

网络出版时间:2015-10-13 08:46 DOI:10.13207/j.cnki.jnwfufu.2015.11.003  
网络出版地址:http://www.cnki.net/kcms/detail/61.1390.S.20151013.0846.006.html

# 牛全基因组测序研究进展

王洪程<sup>1</sup>,梅楚刚<sup>1</sup>,咎林森<sup>1,2</sup>,成 功<sup>1,2</sup>,李安宁<sup>1,2</sup>,王洪宝<sup>1,2</sup>

(1 西北农林科技大学 动物科技学院,陕西 杨凌 712100;2 国家肉牛改良中心,陕西 杨凌 712100)

**[摘 要]** 牛作为反刍动物的代表,具有独特的生物学特征和重要的哺乳动物进化地位,是人类生产活动的重要工具和肉、奶等物质需求的主要来源。近 10 年来,牛全基因组研究取得了很多重要的进展,为解释牛的生物学特征和加快品种的分子育种进程发挥了重要作用,文章重点介绍了牛全基因组测序的相关研究成果,以及测序工作完成后的研究进展,讨论了牛全基因组测序工作的机遇、研究重点,以及今后面临的挑战。

**[关键词]** 牛;全基因组;测序

**[中图分类号]** S823;Q78

**[文献标志码]** A

**[文章编号]** 1671-9387(2015)11-0017-07

## Progress on whole genome sequencing of bovine

WANG Hong-cheng<sup>1</sup>,MEI Chu-gang<sup>1</sup>,ZAN Lin-sen<sup>1,2</sup>,CHENG Gong<sup>1,2</sup>,  
LI An-ning<sup>1,2</sup>,WANG Hong-bao<sup>1,2</sup>

(1 College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi 712100, China;

2 National Beef Cattle Improvement Center in China, Yangling, Shaanxi 712100, China)

**Abstract:** As a representative of ruminants, cattle have unique biological characteristics and important evolutionary position compared to other mammals, acting as an important tool and sources of meat and milk to human. During last decade, great progresses have been made in cattle genome research, which plays an important role in understanding the biology and promoting the molecular breeding process. This article introduces the cattle genome sequencing and researches after sequencing, and discussed the opportunities, emphasis and challenges of future bovine genome sequencing research.

**Key words:** bovine; whole genome; sequencing

反刍动物具有独特的消化能力,可将不适于人类食用的粗纤维植物转化为肉和奶。据统计,全世界现存大约有 36 亿头家牛、水牛、绵羊和山羊等反刍家畜,为全球 66 亿人提供主要的蛋白质营养来源,全世界大约有 3/4 的农业用地生产的饲草被用于反刍动物的饲养。牛作为反刍动物的代表是最早被人类驯化利用的动物之一,在公元前 8 000—10 000 年前,从农业生活方式开始,人类就开始了

牛的驯化,除了能获取大量的食物和皮毛,而且还将牛用于耕作、运输等活动。如今,全世界大约有 15 亿头牛,根据不同的经济、社会和宗教需要,人类共选育出 800 多个牛品种(种群),成为重要的世界遗传资源<sup>[1-4]</sup>。我国幅员辽阔,地形复杂,几千年的农耕文明历史过程中,人们在不同的自然环境中育成了不同的地方黄牛品种。

1977 年, Sanger 等<sup>[5]</sup> 发明了双脱氧链终止法

**[收稿日期]** 2014-03-26

**[基金项目]** “十二五”国家“863”计划项目(2013AA102505,2011AA100307-02);国家自然科学基金项目(31272411);“十二五”国家科技支撑计划项目(2011BAD28B04-03);“十二五”国家转基因育种重大专项(2013ZX08007-002);国家肉牛牦牛产业技术体系建设专项(CARS-38)

**[作者简介]** 王洪程(1985—),男,山东临沂人,在读博士,主要从事基因组学研究。E-mail:besthongcheng@163.com

**[通信作者]** 咎林森(1963—),男,陕西扶风人,教授,博士,主要从事动物遗传育种与繁殖研究。E-mail:zanlinsen@163.com

(Chain Termination Method),这标志着基因组学研究的开端,此后科学家们相继完成了人类<sup>[6]</sup>、水稻<sup>[7]</sup>等物种的基因组测序。经过持续的技术改良和革新,各类测序手段和平台不断提升基因组测序的速度和精确性。2005年,Roche公司发明的454测序平台标志着基因组测序进入了新时代,熊猫<sup>[8]</sup>成为第一次利用新一代测序技术完成基因组从头测序的大型动物。牛作为反刍动物的代表,具有独特的生物学特性,其进化地位异于人和啮齿动物<sup>[9]</sup>,这是其成为最早进行测序的几个哺乳动物之一的重要原因。本文将介绍牛全基因组测序的发展历程,概述牛全基因组测序的后续研究,讨论牛全基因组测序工作的研究重点,以及今后面临的挑战。

