46]基于 10 个有角和 10 个无角牦牛群体的全基因组关联分析,将无角性状定位到一个 200 kb 包含 3 个基因且受到人工选择的区间,并推测其中某个基因表达量的变化引起的角部发育停止。

3 猪

猪肉作为一种重要的动物蛋白来源受到了广泛的欢迎,尤其在亚洲和欧洲的畜牧业生产中占有重要地位。研究表明早在 7.9~9.7 千万年前猪与人拥有共同的祖先^[47],又因为猪拥有和人类相似的饮食结构以及尺寸接近的重要器官,因此猪也适宜作为研究人类疾病的动物模型和异种器官移植供体^[48]。

3.1 参考基因组

国际猪基因组测序联盟(International Swine Genome Sequencing Consortium)^[8]于 2012 年公布了猪的全基因组序列,选择杜洛克猪 2-14 系中的一头雌性个体作为测序材料,分别构建了 BAC 和 fosmid 克隆文库,选用全基因组鸟枪法的测序手段,利用 Phrap、SSAHA 等软件进行组装,其中 Contig N50

为 80 720 bp, Scaffold N50 为 637 332 bp, 组装后的基因组全长为 2.60 Gb, 包括 21 640 个编码基因。同年, 华大基因与中国农业科学院的研究者合作共同测定了五指山猪(*Sus scrofa*)近交系个体的全基因组序列^[14], 连续 20 代近交后的五指山猪杂合度显著降低, 基因组 59.97% 区间杂合度小于 0.01%, 与近交系小鼠 F280 相当(69.15%), 在测序深度 78× 情况下, Contig N50 为 23.5 kb, 但 Scaffold N50 却达到 5.4 Mb, 相对于杜洛克猪的 Scaffold 组装质量有所提升, 组装后基因组全长为 2.6 Gb, 注释出 20 326 个基因以及 11 843 个编码序列。很显然, 由于遗传分化, 个体、群体和种属间必然存在基因组长短的差异, 以单一个体的基因组序列作为参考基因组进行后续的群体遗传学研究存在局限性, 理想参考基因组应该涵盖该物种的全部遗传信息, 因此, 需要采用多品种从头测序后取并集来提升基因组的可参考性。Li 等^[49]正是基于这样的考虑, 选取中国和欧洲的 10 个代表性猪种(中国猪种: 梅山猪、荣昌猪、八眉猪、金华猪、藏猪; 欧洲猪种: 长白猪、大白猪、汉普夏猪、皮特兰猪、巴克夏猪), 采用 *de novo* 策略组装了它们的基因组序列, 获得了原杜洛克猪参考基因组之外的 137.02 Mb 序列, 包含 1737 个编码蛋白的基因, 采用比对策略评估显示, 10 个品种未比对上参考基因组的读段(orphan reads), 其中的 95.04% 都可以比对到这些序列上。因此, 综合这些猪种共有和特有序列可以形成一个更加完整的猪参考基因组。

基因组中高度重复区域通常难以组装, 尤其是性染色体和分化时间较短的旁系同源基因。Skinner 等^[50]基于 array-CGH 芯片等技术鉴定了猪 XY 染色体的同源区域并明确了 X 染色体的拟常染色体(PAR)的边界; Benjamin 等^[51]结合 BAC 和 fosmid 测序、光学图谱以及荧光原位杂交技术组装了杜洛克猪 X 和 Y 染色体序列, 大幅度提高了猪性染色体的序列与注释信息。Dawson 等^[52]采用共表达的方法注释了猪基因组中的免疫组相关基因, 并观察到一些免疫基因家族存在扩张现象。

3.2 基因组后续研究

国际猪基因组测序联盟在公布猪基因组的文章中报道了猪种群差异以及驯化的研究结果^[8], 发现

亚洲猪相对于欧洲猪而言拥有更高的遗传多样性, 产生这一现象的原因可能是由于末次盛冰期(LGM)使整个欧洲猪遭受重创, 群体骤然减少, 造成现存欧洲猪种的遗传多样性相对较低。研究者还发现家猪中一些基因家族经历了相对快速的进化, 免疫系统和嗅觉基因迅速扩增。同年, Rubin 等^[23]利用全基因组重测序技术揭示了欧洲家养猪基因组中一些与表型进化相关的基因座, 发现欧洲家养猪背部变长和椎骨数的增加都是由 *NR6A1*、*PLAG1* 以及 *LCORL* 3 个基因座引起的, 而家养猪中存在过量的非同义突变很可能是驯化后又经历了正、负双向选择; 他们还基因组结构变异进行了深入分析, 验证了 *KIT* 位点多拷贝协同第 17 内含子剪切位点突变调控猪的毛色, 并在多拷贝片段内部鉴定了 3 个重要的嵌套拷贝数变异, 这 3 个变异的出现会产生汉普夏猪白颈圈的独特表型。

中国饲养了世界上近 50% 的猪, 品种资源丰富且分布区域广泛, 但品种间、地理区域间的猪种在基因组上的差异及形成机制在早期受限于研究方法, 进展缓慢。猪基因组测序计划完成后, 以杜洛克猪和五指山猪的参考基因组为基础, 中国学者基于重测序技术开展了多项具有影响力的工作。Li 等^[24]结合藏猪 *de novo* 信息对高原藏区野猪和中国家猪群体进行比较基因组研究, 发现受选择基因多参与缺氧、嗅觉、能量代谢和药物反应等过程, 并揭示了驯化过程中家猪增加唾液分泌的遗传基础。Ai 等^[27]对中国 15 个不同地理居群的 69 头猪进行深度重测序, 进一步揭示了猪环境适应性的分子机理, 这一研究鉴定了 4100 万个基因组变异位点, 其中 52% 为新发现的位点, 发现中欧猪种之间存在广泛的基因交流, 证实了中国地方猪对全球现代商业猪种的培育做出过重要贡献; 人工选择性分析鉴定出不同纬度环境适应性相关基因位点 219 个, 这些位点与毛发生长、皮肤发育、血液循环、能量代谢、肾脏发育、前脑神经元调控等生理功能有关; 该研究还在 X 染色体发现一个长达 14 Mb 的低重组区, 南北方猪在该区域存在两种截然不同且经历了环境适应性自然选择的单倍型, 北方猪单倍型很可能来自猪属另一个已经灭绝的种 *Suide*, 该属间杂交事件据推算发生于数十万年前。

猪早期跨洲际品种杂交以及后期商业化育种必然在基因组中留下印迹,这是目前受到广泛关注的课题。Bosse 等^[25]针对中欧洲猪种的基因渗入进行了研究,欧洲猪在育种史上曾引入中国优秀品种进行杂交,导入了高繁殖力等基因。这项研究对亚洲和欧洲具有代表性的野猪、家养和商业品种共 70 头猪进行测序,明确发现欧洲猪种基因组中渗入了亚洲猪的单倍型,这些渗入的单倍型区域主要与猪的肉质、生长以及繁殖性能有关,同时发现源于亚洲猪的 *AHR* 基因非同义突变位点与产仔数增加显著相关。Frantz 等^[26]根据 100 多只欧洲和亚洲野猪和家猪重测序数据分析它们之间的近万年的驯化历程,发现这一过程存在广泛的基因流动,即驯化与杂交是持续交替进行的,因此不适用于如生殖隔离或驯化瓶颈为假设的驯化模型。对于为何家猪能够保持特有的形态和行为性状,作者推断野猪与家猪杂交后会通过类似于轮回选择的方式来抵消基因流动引发的均化效应,并会形成基因组上的“驯化孤岛”。Choi 等^[53]对韩国本地猪、野猪及欧洲猪 5 个品种共 55 头猪进行重测序,检测到大量新的 SNP 以及一些可能与经济性状有关的受选择基因座。

