

附件 4 吉林大学本科毕业论文（设计）开题报告

学 院	动物科学学院	专业	动物科学
学生姓名	朱明远	（教）学号	85150226
指导教师	赵志辉	职称	教授
合作导师		职称	
论文题目	比较基因组学揭示家鸡黑白羽色变异的遗传学基础		
<p>一、课题研究的背景和意义（综述国内外相关研究现状，阐述课题的研究目的、意义）</p> <p>随着人类生产方式由狩猎向定居转变，在几千年的过程中超过 40 余种物种被驯化，它们在驯化过程中获得了较祖先更多的表型多样性。这其中，鸡作为鸟类中多样性最多的代表，在羽毛颜色、个体大小、视觉等方面有较大的分化。传统遗传学多采用 QTL 定位，工作繁琐且假阳性率高，定位的区域往往很大，无法精确定位。但随着下一代测序技术的飞速发展，从基因组水平鉴定群体遗传多样性的群体基因组学迅速发展起来。群体基因组学是在群体遗传学的基础上，基于测序技术以及基因组图谱，应用统计学和数学原理在群体规模上对全基因组进行研究，探究基因频率和基因型频率的位点特异性效应以及迁移和遗传漂变等与遗传结构的关系的全基因组效应，探究背后的选择效应和突变作用。在重测序的基础上，鸡的许多多样性已经被解释，例如藏鸡的高原机制以及体型的大小多样性，清楚解释了鸡驯化过程中的遗传学基础。元宝鸡是一种著名的观赏名鸡，它有着的金元宝外观，从唐代以来便进行培育。元宝鸡是优良生物模型，尤其是小体型和纯色羽毛等表型被广泛应用于研究之中。为了更好地理解影响家畜羽毛颜色变化的突变，我们用的基于 NGS 的比较群体基因组学研究了元宝鸡的基因组。我们将确定一系列可能影响羽色的基因突变以及区域，为在家鸡中的色素沉积模式提供新的生物学见解。</p>			
<p>二、课题研究已有的工作基础，附证书、报告、文献翻译（总结归纳本人的学习、科研、实习等成果，以及已掌握的前人资料，简述自己初步的学术见解）</p> <p>工作基础：本人经过在吉林大学动物科学学院的 4 年学习，掌握了动物生化、生理、遗传学、生物统计等生命科学知识和实验基础。假期于中科院昆明动物所、</p>			

中科院广州生物医药与健康研究中心进行长时间实习,积累了大量生物信息学相关知识。

文献翻译:

比较群体基因组学揭示了家养鸡体型变异的遗传基础

Ming-Shan Wang^{1,2}, Yong-Xia Huo^{1,3}, Yan Li⁴, Newton O. Otecko^{1,2}, Ling-Yan Su^{2,5}, Hai-Bo Xu^{1,2}, Shi-Fang Wu^{1,2}, Min-Sheng Peng^{1,2}, He-Qun Liu^{1,2}, Lin Zeng^{1,2}, David M. Irwin^{1,6,7}, Yong-Gang Yao^{2,5}, Dong-Dong Wu^{1,2,*}, and Ya-Ping Zhang^{1,2,4,*}

前 言

家养动物作为优秀的生物学模型,广泛用于发育生物学、表型进化和医学研究。它们被驯化为不同品种,在形态学、生理学、行为和适应方面表现出显著差异。作为一种家养动物的经济性状,体型对人类和人类文明的发展极为重要。在家养动物中可以看到大量的体型变异,远远高于其野生祖先。因此育种研究者、进化和发育生物学家们都对探究控制家畜体型变异遗传的调控基础机制感兴趣。