## 1 牛全基因组 *de novo* 测序

*de novo* 来源于拉丁文,意为“重新、开始”。全基因组 *de novo* 测序,顾名思义为从头测序,是直接对某个物种进行测序,利用生物信息学将短片段序列拼接和组装成一个完整的基因组序列,其全过程不参考本物种现有序列资料或者其他物种的基因组。全基因组从头测序对研究物种的基因组结构和功能基因信息,阐明和解释物种的进化及其生物学特性具有重要的意义<sup>[10]</sup>。

2003年12月,由美国、加拿大、澳大利亚和新西兰等国家参与的“牛基因组测序工程”正式启动,经过6年的努力,于2009年完成了世界上第一头牛的全基因组序列<sup>[11]</sup>。“牛基因组测序工程”选择的品种是海福特牛(Hereford,普通家牛, *Bos taurus*),测序采用的是BAC克隆测序和全基因组鸟枪法(Whole-genome shotgun, WGS)测序策略,构建WGS插入片段文库的DNA样品来自一头海福特母牛(编号L1 Dominette 01449),而BAC文库的DNA样品则为该牛的父亲(编号L1 Domino 99375),序列组装方法与之前完成的小鼠<sup>[12]</sup>和海胆<sup>[13]</sup>的序列组装方法类似。测序深度为7×,BAC末端序列用在了基因组 scaffold 组装中,总序列长度为2.98 Gb,其中 contig 的 N50 长度(N50 长度,将 contig 或 scaffold 从大到小排序,并对其长度进行累加,当累加长度达到基因组序列长度一半时,最后一个 contig 或 scaffold 长度即为 N50 长度,是评估基因组组装质量的重要标志,该数值越大,表明基因组组装质量越好)是82.96 kb, scaffold 的 N50 长度为2.59 Mb,104万个表达序列标签(EST)中有95%包含在组装的序列中,估算基因组大小为2.87

Gb。研究小组对组装结果不断更新优化,并先后在NCBI数据库公布了6个版本的基因组序列,最新的为Btau 4.6.1,包括29条常染色体及2条性染色体。研究发现,牛基因组至少包含22 000个编码蛋白基因,从中鉴定出的496个microRNA有135个为新发现。牛基因组的测序成功揭示出了反刍动物很多的生物学特性,例如研究者发现,人类的5个与脂肪酸、甲戊二羧酸、解毒、嘧啶代谢途径相关的基因在牛基因组中缺失或者变异,在牛基因组中有10种C型溶菌酶基因,且EST文库内6/7的C型溶菌酶基因都在瘤胃和肠中表达。

2009年,由美国马里兰大学牵头组织的牛基因组研究团队也发表了关于牛基因组测序的研究成果<sup>[14]</sup>。该团队利用“牛基因组测序工程”测序数据,通过改进组装方法,拼接出了新的牛基因组序列。该团队优化组装的最新牛基因组版本为UMD\_3.1.1,其总序列长度为2.67 Gb, contig 的 N50 长度为96.95 kb, scaffold 的 N50 长度为6.38 Mb,估算基因组大小为2.86 Gb。UMD版本覆盖率、完整性和注释效果等都比Btau版本更优,并且其单核苷酸差异性错误率更低。由于DNA来源问题,当时公布的Btau 4.0版本的基因组序列不包含Y染色体,而UMD基因组通过鉴定筛选出了更多Y染色体的 contig,并完成了Y染色体序列的较为完整的组装。该研究团队还构建了新的人-牛基因组线性图谱,并在该图谱中发现了268个同源线性区域。

家牛分为普通家牛(*Bos taurus*)和瘤牛(*Bos indicus*),瘤牛颈肩隆起的峰瘤是其区别于普通家牛最明显的特征<sup>[15]</sup>,瘤牛主要分布于印度、巴西、美国南部、澳洲北部和中国的南部,具有耐热和抗虫性。2012年,巴西热那亚生物公司联合巴西和加拿大高校<sup>[16]</sup>利用ABI SOLiD测序平台对一头巴西内洛尔公牛(Nellore)进行了52×的测序研究,建立了0.11~2 kb片段长度不同的6个文库,得到4.0×10<sup>9</sup>个reads,比对Btau 4.0和UMD 2.0基因组后分别得到319 222、372 804个contigs,相应覆盖参考基因组的93%和99.1%,比较普通家牛和瘤牛基因组的编码蛋白基因发现,瘤牛缺失存在于普通家牛中的抵抗素(beta-defensin)及另一假想蛋白,研究表明,如果亲缘物种基因组存在的话,利用DNA小片段建库(0.11~2 kb)对某一物种测序,也能得到非常好的组装效果,这对今后的测序工作有重要的借鉴意义。