4 山羊和绵羊

家羊分为两种,即山羊和绵羊,分别属于牛科的山羊属和绵羊属,两者的分歧时间大约在 400 万年前左右。它们是人类肉、皮、毛、绒的重要来源,羊基因组序列的测定对遗传标记辅助育种和改善羊的经济性状具有重要作用。

4.1 参考基因组

2012 年,华大基因和中国科学院的研究者公布了山羊的基因组序列^[9],他们选择了一头云南黑山羊的雌性个体为材料,构建了 fosmid 文库,采用二代测序技术与全基因组酶切光学图谱技术相结合共同绘制出基因组图谱,测序深度为 65.6×,其中 Contig N50 为 18 720 bp, Scaffold N50 为 2 212 139 bp,组装后基因组全长为 2.66 Gb,共包括 22 175 个基因。组装质量检测表明,超过 89% 的双末端读取序列可以映射到组装的山羊基因组上,这些读取序列中的 95% 都已成功装配,说明组装质量已达到较为理想

的水平。2014 年,该团队又公布了绵羊的基因组序列^[10],他们以特克塞尔绵羊为材料,采用同样的技术进行测序和组装,测序深度为 75×,其中 Contig N50 为 39 959 bp, Scaffold N50 为 2 231 873 bp,组装后基因组全长 2.61 Gb,共包括 20 908 个基因。组装质量检测表明,超过 99.3% 的测序个体的 mRNA 可以映射到 Oar v3.1 上(平均水平为 98.4%),说明组装质量已达到较为理想的水平。在绵羊基因组的基础上,Miller 等^[54]采用重测序(12×)比对策略构建了绵羊的近缘物种——大角羊(*Ovis canadensis*)的基因组序列,结果显示有 95% 的序列可以比对到绵羊基因组,说明跨种间比对可以创建近缘种的全基因组序列。

4.2 基因组后续研究

绵羊和山羊的分化时间是物种进化的研究热点,全基因组水平的进化推断是较为令人信服的结论,在绵羊参考基因组的报告中,对哺乳动物保守的 4850 个单拷贝同源基因序列聚类并推断绵羊和山羊的分化时间约在 430 万年前^[10],恰好与新第三纪后期草地的扩张处在同一时期,说明充足的食物为新种属的产生提供了条件。反刍动物一个广受关注的问题是瘤胃是如何产生的。研究人员在绵羊的 EDC (epidermal development complex)区域中首次鉴别到两种反刍动物瘤胃中特异高表达的结构蛋白——毛透明蛋白类似蛋白(Trichohyalin-like 2)和小脯氨酸丰富蛋白家族(PRD-SPRR12),这两个结构蛋白发挥瘤胃表面基板的作用,通过转谷氨酰胺酶介导交联瘤胃表达的角蛋白,从而构成瘤胃壁粘膜层的坚韧的角质化表面,参与瘤胃表层角蛋白角质化形成。绵羊皮肤的转录组数据表明 *MOGAT2* 和 *MOGAT3* 发生高表达且在基因组中发现多个拷贝,它们主要参与甘油二酯(DAG)和三酰甘油(TAG)的合成,说明反刍动物的皮肤也是重要的脂类代谢器官。

羊绒是羊所独有的产品,但其形成和发育的分子机制却知之甚少,在山羊参考基因组的报告中,科研人员对初级毛囊和次级毛囊进行了转录组研究,发现其中有 2 个角蛋白基因和 10 个角蛋白辅助蛋白基因在初级毛囊和次级毛囊中差异表达,并可在次级毛囊中高度表达,这表明角蛋白及角蛋白辅助蛋

白基因在羊绒纤维结构中起着重要作用。此外,在山羊基因组中鉴定出 44 个基因受到正选择,其中 7 个与免疫相关,另外还有 3 个垂体功能相关的基因发生了快速进化,这可能与产奶量、胚胎发育速度以及羊毛变化等相关。在研究山羊基因家族扩张与收缩情况时发现 3 个与味觉受体相关的基因亚家族发生了扩张,而仅有 1 个亚家族收缩,推测这可能与山羊的觅食能力相关;同时还发现铁蛋白重链多肽-1 基因(*FTH1*)扩张至人类的 7 倍,并推测这可能与山羊强大的解毒及觅食能力相关^[10],因为铁蛋白在铁离子螯合、解毒和贮存中发挥着重要作用。Dong 等^[55]对一只伊朗的野山羊进行了 *de novo* 测序,比较野山羊和家养山羊的参考基因组序列发现 13 个拷贝数变异基因与毛色有关,其中的 *ASIP* 基因重复产生了白色的表型,此外还发现与行为、免疫、生产等有关的基因已经快速的进化,这些特异的基因不仅可以给未来山羊功能基因的研究提供候选基因,也为理解动物驯化的遗传机制提供了有用的信息。

绵羊在西南亚被驯化之后,迅速地扩散到了世界上不同的生态地理区域,并对多种极端环境产生了良好的适应性,是研究家畜环境适应性的理想模式物种。近期, Yang 等^[28]对 77 只中国地方品种绵羊和 3 只野羊进行全基因组重测序,比较了来自高原和平原、干旱沙漠和湿润地区样本的基因组,发现了一系列与绵羊极端环境适应性相关的新的候选基因,以及相应的功能类别和信号通路。研究发现在高原环境下,受选择基因的 GO 和通路和低氧耐受反应有关;在沙漠环境下,则与水分子的重吸收有关。而能量代谢和体型大小变异也被确定与绵羊对高原、沙漠等极端环境的适应性相关。该研究所发现的上述遗传机理能够很好地解释极端环境下绵羊的表型变异,从基因组水平上揭示了中国地方绵羊的群体遗传结构和种群历史动态。

5 鸭和鹅

鸭和鹅是中国特色水禽,饲养量分别占全世界的 80%和 94%,为人类提供了肉、蛋、绒等产品。水禽的抗病力较强,是研究天然免疫机制的很好模型。此外,鹅还具有独特的脂肪储存和肝脏代谢特征。

5.1 参考基因组

Huang 等^[11]于 2013 年公布了北京鸭的基因组草图,他们以 10 周龄的北京鸭母鸭为材料,采用经典的 shotgun 策略,测序深度为 64×,组装后的基因组 Contig N50 为 26 kb, Scaffold N50 为 1.2 Mb,基因组全长 1.10 Gb,共包括 19 144 个基因。Rao 等^[56]基于鸭 Scaffold 构建了 RH 图谱,将鸭基因组装配到染色体水平。Lu 等^[12]于 2015 年公布了鹅的全基因组序列草图,他们以一只 70 日龄的雄性浙东白鹅为材料,测序量为 139.55 Gb,采用 SOAP *de novo* 软件进行组装,其中 Contig N50 为 27.5 kb, Scaffold N50 为 5.2 Mb,组装后基因组大小为 1.12 Gb,共包括了 16 150 个基因。组装质量检测表明,参考基因组覆盖了转录组数据库中 98%的序列,平均 GC 含量为 38%,说明基因组组装已达到较高水平。

5.2 基因组后续研究

Huang 等^[11]同时还针对鸭比鸡对流感病毒具有更强的耐受性的科学问题进行了探索,通过对比分析高致病性及低致病性 H5N1 病毒感染后的鸭子和对照组间的肺部组织转录组数据,发现 β -defensin 和 *BTNL* 基因在鸭的免疫反应中起重要作用,而这些免疫相关基因只在鸭基因组中发生了独立复制,鸡却没有。这很好地解释了为何鸭对流感病毒具有更强的耐受性。

相比于其他鸟类,鹅的肝脏能够承受过度的脂肪积累,通常在短期过量饲喂(约 2~3 周)后的脂肪肝比正常肝脏增大 5~10 倍,但通常不会发展为肝纤维化或坏死。对鹅正常肝脏和填饲后的脂肪肝转录组分析发现,鹅 *LEP* 基因的缺失可能是正选择的结果,从而使肝脏采用能量存储机制进行远距离迁移^[12]。