家鸡在鸟类中拥有最多的表型变异。例如,矮脚鸡和浦东鸡是两种极端,这两个神奇的鸡种成年体重分别为 0.5 和 5 千克左右。作为全球分布最广泛且数量最多的家养动物,家鸡也被用于遗传和医学研究。已经被证实的多个性状的突变,特别是体型,在研究和育种中具有重要意义-已经绘制了数百个数量性状基因座并报告其与鸡的生长和体重相关。尽管取得了这些成就,但由于低密度基因座的低分辨率和微卫星和 SNP 组标记的数量有限,大多数这些 QTL 被定位到大的基因组区域。因此,仅鉴定了有限数量的基因座。例如,一些基因,包括 IGF1, TBC1D1, FOXO1A, KPNA3, INTS6 和 HNF4G 与鸡的生长和体重有关。这些研究主要基于变异非常有限的商品代鸡,并且大多数控制体型的突变个体已经丢失。全基因组关联研究有望阐明这种复杂性状的遗传基础,虽然这些方法的难度和收集表型数据的费用高昂,妨碍了其广泛应用。此外,不同品种之间的表型多样性及其复杂的群体统计学历史也阻碍了对鸡体型变异的遗传机制的研究。幸运的是,由比较群体基因组学补充的下一代基因组测序数据研究彻底改变了数量遗传学和进化领域,因此已被证明是解释家畜中复杂性状的遗传基础的有力工具,例如野鸽的驯化,高纬度中国猪的冷适应,以及狗适应的富含淀粉的食物的机制。

元宝鸡是中国著名的观赏鸡品种,以其体型小巧著称,成年雄性体重约 800

克，成年雌性体重约 500 克。它具有悠久的繁殖历史，可以追溯到唐朝。小巧的体型使其易于把握，外观类似于中国古代金元宝，使元宝鸡在古代被视为财富的象征。迄今为止，元宝鸡无疑是中国最受人类器重的鸡种之一。在这里，我们采用比较群体基因组学来研究元宝鸡小型体型的遗传基础。我们确定了四个可能影响家鸡体型变异的新基因座。

结 果

89 个鸡全基因组的分析鉴定了超过 2000 万个 SNP

在该研究中，获得 89 个个体的基因组，其中包括 7 只红色丛林鸡，24 只元宝鸡和 58 只其他家养鸡，每个个体具有约 $12.2\times$ 序列覆盖。基因组序列之间的鉴定了总共 21286312 个 SNP，其中 51.8% 映射到基因间区域，42.6% 映射到内含子区域，并且一小部分映射到基因组的外显子区域。蛋白质编码区的 SNP 的功能注释鉴定了非同义氨基酸替换的 101999 个 SNP 和同义替换的 226713 个 SNP，其中 704 个 SNP 导致基因终止密码子获得或丧失。芯片阵列中使用的 SNP 和基因分型阵列也加入到新的数据集中。我们获得的 SNP 数据集远远大于鸡 SNP 数据库中的数据，这些新的 SNP 可能补充鸡变异的数据库。

与其他鸟类相比，元宝鸡的核苷酸多样性水平较低。使用加权邻接法构建所有个体的系统发生树结果表明，元宝鸡形成相对均匀的祖簇。主成分分析、杂合子和基于单倍型的结构分析表明，元宝鸡和其他鸡一样属于混合血统。

比较群体基因组学作为识别控制鸡体型变异位点的有力工具

群体变异度的比较分析是一个强大的工具，能够成功调查复杂性状下的遗传机制。由于元宝鸡的体型与鸡的平均体型相比要小得多，因此元宝和其他鸡的比较基因组分析将是识别鸡体型变异遗传基础的有效方法。这里我们采用了 FST 和 LSBL 进行估计预测。首先，进行基于异常值方法的筛选候选基因，滑动窗口分析 FST 和 LSBL（50kb 窗口大小和 25kb 步长）分别从之前的数据中鉴定 268 和 275 个基因。这些候选基因的功能富集分析未揭示与体型特异性相关的任何通路。通过组合 FST 的信号和 LSBL，我们确定了四个基因组的区域，表现出极端的群体分化，可能是人工选择的结果。分析显示 chr22: 0.25Mb-0.33Mb 的 BMP10 可能控制元宝鸡的体型；基因组区域 chr22: 0.25 MB-0.33M 是与种群分化的最