牦牛(yak, *Bos grunniens*)是青藏高原等高海拔

地区特有的物种,具有很多适应高原生活的生理结构和习性,大约有 1 400 万头牦牛为生活在这些地区的居民提供着基本的生活资源,被称作“高原之舟”。2012 年,兰州大学联合华大基因公司等科研单位对牦牛进行测序并成功组装基因组<sup>[17]</sup>。科研人员从一头母牦牛上采集 DNA 样品,并采用全基因组鸟枪法策略,利用 Illumina HiSeq 2000 测序平台进行双末端测序,构建出牦牛全基因组序列图谱,测序深度为 65×。牦牛基因组的 contig 和 scaffold N50 长度分别为 20.4 kb 和 1.4 Mb,基因组大小约为 2.66 Gb,该项目组还利用 Sanger 测序技术对牦牛 6 个线粒体基因进行了测序,并对 5 个组织样品进行了转录组测序。基于转录组测序以及同源性、全基因组预测,共检测到 22 282 个编码蛋白基因、220 万个杂合单核苷酸变异。同时,研究人员还利用人基因组为外围序列,构建出了 207 个牦牛祖先共线性同源区域,大约覆盖整个基因组的 94%。研究人员通过分析牦牛、普通牛、人和狗等 4 个物种的 13 810 个同源基因家族发现,牦牛和普通牛共有 362 个基因家族,还发现 100 个与嗅觉、防御、免疫有关的基因家族为牦牛所特有。通过检测正向选择的基因发现,3 个正向选择基因包含的 2 个调控子和 1 个目标基因,均与血管生成、扩张和能量代谢有重要关系,5 个调控牦牛营养通路的关键基因也受到正向选择,有助于牦牛适应高原以及饲草缺乏的恶劣环境。牦牛基因组的破译不仅揭示了牦牛对高原适应性的遗传机理,也促进了人类对高原反应及缺氧相关疾病的认识、预防和治疗。研究结果对加快奶用、肉用牦牛育种进程,提高当地群众经济收入有重要意义。

最近,孟加拉 Lal Teer 畜牧科技公司与华大基因公司联合宣布水牛(*Bos bubalus*)基因组图谱绘制完成<sup>[18]</sup>,其基因组大小约为 2.77 Gb,略小于人类基因组,共有 21 550 个编码基因。该项科研成果将有利于从基因组水平和遗传学角度认识水牛的物种起源、驯化过程和生物学特性,推进水牛品种的选育工作。

随着测序技术和生物信息学的飞速发展,普通家牛(*Bos taurus*)、瘤牛(*Bos indicus*)、牦牛(*Bos grunniens*)、水牛(*Bos bubalus*)的基因组序列均被组装成功,但各个基因组序列还在不断完善过程中,例如普通牛基因组就由不同研究团队组装得到两个版本,基因组序列的覆盖度也不断得到提高。一个物种基因组计划的完成,意味着这一物种学科和产

业发展的新开端,不同种属牛基因组的破译对于人们进一步研究牛不同的生理结构、性状特点和生物学特性奠定了基础,有利于根据不同的需求和特点,加快牛品种选育,使其更好地服务于人类。

## 2 牛全基因组重测序

全基因组重测序是指对已有参考基因组的物种进行不同个体的基因组测序,通过生物信息学的比对方法,可以检测到大量与性状关联的遗传变异信息(包括单核苷酸多态性位点、插入缺失位点、结构变异位点和拷贝数变异等)。通过变异信息注释、系统进化和群体遗传结构分析等方法,能够揭示物种的进化历史、环境适应性、自然选择等特征,有利于缩短分子育种周期。