6 展 望

“人类基因组计划”的完成标志着人类对于生命现象和过程的认识进入了以规模化技术为主线的、以发现为导向的信息密集型阶段。动物遗传学也从以前候选基因、单一性状为主的微观研究转向以物种进化构架内的全基因组水平、多性状、组学化的宏观研究,完成了从抽样到整体的升华。这一转变不仅标志动物遗传学新时代的到来,对于畜禽育种

更具有跨时代的意义。在这样的大数据时代, 畜禽基因组学仍需要不断调整研究思路和策略来迎合和充分利用测序技术进步所带的福祉。

人类对于任何一个学科的认识都有一个由浅入深的过程, 尤其是进化生物学, 着眼于对远古或史前事件时间节点的推测, 存在诸多的不确定性, 基因组学的出现为进化研究提供了更加翔实的数据支撑, 但仍不可避免的存在诸多争议性问题。可喜的是, 面对异议, 科学家可以在杂志上进行理性的辩论, 启发读者从不同角度思考问题, 促进了基因组学科的健康发展^[57-60]。

畜禽基因组学未来在测序技术、群体设计、表型测定、系统学研究方面仍面对诸多挑战。目前, 重测序技术以短片段比对参考基因组获得变异数据, 在深度足够的条件下能够得到非常精确的 SNP、小片段插入缺失和拷贝数变异, 但对于如大片段插入、序列倒置易位等结构变异的检出效率非常有限。因此, 长读段成为了下一代测序技术的主要特点, 不难想像在未来可以实现对整条染色体的一次性完整测序。就当前而言, 当长读段(> 20 kb)测序仪的性能和测序成本达到合理水平后, 基于长读段拼接个体基因组再比对鉴定的群体变异是最为精确和全面的, 这种被称为泛基因组(pan genome)研究^[49, 61]未来会得到更广泛的应用。从前面的描述我们不难发现, 动物表型差异有很大一部分是结构变异引起的^[62], 我们相信在泛基因组时代, 将会有一大批目前难以用基因芯片和重测序技术解析的性状之谜得以解开。

畜禽重要性状的遗传基础是长期以来动物遗传学研究领域的核心科学问题。解答这一问题, 群体设计是至关重要的第一步, 在 GWAS 时代, 虽然随机群体对于一些简单性状如毛色基因定位也能得到较为理想的结果^[63], 但纵观动物遗传学发展历史, 合理设计的遗传资源群体对解析动物复杂性状的机制起到了至关重要的作用。在基因组学时代, 理想群体再结合全基因组测序技术, 人们将能得到更为精准的基因定位信息, 加速对复杂性状形成分子机理的认知。

表型是一个综合性的概念, 与基因型高通量测定不匹配的是表型测定技术和理念至今仍未取得重要突破。随着人们对模式生物、植物相关表型认识的不断深入, 采用化繁为简的思路, 将复杂性状进

行生物学分解测定有助于对遗传机理的探究。以脂类研究为例, 分解为一系列脂肪酸小分子的研究, 此类小分子代谢物的代谢途径和遗传调控较为简单, 易于定位大效应位点, 最后综合所有脂肪酸的基因定位信息则可以绘制整个脂肪遗传调控的网络^[64, 65]。测序技术的突飞猛进也带动了系统遗传学的快速发展^[66], 理清从静态的基因组、动态的转录组和蛋白组, 到稳态的代谢组整个遗传调控过程, 有助于挖掘从基因到表型的因果关系。

畜禽遗传学研究的终极目标是探究家养动物遗传变异的原理, 从而应用于畜禽品种改良来改善人类的生活与健康。基因组学的广泛应用将会推动遗传学成果的批量产出, 这些成果如何应用是要着重思考的问题。自主设计、价格低廉的低密度基因芯片将拥有更为广阔的应用前景。另一方面, 主效基因的致因变异位点为基因编辑提供了明确的靶点, 遗传修饰育种将拥有大量可编辑素材。相信在基因组学的大数据时代, 动物遗传学将会取得长足的进步, 最终造福人类。

参考文献(References):

- [1] International Human Genome Sequencing Consortium, Lander ES, Linton L M, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrum J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Whitehead Institute for Biomedical Research, Center for Genome Research, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, The Sanger Centre, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissole SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Washington University Genome Sequencing Center, Hawkins T,

- Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, US DOE Joint Genome Institute, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Baylor College of Medicine Human Genome Sequencing Center, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Todd Taylor for RIKEN Genomic Sciences Center, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Genoscope and CNRS UMR-, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Andreas Rump for Department of Genome Analysis, Institute of Molecular Biotechnology, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, JoAnn Dubois for GTC Sequencing Center, Yang HM, Yu J, Wang J, Huang GY, Gu J, Beijing Genomics Institute/Human Genome Center, Hood L, Rowen L, Madan A, Qin SZ, Multimegabase Sequencing Center, The Institute for Systems Biology, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Stanford Genome Technology Center, Roe BA, Chen F, Pan HQ, University of Oklahoma's Advanced Center for Genome Technology, Ramser J, Lehrach H, Richard Reinhardt for Max Planck Institute for Molecular Genetics, McCombie WR, de la Bastide M, Dedhia N, Cold Spring Harbor Laboratory, Lita Annenberg Hazen Genome Center, Blöcker H, Hornischer K, Nordsiek G, GBF—German Research Centre for Biotechnology, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglu S, Ewan B, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JGR, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kasprzyk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AFA, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Genome Analysis Group, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Scientific management, National Human Genome Research Institute, US National Institutes of Health, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Stanford Human Genome Center, Olson MV, Kaul R, Raymond C, University of Washington Genome Center, Shimizu N, Kawasaki K, Minoshima S, Department of Molecular Biology, Keio University School of Medicine, Evans GA, Athanasiou M, Schultz R, University of Texas Southwestern Medical Center at Dallas, Aristides Patrinos for Office of Science, US Department of Energy, Morgan MJ, The Wellcome Trust. Initial sequencing and analysis of the human genome. *Nature*, 2001, 409(6822): 860–921. [DOI]
- [2] The International HapMap Consortium. The international HapMap project. *Nature*, 2003, 426(6968): 789–796. [DOI]
- [3] The ENCODE Project Consortium, ENCODE Project Scientific Management: National Human Genome Research Institute, Feingold EA, Good PJ, Guyer MS, Kamholz S, Liefer L, Wetterstrand K, Collins FS, Initial ENCODE Pilot Phase Participants: Affymetrix, Inc., Gingeras TR, Kampa D, Sekinger EA, Cheng J, Hirsch H, Ghosh S, Zhu Z, Patel S, Piccolboni A, Yang A, Tammanna H, Bekiranov S, Kapranov P, Harrison R, Church G, Struhl K, Ludwig Institute for Cancer Research, Ren B, Kim TH, Barrera LO, Qu C, Van Calcar S, Luna R, Glass CK, Rosenfeld MG, Municipal Institute of Medical Research, Guigo R, Antonarakis SE, Birney E, Brent M, Pachter L, Reymond A, Dermitzakis ET, Dewey C, Keefe D, Denoeud F, Lagarde J, Ashurst J, Hubbard T, Wesselink JJ, Castelo R, Eyraes E, Stanford University, Myers RM, Sidow A, Batzoglu S, Trinklein ND, Hartman SJ, Aldred SF, Anton E, Schroeder DI, Marticke SS, Nguyen L, Schmutz J, Grimwood J, Dickson M, Cooper GM, Stone EA, Asimenos G, Brudno M, University of Virginia, Dutta A, Karnani N, Taylor CM, Kim HK, Robins G, University of Washington, Stamatoyannopoulos G, Stamatoyannopoulos JA, Dorschner M, Sabo P, Hawrylycz M, Humbert R, Wallace J, Yu M, Navas PA, McArthur M, Noble WS, Wellcome Trust Sanger Institute, Dunham I, Koch CM, Andrews RM, Clelland GK, Wilcox S, Fowler JC, James KD, Groth P, Dovey OM, Ellis PD, Wraight VL, Mungall AJ, Dhami P, Fiegler H, Langford CF, Carter NP, Vetric D, Yale University, Snyder M, Euskirchen G, Urban AE, Nagalakshmi U, Rinn J, Popescu G, Bertone P, Hartman S, Rozowsky J, Emanuelsson O, Royce T, Chung S, Gerstein M, Lian Z, Lian J, Nakayama Y, Weissman S, Stolc V, Tongprasit W, Sethi H, Additional ENCODE Pilot Phase Participants: British Columbia Cancer Agency Genome Sciences Centre, Jones S, Marra M, Shin H, Schein J, Broad Institute, Clamp M, Lindblad-Toh K, Chang J, Jaffe DB, Kamal M, Lander ES, Mikkelsen TS, Vinson J, Zody