高水平。有三个基因 GKN1, GKN2 和 BMP10 位于该区域。GKN1 和 GKN2 是在胃中特异性大量表达的旁系同源物。这两个基因都证明了维持胃粘膜完整性和正常功能的功能有关, 它们的异常与胃癌相关。在元宝和其他鸡中, GKN1 和 GKN2 在肾, 脾, 肌肉, 肝和肺中均未表达或极低的表达水平。BMP10 作为转化生长因子 β 家族的成员, 在 FST 和 LSBL 分析和等位基因频率分析有较高的差异。单倍型比较分析揭示 BMP10 与其他家鸡存在一致分化。到目前为止, BMP10 在体型中的作用尚未见报道。此外, 没有与鸡体型相关的 QTL 定位于该基因组区域。

进一步的研究表明, BMP10 在心脏中具有高表达, 并且与其他鸡相比, 元宝鸡的表达水平显著上调。基于 sanger 测序演化网络分析验证支持启动子区域 BMP10 受调控的结论。为了检测启动子活性, 我们用荧光素酶报告基因测定元宝鸡的 BMP10 启动子序列与参考序列的差异。同样地, 我们观察到由 BMP10 驱动的启动子活性与参考启动子相比增加元宝鸡的启动子。接下来, 基于系统发育网络, 我们提取了 BMP10 上游的 21 个 SNP, 结果显示了元宝鸡与其他家鸡之间的高度分化。进一步检查这些 SNP, 通过电泳迁移率变动分析与来自鸡心的核提取物来研究 DNA-蛋白质相互作用的差异。对于 SNP 五个探针元宝和其它鸡在 chr22: 274606 和 chr22: 274758 表现出突变导致 DNA 和蛋白质之间相互作用的减少或丧失。另一方面, 元宝鸡在 chr22: 274670, chr22: 276118 和 chr22: 276140 的突变增加了 DNA-蛋白质相互作用。根据这些观察结果, 我们推断由于潜在的 SNP, 蛋白质结合的差异可能提高元宝鸡心脏中 BMP10 表达的上调。我们这个强连锁区域显示出与体重的差异。这些 SNP 可以解释五种鸡品系中总变异的 22.41%, 包含济宁白鸡, 鲁花鸡, 矮脚鸡, 元宝鸡和观赏鸡。

BMP10 过表达抑制血管生成并诱导斑马鱼体长短

BMP10 的过表达诱导小鼠体重减轻。为了检查它是否是 BMP10 控制脊椎动物的体型的一个保守功能区域, 我们在斑马鱼中进行了野生型 BMP10 的过表达测定。来自 flil1a-EGFP 转基因系的受精单细胞斑马鱼胚胎分别每个胚胎注射 200pg, 野生型斑马鱼 BMP10 mRNA。与未注射的对照胚胎相比, 斑马鱼中 BMP10 的过表达导致体长减小和体轴弯曲。这些结果证实了这一点 BMP10 在决定体型方面具有重要作用。血管生成是生长和发育过程中正常而重要的过程。为了研究

BMP10 过表达对斑马鱼血管生成的表型后果，我们用 0.016% 甲磺酸三卡因麻醉转基因入胚胎，并在受精后 32 小时计数完整节间血管的数量。与对照组斑马鱼相比，过表达 BMP10 的斑马鱼表现出更多的不完全 ISV 并且偶尔出现背主动脉的芽。这些结果表明 BMP10 的过表达抑制斑马鱼血管生成血管生长，抑制生长，导致身体缩短。所有这些数据表明，BMP10 的上调可能有助于减少元宝鸡的体型。