随着牛全基因组序列的公布,牛全基因组重测序研究取得了较大进展。2009 年,牛基因组图谱研究团队利用 6 个品种牛的序列文库与海福特牛参考基因组比对得到的探针,检测到 19 个不同品种 497 头牛的 37 470 个 SNPs<sup>[19]</sup>;发现普通家牛品种的平均最小等位基因频率(Minor allele frequencies, MAFs)比瘤牛高。研究发现,基于个体间的系谱和共享基因位点来比较个体祖先时,即使在系谱未知的情况下也能准确预测出祖先群体大小,这是很好的保护濒危灭绝牛群体的方法。研究者还分析了牛品种间的亲缘关系和连锁不平衡(LD)模型,发现非洲的 N'Dama 牛和 Sheko 牛均起源于欧洲牛,非洲的 N'Dama 牛的历史群体规模很小,说明该群体没有受到强烈的驯化瓶颈。普通家牛的历史群体要比瘤牛大得多,都经历了由远古时代的大群体到当代群体迅速下降的一个过程。

2009 年,德国一研究小组利用新一代测序平台 Illumina Genome Analyzer II,对一头弗莱维赫公牛(Fleckvieh)进行了深度为 7.4×的重测序,得到 24 G 的测序数据,检测到了  $2.44 \times 10^6$  个 SNPs(其中 82%是未知 SNPs)以及  $1.15 \times 10^5$  个 indels,为评估测序数据的准确性,该研究利用 50 K 基因芯片对同一头牛进行了基因型比较,发现纯合子和杂合子的检测率分别为 74%和 30%,比较随机选择的 196 个基因型得到的假阳性率(False positive rate)为 1.1%,利用 48 头弗莱维赫牛和 48 头瑞士褐牛(Braunvieh)检测这 196 个 SNPs 的等位基因频率发现,95%的 SNPs 呈多态性,平均最小等位基因频率(Minor allele frequency)是 24.5%,83%的 SNPs 的最小等位基因频率大于 5%,可用于关联分析研

究<sup>[20]</sup>。这是第一次利用二代测序技术开展牛基因组测序研究,利用中低深度的重测序技术得到 200 多万的新 SNPs,进一步丰富了现在的 SNPs 数据库,为构建高密 SNP 芯片,开展全基因组关联研究提供了有价值的资源。

2011 年,日本科学家对一头日本当地牛(Kuchinoshima-Ushi)进行了重测序研究,评估了这个牛品种的 SNPs 等遗传特性,研究人员利用 Illumina GA II 测序平台共得到 64.2 G 的测序数据,测序深度达到  $15.8\times$ ,其中 86% 的 reads 比对到参考基因组序列上(Btau 4.0),可以覆盖 93% 的基因组,共检测到 630 万个 SNPs(其中 550 万个为新发现)以及约 69 万个插入缺失信息,除已经在牛基因组注释过的 SNPs 外,该研究在蛋白编码区还发现了 20 432 个 SNPs,包括分布在 4 643 个基因中的 11 713 个非同义 SNPs,该研究通过基因聚类发现,含有非同义 SNP 数量最多的 100 个基因大多与蛋白锚定、活性催化、代谢通路等有关<sup>[21]</sup>。而已有相关报道证明,在 Kuchinoshima-Ushi 牛发现的 SNP 与其他品种牛的表型性状有关联<sup>[22-25]</sup>。另外,该研究通过对 10 个基因序列进行系统发育分析,发现 Kuchinoshima-Ushi 牛明显不同于欧洲家养牛品种。为进一步开展 Kuchinoshima-Ushi 牛群体遗传研究和含有 SNP 位点基因功能的研究提供了框架,有助于探究重要经济性状表型变异的分子机制,改善牛的内在缺陷。

Stothard 等<sup>[26]</sup>采用 SOLiD system 测序平台对一头黑安格斯牛和一头美国荷斯坦牛进行了深度为  $22\times$  和  $19\times$  的全基因组重测序,通过比对分别得到 320 万和 370 万个 SNPs,其中 81% 和 75% 为新发现,24% 为共有。另外,该研究检测出 790 个拷贝数变异(Copy number variations, CNVs),注释发现,黑安格斯牛基因组中含有更多的与运动相关的 CNVs,荷斯坦奶牛中含有更多与哺乳、激活酶等繁殖功能相关的 CNVs,注释结果与动物各自的优秀性状保持一致。一些 CNVs 与牛肉和乳制品(如产奶、健康或者肉品质)相关基因重叠,且荷斯坦牛的 CNVs 更丰富,推测可能与选择强度有关。

CNVs 是某个物种 2 个个体基因组序列中 50 bp 以上片段的增加或者减少<sup>[27]</sup>,尽管 SNPs 在全基因组水平中很常见,但是 CNVs 对序列长度的影响比较大,对改变基因结构、数量及基因调控和暴露隐性等位基因等有着更重要的潜在影响<sup>[28]</sup>。Bickhart 等<sup>[29]</sup>用 Illumina GA II x 第二代测序技术检测了 5