- MC, Children's Hospital Oakland Research Institute, de Jong PJ, Osoegawa K, Nefedov M, Zhu B, National Human Genome Research Institute/Computational Genomics Unit, Baxeavanis AD, Wolfsberg TG, National Human Genome Research Institute/Molecular Genetics Section, Collins FS, Crawford GE, Whittle J, Holt IE, Vasicek TJ, Zhou D, Luo S, NIH Intramural Sequencing Center/National Human Genome Research Institute, Green ED, Bouffard GG, Margulies EH, Portnoy ME, Hansen NF, Thomas PJ, McDowell JC, Maskeri B, Young AC, Idol JR, Blakesley RW, National Library of Medicine, Schuler G, Pennsylvania State University, Miller W, Hardison R, Elnitski L, Shah P, The Institute for Genomic Research, Salzberg SL, Pertea M, Majoros WH, University of California, Santa Cruz, Haussler D, Thomas D, Rosenbloom KR, Clawson H, Siepel A, Kent WJ, ENCODE Technology Development Phase Participants: Boston University, Weng Z, Jin S, Halees A, Burden H, Karaoz U, Fu Y, Yu Y, Ding C, Cantor CR, Massachusetts General Hospital, Kingston RE, Dennis J, NimbleGen Systems, Inc., Green RD, Singer MA, Richmond TA, Norton JE, Farnham PJ, Oberley MJ, Inman DR, NimbleGen Systems, Inc., McCormick MR, Kim H, Middle CL, Pirrung MC, University of California, San Diego, Fu XD, Kwon YS, Ye Z, University of Massachusetts Medical School, Dekker J, Tabuchi TM, Gheldof N, Dostie J, Harvey SC. The ENCODE (ENCyclopedia Of DNA elements) project. *Science*, 2004, 306(5696): 636–640. [DOI]
- [4] Siva N. 1000 Genomes project. *Nat Biotechnol*, 2008, 26(3): 256. [DOI]
- [5] Andersson L, Archibald AL, Bottema CD, Brauning R, Burgess SC, Burt DW, Casas E, Cheng HH, Clarke L, Couldrey C, Dalrymple BP, Elsik CG, Foissac S, Giuffra E, Groenen MA, Hayes BJ, Huang LS, Khatib H, Kijas JW, Kim H, Lunney JK, McCarthy FM, McEwan JC, Moore S, Nanduri B, Notredame C, Palti Y, Plastow GS, Reecy JM, Rohrer GA, Sarropoulou E, Schmidt CJ, Silverstein J, Tellam RL, Tixier-Boichard M, Tosser-Klopp G, Tuggle CK, Vilki J, White SN, Zhao SH, Zhou HJ. Coordinated international action to accelerate genome-to-phenome with FAANG, the Functional Annotation of Animal Genomes project. *Genome Biol*, 2015, 16(1): 57. [DOI]
- [6] Hillier LW, Miller W, Birney E, Warren W, Hardison RC, Ponting CP, Bork P, Burt DW, Groenen MAM, Delany ME, Dodgson JB, Chinwalla AT, Cliften PF, Clifton SW, Delehaunty KD, Fronick C, Fulton RS, Graves TA, Kremitzki C, Layman D, Magrini V, McPherson JD, Miner TL, Minx P, Nash WE, Nhan MN, Nelson JO, Oddy LG, Pohl CS, Randall-Maher J, Smith SM, Wallis JW, Yang SP, Romanov MN, Rondelli CM, Paton B, Smith J, Morrice D, Daniels L, Tempest HG, Robertson L, Masabanda JS, Griffin DK, Vignal A, Fillon V, Jacobsson L, Kerje S, Andersson L, Crooijmans RPM, Aerts J, van der Poel JJ, Ellegren H, Caldwell RB, Hubbard SJ, Grafham DV, Kierzek AM, McLaren SR, Overton IM, Arakawa H, Beattie KJ, Bezubov Y, Boardman PE, Bonfield JK, Croning MDR, Davies RM, Francis MD, Humphray SJ, Scott CE, Taylor RG, Tickle C, Brown WRA, Rogers J, Buerstedde JM, Wilson SA, Stubbs L, Ovcharenko I, Gordon L, Lucas S, Miller MM, Inoko H, Shiina T, Kaufman J, Salomonsen J, Skjoedt K, Wong GKS, Wang J, Liu B, Wang J, Yu J, Yang HM, Nefedov M, Koriabine M, deJong PJ, Goodstadt L, Webber C, Dickens NJ, Letunic I, Suyama M, Torrents D, von Mering C, Zdobnov EM, Makova K, Nekrutenko A, Elnitski L, Eswara P, King DC, Yang S, Tyekucheva S, Radakrishnan A, Harris RS, Chiaromonte F, Taylor J, He JB, Rijnkels M, Griffiths-Jones S, Ureta-Vidal A, Hoffman MM, Severin J, Searle SMJ, Law AS, Speed D, Waddington D, Cheng Z, Tuzun E, Eichler E, Bao ZR, Flicek P, Shteynberg DD, Brent MR, Bye JM, Huckle EJ, Chatterji S, Dewey C, Pachter L, Kouranov A, Mourelatos Z, Hatzigeorgiou AG, Paterson AH, Ivarie R, Brandstrom M, Axelsson E, Backstrom N, Berlin S, Webster MT, Pourquie O, Reymond A, Ucla C, Antonarakis SE, Long MY, Emerson J, Betrán E, Dupanloup I, Kaessmann H, Hinrichs AS, Bejerano G, Furey TS, Harte RA, Raney B, Siepel A, Kent WJ, Haussler D, Eyraes E, Castelo R, Abril JF, Castellano S, Camara F, Parra G, Guigo R, Bourque G, Tesler G, Pevzner PA, Smit A, Fulton LA, Mardis ER, Wilson RK. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*, 2004, 432(7018): 695–716. [DOI]
- [7] The Bovine Genome Sequencing and Analysis Consortium, Elsik CG, Tellam RL, Worley KC. The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science*, 2009, 324(5926): 522–528. [DOI]
- [8] Groenen MAM, Archibald AL, Uenishi H, Tuggle CK, Takeuchi Y, Rothschild MF, Rogel-Gaillard C, Park C, Milan D, Megens HJ, Li ST, Larkin DM, Kim H, Frantz LAF, Caccamo M, Ahn H, Aken BL, Anselmo A, Anthon C, Auvil L, Badaoui B, Beattie CW, Bendixen C, Berman D, Blecha F, Blomberg J, Bolund L, Bosse M, Botti S, Bujie Z, Bystrom M, Capitanu B, Carvalho-Silva D, Chardon P, Chen C, Cheng R, Choi SH, Chow W, Clark