BMP10 的过度表达诱导斑马鱼的发育缺陷。32 h 的总形态与未注射的野生型对照胚胎相比，具有野生型斑马鱼 BMP10 过表达的胚胎表现出体长减小和弯曲的体轴。

chr1 的基因：147.55Mb-147.82Mb，chr2：57.05Mb-57.22Mb，chr24：6.17Mb-6.25Mb 区域可能参与体型的发育

CHR1：147.55Mb-147.82Mb 是选择性扫描 SNP 的一个集合，表现出显著高水平的种群分化水平-更高的 FST 和 LSBL。但是没有注释基因位于该区域。然而该区域位于先前报道的与“体重”和“生长”相关的 QTL 内。在该映射位置附近发现了在控制细胞分裂和生长调节中起作用的基因 GPC5。RNA-seq 分析显示 GPC5 的表达在元宝鸡的心脏中下调。

与 chr1：147.55Mb-147.82Mb 中的相似，尽管它也显示出强烈的选择信号，FST 和 LSBL 所示具有高水平的群体分化，但没有蛋白质编码基因位于 chr2：57.05Mb-57.22Mb 基因组区域。QTL 图显示该区域与鸡的“体重”和“生长速率”密切相关。Chr2：57.05Mb-57.22Mb 是基因 *spalt* 样转录因子 3 的上游，这是一种在动物发育中起基础作用的转录因子。与家鸡和红原鸡相比，我们观察到元宝鸡肾脏中 *SALL3* 的表达上调。

第四个定位区域 chr24：6.17Mb-6.25Mb 含有 13 个蛋白质编码基因。该区域没有与体重或生长速率相关的 QTL，但这 13 种基因参与各种生物过程。例如 *BCO2* 与鸡的皮肤颜色相关，而 *DIXDC1* 和 *CRYAB* 参与神经系统。*HSPB2* 热休克蛋白基因在心脏和骨骼肌中高度表达，这对于维持一些小鼠骨骼肌的肌细胞完整性至关重要。敲除 *HSPB2* 导致小鼠骨骼肌退化。在我们的研究中，与家鸡和红原鸡相比，元宝鸡肌肉中 *HSPB2* 的表达下调。

讨 论

鸡是全世界不可或缺的饮食，它们也被用于生物研究和娱乐等其他目的。作为食物和经济收入来源，育种者已经做出巨大努力来开发体型庞大且肉类生产快速增长的鸡。我们的研究报告了四个可能控制体型的基因座，为鸡的繁殖提供了重要信息和候选遗传标记。我们的研究还提供了一种比较群体基因组学的策略，以确定考虑到鸡体大小变化的候选基因/突变。与以前的方法相比，该策略更具成本效益和节省时间。

在几种驯养动物中观察到体型的巨大变化，包括狗，猪和鸡。体尺的大小不仅是食品生产的重要商业特性，也是进化和发育生物学研究的重要课题。控制体型变异的基因/遗传变异的研究引起了动物育种者、进化和发育生物学家甚至医学科学家的关注。

体型大小作为典型的复杂性状，通常被认为受到涉及许多基因的影响功能通路。例如，在人类研究中，确定了 97 个与体重指数相关的基因座，但这些基因座仅占 BMI 变异的约 2.7%。与之形成鲜明对比的是，据报道很少有基因参与家畜体型的进化。例如，六个基因的突变解释了近半数狗类品种中大小减少的变异。在马中报告了类似的模式，其中四个位点解释了 83% 的体型变异。在这里，我们还发现 BMP10 可以解释五种鸡品系中 22.41% 的体型变异，包括济宁白鸡，鲁花鸡，矮脚鸡，元宝鸡和观赏鸡。人类与家畜之间的这些共同进化可能是由自然选择与人工选择的差异所导致的。