头普通家牛(3 头安格斯、1 头荷斯坦牛、1 头海福特牛)和 1 头瘤牛(Nelore)全基因组 CNVs 的差异,在比对序列中共检测到了 1 265 个 CNV 区域,其中有 476 个(38%)是首次发现,研究者检测了变异最多的 25 个基因,发现瘤牛个体中 13 个基因拷贝数较低,而如 CATHL4、ULBP17 等抗病基因则在瘤牛中高度复制,而在普通家牛中与脂质运输和代谢相关的基因(APOL3 和 FABP2)高度复制,这些结果表明 CNVs 与牛品种的适应性、健康、生产性状之间的差异有关联,这与 Stothard 等<sup>[26]</sup>的研究结果一致。Bickhart 等<sup>[29]</sup>还构建了第一张牛个体 CNV、重复片段图谱和估算全基因组 CNVs,为将来在全基因组高重复片段中进一步研究 CNV 奠定了基础。

通过全基因组群体水平重测序可以寻找到大部分的有效突变。尽管近几年二代测序费用大幅度下降,但是大规模重测序费用依然难以承受。人类千人基因组计划对 179 个样本进行了平均深度为  $3.6\times$  的测序,也得到了绝大多数的突变信息<sup>[30]</sup>。基于此,Jansen 等<sup>[31]</sup>选择可代表群体 69% 的遗传多样性的 43 头弗莱维赫牛进行重测序研究,测序平台为 Illumina GA II x 和 HiSeq2000,测序深度从  $4.17\times$  到  $24.98\times$ ,平均深度为  $7.46\times$ ,通过与参考基因组(UMD3.1)进行比对,共检测到约  $1.7\times 10^7$  个遗传突变,其中 67.95% 为新发现,新发现的突变中有 90% 为等位基因的 SNPs,10% 为 indels,该研究发现在 18 444 个基因的编码区存在 91 733 个突变,其中 46% 为非同义突变,在这些非同义突变中有 575 个突变预测为提前终止密码子,该研究表明,基于高密度基因芯片分型的测序的敏感度和特异性分别为 92% 和 81%,如果填补数据过程的话,两者能达到 97% 和 93%,基因型填补在动物低覆盖度的重测序中能够显著提高基因型质量,也为群体重测序提供了新的策略。

韩牛(Hanwoo)是韩国的牛品种,具有普通家牛和瘤牛血统,由中国北方迁徙到朝鲜半岛,有 5 000 多年的使役历史<sup>[32-33]</sup>。韩国研究团队用 ABI SOLiD 平台对一头韩牛公牛进行了深度为  $45.6\times$  的测序,通过与 Btau 4.0 基因组比对得到  $4.7\times 10^6$  个 SNPs,其中 58% 为新发现, $4.0\times 10^5$  个 indels 中 87% 为新发现,通过牛基因芯片(BovineSNP50)分型结果评估发现,SNPs 检测结果一致性达 96.2%<sup>[9]</sup>。该研究同时对文献[26]中黑安格斯和荷斯坦牛的测序数据进行了比较分析,利用 NCBI 数据库中的 20 955 个基因注释,在韩牛 8 360 个基因

中得到 25 000 个非同义 SNPs、剪贴变异体、编码 indels(non-synonymous SNPs, splice-site variants, and coding indels, NS/SS/Is), 多于黑安格斯牛和荷斯坦牛, 说明韩牛遗传距离要远于后两者<sup>[34]</sup>。在研究的 3 个品种中, 10 906 个基因有 NS/SS/Is, 737 个基因有 10 个以上的 NS/SS/Is, 研究表明, 含有多个 NS/SS/Is 的基因是为了适应环境而进化成的多拷贝基因, 或者是由于错误的参考基因组而形成的。检测纯合子区域(Regions of homozygosity, ROH)发现, 韩牛 ROHs 的长度和数量都不及黑安格斯牛和荷斯坦牛, 推测是由于黑安格斯牛和荷斯坦牛经历了更长的选育时间而致。在韩牛、黑安格斯牛和荷斯坦牛的 ROHs 中分别检测到 753、1 320 和 2 482 个基因, 通过聚类发现, 在 ROHs 中的基因与牛的生物特性和外貌(抗病性、肉质、黑白毛色)有相关性<sup>[35-38]</sup>, ROHs 的利用为开展牛遗传改良工作提供了一个很好的全基因组选择策略。