- RC, Clee C, Crooijmans RPMA, Dawson HD, Dehais P, De Sapio F, Dibbits B, Drou N, Du ZQ, Eversole K, Faddista J, Fairley S, Faraut T, Faulkner GJ, Fowler KE, Fredholm M, Fritz E, Gilbert JGR, Giuffra E, Gorodkin J, Griffin DK, Harrow JL, Hayward A, Howe K, Hu ZL, Humphray SJ, Hunt T, Hornshøj H, Jeon JT, Jern P, Jones M, Jurka J, Kanamori H, Kapetanovic R, Kim J, Kim JH, Kim KW, Kim TH, Larson G, Lee K, Lee KT, Leggett R, Lewin HA, Li YR, Liu WS, Loveland JE, Lu Y, Lunney JK, Ma J, Madsen O, Mann K, Matthews L, McLaren S, Morozumi T, Murtaugh MP, Narayan J, Nguyen DT, Ni PX, Oh SJ, Onteru S, Panitz F, Park EW, Park HS, Pascal G, Paudel Y, Perez-Enciso M, Ramirez-Gonzalez R, Reecy JM, Rodriguez-Zas S, Rohrer GA, Rund L, Sang YM, Schachtschneider K, Schraiber JG, Schwartz J, Scobie L, Scott C, Searle S, Servin B, Southey BR, Sperber G, Stadler P, Sweedler JV, Tafer H, Thomsen B, Wali R, Wang J, Wang J, White S, Xu X, Yerle M, Zhang GJ, Zhang JG, Zhang J, Zhao SH, Rogers J, Churcher C, Schook LB. Analyses of pig genomes provide insight into porcine demography and evolution. *Nature*, 2012, 491(7424): 393–398. [DOI]
- [9] Dong Y, Xie M, Jiang Y, Xiao NQ, Du XY, Zhang WG, Tosser-Klopp G, Wang JH, Yang S, Liang J, Chen WB, Chen J, Zeng P, Hou Y, Bian C, Pan SK, Li YX, Liu X, Wang WL, Servin B, Sayre B, Zhu B, Sweeney D, Moore R, Nie WH, Shen YY, Zhao RP, Zhang GJ, Li JQ, Faraut T, Womack J, Zhang YP, Kijas J, Cockett N, Xu X, Zhao SH, Wang J, Wang W. Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (*Capra hircus*). *Nat Biotechnol*, 2013, 31(2): 135–141. [DOI]
- [10] Jiang Y, Xie M, Chen WB, Talbot R, Maddox JF, Faraut T, Wu CH, Muzny DM, Li YX, Zhang WG, Stanton JA, Brauning R, Barris WC, Hourlier T, Aken BL, Searle SMJ, Adelson DL, Bian C, Cam GR, Chen YL, Cheng SF, DeSilva U, Dixen K, Dong Y, Fan GY, Franklin IR, Fu SY, Fuentes-Utrilla P, Guan R, Highland MA, Holder ME, Huang GD, Ingham AB, Jhangiani SN, Kalra D, Kovar CL, Lee SL, Liu WQ, Liu X, Lu CX, Lv T, Mathew T, McWilliam S, Menzies M, Pan SK, Robelin D, Servin B, Townley D, Wang WL, Wei B, White SN, Yang XH, Ye C, Yue YJ, Zeng P, Zhou Q, Hansen JB, Kristiansen K, Gibbs RA, Flicek P, Warkup CC, Jones HE, Oddy VH, Nicholas FW, McEwan JC, Kijas JW, Wang J, Worley KC, Archibald AL, Cockett N, Xu X, Wang W, Dalrymple BP. The sheep genome illuminates biology of the rumen and lipid metabolism. *Science*, 2014, 344(6188): 1168–1173. [DOI]
- [11] Huang YH, Li YR, Burt DW, Chen HL, Zhang Y, Qian WB, Kim H, Gan SQ, Zhao YQ, Li JW, Yi K, Feng HP, Zhu PY, Li B, Liu QY, Fairley S, Magor KE, Du Z, Hu XX, Goodman L, Tafer H, Vignal A, Lee T, Kim KW, Sheng ZY, An Y, Searle S, Herrero J, Groenen MAM, Crooijmans RPMA, Faraut T, Cai QL, Webster RG, Aldridge JR, Warren WC, Bartschat S, Kehr S, Marz M, Stadler PF, Smith J, Kraus RHS, Zhao YF, Ren LM, Fei J, Morisson M, Kaiser P, Griffin DK, Rao M, Pitel F, Wang J, Li N. The duck genome and transcriptome provide insight into an avian influenza virus reservoir species. *Nat Genet*, 2013, 45(7): 776–783. [DOI]
- [12] Lu LZ, Chen Y, Wang Z, Li XF, Chen WH, Tao ZR, Shen JD, Tian Y, Wang DQ, Li GQ, Chen L, Chen F, Fang DM, Yu LL, Sun YD, Ma Y, Li JJ, Wang J. The goose genome sequence leads to insights into the evolution of waterfowl and susceptibility to fatty liver. *Genome Biol*, 2015, 16(1): 89. [DOI]
- [13] Qiu Q, Zhang GJ, Ma T, Qian WB, Wang JY, Ye ZQ, Cao CC, Hu QJ, Kim J, Larkin DM, Auvil L, Capitanu B, Ma J, Lewin HA, Qian XJ, Lang YS, Zhou R, Wang LZ, Wang K, Xia JQ, Liao SG, Pan SK, Lu X, Hou HL, Wang Y, Zang XT, Yin Y, Ma H, Zhang J, Wang ZF, Zhang YM, Zhang DW, Yonezawa T, Hasegawa M, Zhong Y, Liu WB, Zhang Y, Huang ZY, Zhang SX, Long RJ, Yang HM, Wang J, Lenstra JA, Cooper DN, Wu Y, Wang J, Shi P, Wang J, Liu JQ. The yak genome and adaptation to life at high altitude. *Nat Genet*, 2012, 44(8): 946–949. [DOI]
- [14] Fang XD, Mou YL, Huang ZY, Li Y, Han LJ, Zhang YF, Feng Y, Chen YX, Jiang XT, Zhao W, Sun XQ, Xiong ZQ, Yang L, Liu H, Fan DD, Mao LK, Ren LJ, Liu CX, Wang J, Li K, Wang GB, Yang SL, Lai LX, Zhang GJ, Li YR, Wang J, Bolund L, Yang HM, Wang J, Feng ST, Li SG, Du YT. The sequence and analysis of a Chinese pig genome. *GigaScience*, 2012, 1(1): 16. [DOI]
- [15] Wong GKS, Liu B, Wang J, Zhang Y, Yang X, Zhang ZJ, Meng QS, Zhou J, Li DW, Zhang JJ, Ni PX, Li SG, Ran LH, Li H, Zhang JG, Li RQ, Li ST, Zheng HK, Lin W, Li GY, Wang XL, Zhao WM, Li J, Ye C, Dai MT, Ruan J, Zhou Y, Li YZ, He XM, Zhang YZ, Wang J, Huang XG, Tong W, Chen J, Ye J, Chen C, Wei N, Li GQ, Dong L, Lan FD, Sun YQ, Zhang ZP, Yang Z, Yu YP, Huang YQ, He DD, Xi Y, Wei D, Qi QH, Li WJ, Shi JP, Wang MH, Xie F, Wang JJ, Zhang XW, Wang P, Zhao YQ, Li N, Yang N, Dong W, Hu SN, Zeng CQ, Zheng WM, Hao BL. Genome Sequence of Red Jungle Fowl: Washington University School of Medicine, Hillier LW, Yang SP, Warren WC,