### 3 牛全基因组测序研究面临的机遇和挑战

近 10 年来, 第二代高通量测序技术发展迅猛, 国际上主要的几家测序公司 454 LifeScience、Illumina、Life Technologies、Pacific Biosciences 在不断地进行技术变革, 使检测成本不断降低, 而测序通量和读长都得到明显改善。如 Illumina 公司发布的 HiSeq X Ten 测序系统, 能够以千元成本测序完整的人类基因组, 每天最多可以对  $6 \times 10^{11}$  bp 进行测序, 产量增加的同时, 成本直线下降; 而 Pacific Biosciences 公司的 PacBio 测序读长已经达到 8.5 kb, 极大地方便了后续的序列拼接、组装以及注释等生物信息学分析。此外, 采用单分子读取技术的第三代测序技术已经发展起来<sup>[39]</sup>, 增加了测序读长和通量, 并且无 DNA 扩增环节, 极大地降低了测序成本, 在全基因组测序<sup>[40]</sup>和重测序<sup>[41]</sup>方面都得到了很好的应用。新一代测序技术以其高通量、低成本、高效率的特点, 已然成为科学家探究各类生物基因组奥秘的重要工具, 也必将为牛全基因组测序研究带来更广阔的发展空间和新的机遇。

牛是最具代表性的反刍动物之一, 其遗传资源非常丰富, 世界上有 800 多个牛品种, 仅中国就拥有 52 个地方黄牛品种, 牛全基因组测序工作未来将主要集中在以下几点: (1) 在从头测序方面, 家牛、牦牛和水牛基因组序列已经破译, 而野牛、大额牛等特殊牛品种的全基因组信息还属未知, 其基因组信息的

解析对全面研究牛的物种起源进化和探究相关物种特异性具有重要作用。(2) 在全基因组重测序方面, 研究人员已经开展了大量工作, 在群体水平上研究了牛的进化历史、环境适应性和生物特异性等, 并且在不同群体发现了高密度的 SNPs、indels 和 SVs (Structural variations) 等变异信息, 但对很多牛品种核心种质的重测序工作还未开展。世界不同地区分布的很多地方牛品种保存了牛物种的很多重要遗传多样性, 开展核心种质资源重测序对牛遗传资源保护及特殊种质资源的利用具有重要科学意义。(3) 开展全基因组关联分析 (Genome-wide association study, GWAS), 随着测序技术的发展, 基于全基因组测序的 GWAS 研究得到普及和广泛应用, 可以筛选、鉴定与牛重要经济性状的相关基因和位点, 开展全基因组选择, 缩短分子育种的试验周期。

高通量测序技术为基因组学研究提供了一个高效的新平台和巨大的发展机遇, 先进的测序技术不断产生海量的测序数据, 如何充分挖掘隐藏在原始数据中的生物学信息<sup>[42]</sup>, 并据此解释许多复杂的生物学现象和生理机制, 以及应用已知基因组信息促进牛品种保护和选育工作, 是今后全基因组研究的难点和最重要的挑战。

### 参考文献

- [1] FAO. FAO statistical yearbooks world food and agriculture [M/OL]. Rome: Food and Agriculture Organization of the United Nations (2013) [2015-05-05]. <http://issuu.com/faoofttheun/docs/syb2013issuu>.
- [2] Willham R L. From husbandry to science: A highly significant facet of our livestock heritage [J]. Journal of Animal Science, 1986, 62: 1742-1758.
- [3] Diamond J. Guns, germs and steel: The fates of human societies [M]. New York: WW Norton & Company, 1997.
- [4] Mason I L. A mason's world dictionary of livestock breeds, types and varieties [M]. UK: Oxford Univ Pr, 2002.
- [5] Sanger F, Nicklen S, Coulson A R. DNA sequencing with chain-terminating inhibitors [J]. Proceedings of the National Academy of Sciences, 1977, 74(12): 5463-5467.
- [6] Venter J C, Adams M D, Myers E W, et al. The sequence of the human genome [J]. Science, 2001, 291(5507): 1304-1351.
- [7] Goff S A, Ricke D, Lan T H, et al. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*) [J]. Science, 2002, 296(5565): 92-100.
- [8] Li R, Fan W, Tian G, et al. The sequence and de novo assembly of the giant panda genome [J]. Nature, 2009, 463(7279): 311-317.
- [9] Lee K T, Chung W H, Lee S Y, et al. Whole-genome resequencing of Hanwoo (Korean cattle) and insight into regions of