- Wilson RK, Molecular Evolution: Uppsala University, Brandström M, Ellegren H, Population Genotyping, BAC Sequences and Haplotypes: Wageningen University, Crooijmans RPMA, van der Poel JJ, Bovenhuis H, Groenen MAM, Lawrence Livermore National Laboratory, Ovcharenko I, Gordon L, Stubbs L, DOE Joint Genome Institute, Lucas S, Glavina T, Aerts A, Examples of Application to Complex Traits: Institute for Animal Health, Kaiser P, Rothwell L, Young JR, Rogers S, Walker BA, van Hateren A, Kaufman J, Bumstead N, Iowa State University, Lamont SJ, Zhou HJ, Institute R, Hocking PM, Morrice D, de Koning DJ, Law A, Bartley N, Burt DW, USDA-ARS Avian Disease and Oncology Laboratory, Henry Hunt, Cheng HH, Domestication and Selection: Uppsala University, Gunnarsson U, Wahlberg P, Andersson L, Institutet K, Kindlund E, Tammi MT, Andersson B, Human Disease Genes: University of Oxford, Webber C, Ponting CP, EST-Based SNP Data: University of Manchester Institute of Science and Technology, Overton IM, Boardman PE, Tang HZ, Hubbard SJ, University of Sheffield, Wilson SA, Scientific Management: Beijing Institute of Genomics of Chinese Academy of Sciences, Yu J, Wang J, Yang HM, International Chicken Polymorphism Map Consortium Polymorphism Discovery and Analysis: Beijing Institute of Genomics of Chinese Academy of Sciences, Watson JD, Institute of Genome Sciences of Zhejiang University. A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. *Nature*, 2004, 432(7018): 717–722. [DOI]
- [16] Rubin CJ, Zody MC, Eriksson J, Meadows JRS, Sherwood E, Webster MT, Jiang L, Ingman M, Sharpe T, Ka S, Hallböök F, Besnier F, Carlborg Ö, Bed'hom B, Tixier-Boichard M, Jensen P, Siegel P, Lindblad-Toh K, Andersson L. Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature*, 2010, 464(7288): 587–591. [DOI]
- [17] Wang MS, Li Y, Peng MS, Zhong L, Wang ZJ, Li QY, Tu XL, Dong Y, Zhu CL, Wang L, Yang MM, Wu SF, Miao YW, Liu JP, Irwin DM, Wang W, Wu DD, Zhang YP. Genomic analyses reveal potential independent adaptation to high altitude in Tibetan chickens. *Mol Biol Evol*, 2015, 32(7): 1880–1889. [DOI]
- [18] Wang MS, Zhang RW, Su LY, Li Y, Peng MS, Liu HQ, Zeng L, Irwin DM, Du JL, Yao YG, Wu DD, Zhang YP. Positive selection rather than relaxation of functional constraint drives the evolution of vision during chicken domestication. *Cell Res*, 2016, 26(5): 556–573. [DOI]
- [19] The Bovine HapMap Consortium. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science*, 2009, 324(5926): 528–532. [DOI]
- [20] Daetwyler HD, Capitan A, Pausch H, Stothard P, Van Binsbergen R, Brøndum RF, Liao XP, Djari A, Rodriguez SC, Grohs C, Esquerré D, Bouchez O, Rossignol MN, Klopp C, Rocha D, Fritz S, Eggen A, Bowman PJ, Coote D, Chamberlain AJ, Anderson C, VanTassell CP, Hulsege I, Goddard ME, Guldbrandtsen B, Lund MS, Veerkamp RF, Boichard DA, Fries R, Hayes BJ. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. *Nat Genet*, 2014, 46(8): 858–865. [DOI]
- [21] Qanbari S, Pausch H, Jansen S, Somel M, Strom TM, Fries R, Nielsen R, Simianer H. Classic selective sweeps revealed by massive sequencing in cattle. *PLoS Genet*, 2014, 10(2): e1004148. [DOI]
- [22] Qiu Q, Wang LZ, Wang K, Yang YZ, Ma T, Wang ZF, Zhang X, Ni ZQ, Hou FJ, Long RJ, Abbott R, Lenstra J, Liu JQ. Yak whole-genome resequencing reveals domestication signatures and prehistoric population expansions. *Nat Commun*, 2015, 6: 10283. [DOI]
- [23] Rubin CJ, Megens HJ, Barrio AM, Maqbool K, Sayyab S, Schwochow D, Wang C, Carlborg Ö, Jern P, Jørgensen CB, Archibald AL, Fredholm M, Groenen MAM, Andersson L. Strong signatures of selection in the domestic pig genome. *Proc Natl Acad Sci USA*, 2012, 109(48): 19529–19536. [DOI]
- [24] Li MZ, Tian SL, Jin L, Zhou GY, Li Y, Zhang Y, Wang T, Yeung CKL, Chen L, Ma JD, Zhang JB, Jiang AA, Li J, Zhou CW, Zhang J, Liu YK, Sun XQ, Zhao HW, Niu ZX, Lou PE, Xian LJ, Shen XY, Liu SQ, Zhang SH, Zhang MW, Zhu L, Shuai SR, Bai L, Tang GQ, Liu HF, Jiang YZ, Mai MM, Xiao J, Wang X, Zhou Q, Wang ZQ, Stothard P, Xue M, Gao XL, Luo ZG, Gu YR, Zhu HM, Hu XX, Zhao YF, Plastow GS, Wang JY, Jiang Z, Li K, Li N, Li XW, Li RQ. Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars. *Nat Genet*, 2013, 45(12): 1431–1438. [DOI]
- [25] Bosse M, Megens HJ, Frantz LAF, Madsen O, Larson G, Paudel Y, Duijvesteijn N, Harlizius B, Hagemeijer Y, Crooijmans RPMA, Groenen MAM. Genomic analysis reveals selection for Asian genes in European pigs following human-mediated introgression. *Nat Commun*, 2014, 5: 4392. [DOI]
- [26] Frantz LAF, Schraiber JG, Madsen O, Megens HJ, Cagan A, Bosse M, Paudel Y, Crooijmans RPMA, Larson G,

- Groenen MAM. Evidence of long-term gene flow and selection during domestication from analyses of Eurasian wild and domestic pig genomes. *Nat Genet*, 2015, 47(10): 1141–1148. [DOI]
- [27] Ai HS, Fang XD, Yang B, Huang ZY, Chen H, Mao LK, Zhang F, Zhang L, Cui LL, He WM, Yang J, Yao XM, Zhou LS, Han LJ, Li J, Sun SL, Xie XH, Lai BX, Su Y, Lu Y, Yang H, Huang T, Deng WJ, Nielsen R, Ren J, Huang LS. Adaptation and possible ancient interspecies introgression in pigs identified by whole-genome sequencing. *Nat Genet*, 2015, 47(3): 217–225. [DOI]
- [28] Yang J, Li WR, Lv FH, He SG, Tian SL, Peng WF, Sun YW, Zhao YX, Tu XL, Zhang M, Xie XL, Wang YT, Li JQ, Liu YG, Shen ZQ, Wang F, Liu GJ, Lu HF, Kantanen J, Han JL, Li MH, Liu MJ. Whole-genome sequencing of native sheep provides insights into rapid adaptations to extreme environments. *Mol Biol Evol*, 2016, 33(10): 2576–2592. [DOI]
- [29] Wright D, Boije H, Meadows JRS, Bed'hom B, Gourichon D, Vieaud A, Tixier-Boichard M, Rubin CJ, Imsland F, Hallböök F, Andersson L. Copy number variation in intron 1 of *SOX5* causes the Pea-comb phenotype in chickens. *PLoS Genet*, 2009, 5(6): e1000512. [DOI]
- [30] Imsland F, Feng CG, Boije H, Bed'Hom B, Fillon V, Dorshorst B, Rubin CJ, Liu RR, Gao Y, Gu XR, Wang YQ, Gourichon D, Zody MC, Zecchin W, Vieaud A, Tixier-Boichard M, Hu XX, Hallböök F, Li N, Andersson L. The *Rose-comb* mutation in chickens constitutes a structural rearrangement causing both altered comb morphology and defective sperm motility. *PLoS Genet*, 2012, 8(6): e1002775. [DOI]
- [31] Dorshorst B, Harun-Or-Rashid M, Bagherpoor AJ, Rubin CJ, Ashwell C, Gourichon D, Tixier-Boichard M, Hallböök F, Andersson L. A genomic duplication is associated with ectopic Eomesodermin expression in the embryonic chicken comb and two duplex-comb phenotypes. *PLoS Genet*, 2015, 11(3): e1004947. [DOI]
- [32] Eriksson J, Larson G, Gunnarsson U, Bed'hom B, Tixier-Boichard M, Strömstedt L, Wright D, Jungerius A, Vereijken A, Randi E, Andersson L, Georges M. Identification of the *Yellow Skin* gene reveals a hybrid origin of the domestic chicken. *PLoS Genet*, 2008, 4(2): e1000010. [DOI]
- [33] Dorshorst B, Molin AM, Rubin CJ, Johansson AM, Strömstedt L, Pham MH, Chen CF, Hallböök F, Ashwell C, Andersson L. A complex genomic rearrangement involving the *endothelin 3* locus causes dermal hyperpigmentation in the chicken. *PLoS Genet*, 2011, 7(12): e1002412. [DOI]
- [34] Wang ZP, Qu LJ, Yao JF, Yang XL, Li GQ, Zhang YY, Li JY, Wang XT, Bai JR, Xu GY, Deng XM, Yang N, Wu CX. An *EAV-HP* insertion in 5' flanking region of *SLCO1B3* causes blue eggshell in the chicken. *PLoS Genet*, 2013, 9(1): e1003183. [DOI]
- [35] Feng CG, Gao Y, Dorshorst B, Song C, Gu XR, Li QY, Li JX, Liu TX, Rubin CJ, Zhao YQ, Wang YQ, Fei J, Li HF, Chen KW, Qu H, Shu DM, Ashwell C, Da Y, Andersson L, Hu XX, Li N. A *cis*-regulatory mutation of *PDSS2* causes silky-feather in chickens. *PLoS Genet*, 2014, 10(8): e1004576. [DOI]
- [36] Jin SH, Zhu F, Wang YY, Yi GQ, Li JY, Lian L, Zheng JX, Xu GY, Jiao RG, Gong Y, Hou ZC, Yang N. Deletion of Indian hedgehog gene causes dominant semi-lethal Creeper trait in chicken. *Sci Rep*, 2016, 6: 30172. [DOI]
- [37] Jacobsson L, Park HB, Wahlberg P, Fredriksson R, Perez-Enciso M, Siegel PB, Andersson L. Many QTLs with minor additive effects are associated with a large difference in growth between two selection lines in chickens. *Genet Res*, 2005, 86(2): 115–125. [DOI]
- [38] Flink LG, Allen R, Barnett R, Malmström H, Peters J, Eriksson J, Andersson L, Dobney K, Larson G. Establishing the validity of domestication genes using DNA from ancient chickens. *Proc Natl Acad Sci USA*, 2014, 111(17): 6184–6189. [DOI]
- [39] Zhang Q, Gou WY, Wang XT, Zhang YW, Ma J, Zhang HL, Zhang Y, Zhang H. Genome resequencing identifies unique adaptations of tibetan chickens to hypoxia and high-dose ultraviolet radiation in high-altitude environments. *Genome Biol Evol*, 2016, 8(3): 765–776. [DOI]
- [40] Zhang GJ, Li C, Li QY, Li B, Larkin DM, Lee C, Storz JF, Antunes A, Greenwold MJ, Meredith RW, Ödeen A, Cui J, Zhou Q, Xu LH, Pan HL, Wang ZJ, Jin LJ, Zhang P, Hu HF, Yang W, Hu J, Xiao J, Yang ZK, Liu Y, Xie QL, Yu H, Lian JM, Wen P, Zhang F, Li H, Zeng YL, Xiong ZJ, Liu SP, Zhou L, Huang ZY, An N, Wang J, Zheng QM, Xiong YQ, Wang GB, Wang B, Wang JJ, Fan Y, da Fonseca R R, Alfaro-Núñez A, Schubert M, Orlando L, Mourier T, Howard JT, Ganapathy G, Pfenning A, Whitney O, Rivas MV, Hara E, Smith J, Farré M, Narayan J, Slavov G, Romanov MN, Borges R, Machado JP, Khan I, Springer MS, Gatesy J, Hoffmann FG, Opazo JC, Håstad O, Sawyer RH, Kim H, Kim KW, Kim HJ, Cho S, Li N, Huang YH, Bruford MW, Zhan XJ, Dixon A, Bertelsen MF, Derryberr E, Warren W, Wilson RK, Li SB, Ray DA, Green RE,