- homozygosity [J]. BMC Genomics, 2013, 14(1): 519.
- [10] Yue G D, Gao Q, Luo L H, et al. The application of high-throughput sequencing technology in plant and animal research [J]. Scientia Sinica Vitae, 2012, 42: 107-124.
  - [11] Elsik C G, Tellam R L, Worley K C. The genome sequence of taurine cattle: A window to ruminant biology and evolution [J]. Science, 2009, 324(5926): 522-528.
  - [12] Gibbs R A, Weinstock G M, Metzker M L, et al. Genome sequence of the Brown Norway rat yields insights into mammalian evolution [J]. Nature, 2004, 428(6982): 493-521.
  - [13] Sodergren E, Weinstock G M, Davidson E H, et al. The genome of the sea urchin *strongylocentrotus purpuratus* [J]. Science, 2006, 314(5801): 941-952.
  - [14] Zimin A V, Delcher A L, Florea L, et al. A whole-genome assembly of the domestic cow, *Bos taurus* [J]. Genome Biol, 2009, 10(4): R42.
  - [15] Bradley D G, MacHugh D E, Cunningham P, et al. Mitochondrial diversity and the origins of African and European cattle [J]. Proceedings of the National Academy of Sciences, 1996, 93(10): 5131-5135.
  - [16] Canavez F, Luche D D, Stothard P, et al. Genome sequence and assembly of *Bos indicus* [J]. Journal of Heredity, 2012, 103(3): 342-348.
  - [17] Qiu Q, Zhang G, Ma T, et al. The yak genome and adaptation to life at high altitude [J]. Nature Genetics, 2012, 44(8): 946-949.
  - [18] BGI. 孟加拉 Lal Teer 畜牧科技公司联合华大基因成功绘制水牛基因组图谱助力增强肉类及奶制品质量及安全 [EB/OL]. (2014-01-24) [2015-05-05] [http://www.genomics.cn/news/show\\_news?nid=99905](http://www.genomics.cn/news/show_news?nid=99905).
  - [19] Bovine HapMap Consortium. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds [J]. Science, 2009, 324(5926): 528-532.
  - [20] Eck S H, Benet-Pagès A, Flisikowski K, et al. Whole genome sequencing of a single *Bos taurus* animal for single nucleotide polymorphism discovery [J]. Genome Biol, 2009, 10(8): R82.
  - [21] Kawahara-Miki R, Tsuda K, Shiwa Y, et al. Whole-genome resequencing shows numerous genes with nonsynonymous SNPs in the Japanese native cattle Kuchinoshima-Ushi [J]. BMC Genomics, 2011, 12(1): 103.
  - [22] Hoashi S, Hinenoya T, Tanaka A, et al. Association between fatty acid compositions and genotypes of FABP4 and LXR- $\alpha$  in Japanese blackcattle [J]. BMC Genet, 2008, 9: 84.
  - [23] Jiang Z, Michal J J, Tobey D J, et al. Comparative understanding of UTS2 and UTS2R genes for their involvement in type 2 diabetes mellitus [J]. Int J Biol Sci, 2008, 4: 96-102.
  - [24] Gill J L, Bishop S C, McCorquodale C, et al. Association of selected SNP with carcass and taste panel assessed meat quality traits in a commercial population of Aberdeen Angus-sired beef cattle [J]. Genet Sel Evol, 2009, 41: 36.
  - [25] Pant S, Schenkel F, Leyva-Baca I, et al. Identification of single nucleotide polymorphisms in bovine CARD15 and their associations with health and production traits in Canadian Holsteins [J]. BMC Genomics, 2007, 8: 421.
  - [26] Stothard P, Choi J W, Basu U, et al. Whole genome resequencing of black Angus and Holstein cattle for SNP and CNV discovery [J]. BMC Genomics, 2011, 12(1): 559.
  - [27] Mills R E, Walter K, Stewart C, et al. Mapping copy number variation by population-scale genome sequencing [J]. Nature, 2011, 470(7332): 59-65.
  - [28] Zhang F, Gu W, Hurles M E, et al. Copy number variation in human health, disease, and evolution [J]. Annual Review of Genomics and Human Genetics, 2009, 10: 451-481.
  - [29] Bickhart D M, Hou Y, Schroeder S G, et al. Copy number variation of individual cattle genomes using next-generation sequencing [J]. Genome Research, 2012, 22(4): 778-790.
  - [30] 1 000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing [J]. Nature, 2010, 467(7319): 1061-1073.
  - [31] Jansen S, Aigner B, Pausch H, et al. Assessment of the genomic variation in a cattle population by re-sequencing of key animals at low to medium coverage [J]. BMC Genomics, 2013, 14(1): 446.
  - [32] Lee C, Pollak E J. Genetic antagonism between body weight and milk production in beef cattle [J]. Journal of Animal Science, 2002, 80(2): 316-321.
  - [33] Han S W. The breed of cattle [M]. 1st ed. Seoul: Sunjin publishing, 1996: 148-160.
  - [34] Decker J E, Pires J C, Conant G C, et al. Resolving the evolution of extant and extinct ruminants with high-throughput phylogenomics [J]. Proceedings of the National Academy of Sciences, 2009, 106(44): 18644-18649.
  - [35] Marquez B, Ameye G, Vallet C M, et al. Characterization of Abcc4 gene amplification in stepwise-selected mouse J774 macrophages resistant to the topoisomerase II inhibitor ciprofloxacin [J]. PloS One, 2011, 6(12): e28368.
  - [36] Huq M D M, Tsai N P, Lin Y P, et al. Vitamin B<sub>6</sub> conjugation to nuclear corepressor RIP140 and its role in gene regulation [J]. Nature Chemical Biology, 2007, 3(3): 161-165.
  - [37] Brandes R, Arad R, Bar-Tana J. Inducers of adipose conversion activate transcription promoted by a peroxisome proliferators response element in 3T3-L1 cells [J]. Biochemical Pharmacology, 1995, 50(11): 1949-1951.
  - [38] Kühn C, Weikard R. An investigation into the genetic background of coat colour dilution in a Charolais  $\times$  German Holstein F<sub>2</sub> resource population [J]. Animal Genetics, 2007, 38(2): 109-113.
  - [39] 张得芳, 马秋月, 尹佟明, 等. 第三代测序技术及其应用 [J]. 中国生物工程杂志, 2013, 33(5): 125-131.  
Zhang D F, Ma Q Y, Yin T M, et al. The third generation sequencing technology and its application [J]. China Biotechnology, 2013, 33(5): 125-131. (in Chinese)
  - [40] Perry G H, Reeves D, Melsted P, et al. A genome sequence resource for the aye-aye (*Daubentonia madagascariensis*), a