- O'Brien SJ, Griffin D, Johnson WE, Haussle D, Ryder OA, Willerslev E, Graves GR, Alström P, Fjeldså J, Mindell DP, Edwards SV, Braun EL, Rahbek C, Burt DW, Houde P, Zhang Y, Yang HM, Wang J, Consortium AG, Jarvis ED, Gilbert MTP, Wang J. Comparative genomics reveals insights into avian genome evolution and adaptation. *Science*, 2014, 346(6215): 1311–1320. [DOI]
- [41] Romanov MN, Farré M, Lithgow PE, Fowler KE, Skinner BM, O'Connor R, Fonseca G, Backström N, Matsuda Y, Nishida C, Houde P, Jarvis ED, Ellegren H, Burt DW, Larkin DM, Griffin DK. Reconstruction of gross avian genome structure, organization and evolution suggests that the chicken lineage most closely resembles the dinosaur avian ancestor. *BMC Genomics*, 2014, 15(1): 1060. [DOI]
- [42] Zhou Q, Zhang JL, Bachtrog D, An N, Huang QF, Jarvis ED, Gilbert MTP, Zhang GJ. Complex evolutionary trajectories of sex chromosomes across bird taxa. *Science*, 2014, 346(6215): 1246338. [DOI]
- [43] Larkin DM, Daetwyler HD, Hernandez AG, Wright CL, Hetrick LA, Boucek L, Bachman SL, Band MR, Akraiko TV, Cohen-Zinder M, Thimmapuramc J, Macleode IM, Harkinsf TT, McCagueg JE, Goddardb ME, Hayes BJ, Lewina HA. Whole-genome resequencing of two elite sires for the detection of haplotypes under selection in dairy cattle. *Proc Natl Acad Sci USA*, 2012, 109(20): 7693–7698. [DOI]
- [44] Park SDE, Magee DA, McGettigan PA, Teasdale MD, Edwards CJ, Lohan AJ, Murphy A, Braud M, Donoghue MT, Liu Y, Chamberlain AT, Rue-Albrecht K, Schroeder S, Spillane C, Tai SS, Bradley DG, Sonstegard TS, Loftus BJ, MacHugh DE. Genome sequencing of the extinct Eurasian wild aurochs, *Bos primigenius*, illuminates the phylogeography and evolution of cattle. *Genome Biol*, 2015, 16(1): 234. [DOI]
- [45] Mei CG, Wang HC, Zhu WJ, Wang HB, Cheng G, Qu KX, Guang XM, Li AN, Zhao CP, Yang WC, Wang CZ, Xin YP, Zan LS. Whole-genome sequencing of the endangered bovine species Gayal (*Bos frontalis*) provides new insights into its genetic features. *Sci Rep*, 2016, 6: 19787. [DOI]
- [46] Liang CN, Wang LZ, Wu XY, Wang K, Ding XZ, Wang MC, Chu M, Xie XY, Qiu Q, Yan P. Genome-wide association study identifies loci for the polled phenotype in yak. *PLoS One*, 2016, 11(7): e0158642. [DOI]
- [47] Larson G, Albarella U, Dobney K, Rowley-Conwy P, Schibler J, Tresset A, Vigne JD, Edwards CJ, Schlumbaum A, Dinu A, Bălăşescu A, Dolman G, Tagliacozzo A, Manaseryan N, Miracle P, Van Wijngaarden-Bakker L, Maseti M, Bradley DG, Cooper A. Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. *Proc Natl Acad Sci USA*, 2007, 104(39): 15276–15281. [DOI]
- [48] Lunney JK. Advances in swine biomedical model genomics. *Int J Biol Sci*, 2007, 3(3): 179–184. [DOI]
- [49] Li MZ, Chen L, Tian SL, Lin Y, Tang QZ, Zhou XM, Li DY, Yeung CK, Che TD, Jin L, Fu YH, Ma JD, Wang X, Jiang AA, Lan J, Pan Q, Liu YK, Luo ZG, Guo ZY, Liu HF, Zhu L, Shuai SR, Tang GQ, Zhao JG, Jiang YZ, Bai L, Zhang SH, Mai MM, Li CC, Wang DW, Gu YR, Wang GS, Lu HF, Li Y, Zhu HH, Li ZW, Li M, Gladyshev VN, Jiang Z, Zhao SH, Wang JY, Li RQ, Li XW. Comprehensive variation discovery and recovery of missing sequence in the pig genome using multiple de novo assemblies. *Genome Res*, 2016, doi: 10.1101/gr.207456.116. (in Press) [DOI]
- [50] Skinner BM, Lachani K, Sargent CA, Affara NA. Regions of XY homology in the pig X chromosome and the boundary of the pseudoautosomal region. *BMC Genet*, 2013, 14(1): 3. [DOI]
- [51] Skinner BM, Sargent CA, Churcher C, Hunt T, Herrero J, Loveland JE, Dunn M, Louzada S, Fu BY, Chow W, Gilbert J, Austin-Guest S, Beal K, Carvalho-Silva D, Cheng W, Gordon D, Grafham D, Hardy M, Harley J, Hauser H, Howden P, Howe K, Lachani K, Ellis PJI, Kelly D, Kerry G, Kerwin J, Ng BL, Threadgold G, Wileman T, Wood JMD, Yang FT, Harrow J, Affara NA, Tyler-Smith C. The pig X and Y Chromosomes: structure, sequence, and evolution. *Genome Res*, 2016, 26(1): 130–139. [DOI]
- [52] Dawson HD, Loveland JE, Pascal G, Gilbert JGR, Uenishi H, Mann KM, Sang YM, Zhang J, Carvalho-Silva D, Hunt T, Hardy M, Hu ZL, Zhao SH, Anselmo A, Shinkai H, Chen C, Badaoui B, Berman D, Amid C, Kay M, Lloyd D, Snow C, Morozumi T, Cheng RPY, Bystrom M, Kapetanovic R, Schwartz JC, Kataria R, Astley M, Fritz E, Steward C, Thomas M, Wilming L, Toki D, Archibald AL, Bed'Hom B, Beraldi D, Huang TH, Ait-Ali T, Blecha F, Botti S, Freeman TC, Giuffra E, Hume DA, Lunney JK, Murtaugh MP, Reecy JM, Harrow JL, Rogel-Gaillard C, Tuggle CK. Structural and functional annotation of the porcine immunome. *BMC Genomics*, 2013, 14(1): 332. [DOI]
- [53] Choi JW, Chung WH, Lee KT, Cho ES, Lee SW, Choi BH, Lee SH, Lim W, Lim D, Lee YG, Hong JK, Kim DW, Jeon HJ, Kim J, Kim N, Kim TH. Whole-genome resequencing analyses of five pig breeds, including Korean wild and native, and three European origin breeds. *DNA Res*, 2015,

- 22(4): 259–267. [DOI]
- [54] Miller JM, Moore SS, Stothard P, Liao XP, Coltman DW. Harnessing cross-species alignment to discover SNPs and generate a draft genome sequence of a bighorn sheep (*Ovis canadensis*). *BMC Genomics*, 2015, 16(1): 397. [DOI]
- [55] Dong Y, Zhang XL, Xie M, Arefnezhad B, Wang ZJ, Wang WL, Feng SH, Huang GD, Guan R, Shen WJ, Bunch R, McCulloch R, Li QY, Li B, Zhang GJ, Xu X, Kijas JW, Salekdeh GH, Wang W, Jiang Y. Reference genome of wild goat (*capra aegagrus*) and sequencing of goat breeds provide insight into genic basis of goat domestication. *BMC Genomics*, 2015, 16(1): 431. [DOI]
- [56] Rao M, Morisson M, Faraut T, Bardes S, Fève K, Labarthe E, Fillon V, Huang YH, Li N, Vignal A. A duck RH panel and its potential for assisting NGS genome assembly. *BMC Genomics*, 2012, 13(1): 513. [DOI]
- [57] Pérez Enciso M, Burgos-Paz W, Ramos Onsins SE. On genetic differentiation between domestic pigs and Tibetan wild boars. *Nat Genet*, 2015, 47(3): 190–192. [DOI]
- [58] Frantz LAF, Madsen O, Megens HJ, Schraiber JG, Paudel Y, Bosse M, Crooijmans RPMA, Larson G, Groenen MAM. Evolution of Tibetan wild boars. *Nat Genet*, 2015, 47(3): 188–189. [DOI]
- [59] Wang MS, Yang HC, Otecko NO, Wu DD, Zhang YP. Olfactory genes in Tibetan wild boar. *Nat Genet*, 2016, 48(9): 972–973. [DOI]
- [60] Civián P, Craig H, Cox CJ, Brown TA. Three geographically separate domestications of Asian rice. *Nature Plants*, 2015, 1(11): 15164. [DOI]
- [61] Li RQ, Li YR, Zheng HC, Luo RB, Zhu HM, Li QB, Qian WB, Ren YY, Tian G, Li JX, Zhou GY, Zhu X, Wu HL, Qin JJ, Jin X, Li DF, Cao HZ, Hu XD, Blanche H, Cann H, Zhang XQ, Li SG, Bolund L, Kristiansen K, Yang HM, Wang J, Wang J. Building the sequence map of the human pan-genome. *Nat Biotechnol*, 2010, 28(1): 57–63. [DOI]
- [62] Andersson L. Molecular consequences of animal breeding. *Curr Opin Genet Dev*, 2013, 23(3): 295–301. [DOI]
- [63] Karlsson EK, Baranowska I, Wade CM, Hillbertz NHCS, Zody MC, Anderson N, Biagi TM, Patterson N, Pielberg GR, Kulbokas EJ III, Comstock KE, Keller ET, Mesirov JP, von Euler H, Kämpe O, Hedhammar Å, Lander ES, Andersson G, Andersson L, Lindblad-Toh K. Efficient mapping of mendelian traits in dogs through genome-wide association. *Nat Genet*, 2007, 39(11): 1321–1328. [DOI]
- [64] Zhang WC, Yang B, Zhang JJ, Cui LL, Ma JW, Chen CY, Ai HS, Xiao SJ, Ren J, Huang LS. Genome-wide association studies for fatty acid metabolic traits in five divergent pig populations. *Sci Rep*, 2016, 6: 24718. [DOI]
- [65] Li H, Peng ZY, Yang XH, Wang WD, Fu JJ, Wang JH, Han YJ, Chai YC, Guo TT, Yang N, Liu J, Warburton ML, Cheng YB, Hao XM, Zhang P, Zhao JY, Liu YJ, Wang GY, Li JS, Yan JB. Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nat Genet*, 2013, 45(1): 43–50. [DOI]
- [66] Civelek M, Lusis AJ. Systems genetics approaches to understand complex traits. *Nat Rev Genet*, 2014, 15(1): 34–48. [DOI]

(责任编辑: 李明洲)