nocturnal lemur from Madagascar [J]. Genome Biology and Evolution,2012,4(2):126-135.

[41] Harris T D,Buzby P R,Babcock H,et al. Single-molecule DNA sequencing of a viral genome [J]. Science,2008,320(5872):106-109.

[42] 张全芳,李 军,范仲学,等. 高通量测序技术在农业研究中的应用 [J]. 山东农业科学,2013,45(1):137-140.

Zhang Q F,Li J,Fan Z X,et al. Application of high-throughput sequencing technology in agricultural research [J]. Shandong Agricultural Sciences,2013,45(1):137-140. (in Chinese)

(上接第 8 页)

[35] 杨丽萍,杨慧荣,张 勇,等. 中国大鲵脑 cDNA 文库构建及促甲状腺激素  $\beta$  亚基基因 cDNA 的克隆和序列分析 [J]. 水产学报,2008,32(4):507-516.

Yang L P,Yang H R,Zhang Y,et al. Construction of cDNA library from the brain of Chinese giant salamander *Andrias davidianus* with cloning and sequence analysis of the cDNA encoding thyroid-stimulating hormone  $\beta$  subunit [J]. Journal of Fisheries of China,2008,32(4):507-516. (in Chinese)

(上接第 16 页)

[35] Sacco A G. Antigenic cross-reactivity between human and pig zona pellucida [J]. Biology of Reproduction,1977,16:164-173.

[36] Sacco A G,Pierce D L,Subramanian M G,et al. Ovaries remain functional in squirrel monkeys (*Saimiri sciureus*) immunized with porcine zona pellucida 55 000 macromolecule [J]. Biology of Reproduction,1987,36:481-490.

[37] 张 浩. 草兔卵透明带 3(ZP3)基因的克隆分析及原核表达初探 [D]. 陕西杨凌:西北农林科技大学,2013.

Zhang H. Cloning, sequence analysis and prokaryotic expression of hare (*Lepus capensis*) zona pellucida 3 [D]. Yangling, Shaanxi:Northwest A&F University,2013. (in Chinese)

[38] Bleil J D,Wassarman P M. Identification of a ZP3-binding protein on acrosome-intact mouse sperm by photoaffinity crosslinking [J]. Proceedings of the National Academy of Sciences of the United States of America,1990,87(14):5563-5567.