在评估元宝鸡较小体型的遗传基础时，我们鉴定了几个相关基因的基因组区域。在染色体 1 上鉴定出非基因区域。该序列位于蛋白质编码基因 GPC5 的上游，并显示出阳性选择的证据，其表达在元宝鸡的心脏中下调。据报道人 GPC5 基因座与高度有关。基于 Illumina 60 K Chicken SNP Beadchip 的 GWAS 也发现了 GPC5 与家养鸡体重的潜在关联的证据。类似地，在染色体 2 的选定区域中没有发现蛋白质编码基因，其位于另一个参与发育的基因 SRL3 的上游。表达 SALL3 在元宝鸡肾脏上调。SALL3 是否具有控制体型的功能尚不清楚。只知道 SALL3 蛋白直接结合 DNA 甲基转移酶并减少 DNMT3A 介导的 CpG 岛甲基化。DNMT3A 是控制体重和能量稳态所必需的，与人体身高有关。此外，SALL3 的表达可以由 BMP4

诱导，在骨骼系统的发育、骨形态发生蛋白中起重要作用。具有双失活 Sa111 / Sa113 的小鼠突变体在肢体形态发生中表现出畸形。在 24 号染色体的选择性扫描区域，13 种蛋白质编码基因具有不同的生物学功能。HSPB2 是一种必不可少的骨骼肌发育基因，在元宝鸡中下调，可能与元宝鸡独特的生长特性有关。在染色体 22 的选定区域中发现了三种蛋白质编码基因 GKN1, GKN2 和 BMP10。无论是 GKN1 还是 GKN2 都在胃中高度和独特的表达，在维持其正常功能的重要作用。BMP10 在心脏中特异性表达，在调节小鼠心脏发育中起关键作用。BMP10 缺陷小鼠由于严重受损的心脏发育和功能而死亡。BMP10 在小鼠中的过度表达导致胚胎中的心肌过度生长和过度蠕动。在心肌中出生后过度表达 BMP10 的转基因小鼠在 1 个月大时显示心脏大小减少 50% 并且体重和体积减小。据报道，BMP10 还具有诱导细胞凋亡、增殖和生长的作用。例如 BMP10 在前列腺肿瘤中观察到表达降低或不存在，并且强制 BMP10 过表达降低了身体生长、细胞基质粘附、侵袭和前列腺癌细胞的迁移。进一步检查显示，元宝鸡心脏中 BMP10 表达上调，可能是由于 BMP10 上游的 5 个潜在突变可能增加启动子活性。与小鼠相似，BMP10 的过度表达诱导斑马鱼体长减少，这意味着 BMP10 的保守功能控制脊椎动物的体型。在研究中发现四个位点在元宝和其他家鸡之间具有高的群体分化，它们可能影响基因表达而不是直接编码蛋白质。这可能支持一种观点，即基因表达的变化对体型的进化有显著贡献，这一观点与基因表达变化在形态学演变中特别重要的假设一致。

我们的研究仍有一些局限性。首先，我们认为 5 个突变可能会改变 BMP10 的启动子活性，导致元宝鸡 BMP10 的高表达。两个突变 chr22: 274606 和 chr22: 274758 导致 DNA-蛋白质相互作用的减少或丧失，而其他三个突变 chr22: 274670, chr22: 276118 和 chr22: 276140 增加 DNA-蛋白质相互作用。我们的研究无法指出哪些蛋白质参与这些相互作用，无论这些蛋白质中的所有蛋白质是否全部共同作用以激活 BMP10 的启动子。此外，一些微型家养鸡品种的体型与元宝鸡的体型相似，但在我们的研究中我们只包括元宝鸡作为小体型鸡品种。此外，大围山鸡中含有 BMP10 基因的染色体区域与红原鸡和其他家养鸡品系处于相同的野生型状态。因此，我们不能得到确切的结论，BMP10 在其他国内家鸡品系中具有共同的结果，特别是由于家养鸡的复杂起源和群体统计学。更广泛的抽样包括更多小

体型的鸡品种、额外的工作将有助于解决这些问题。

材料和方法

动物实验伦理

所有动物实验程序均按照昆明动物研究所伦理委员会批准的指导方针进行。

采样和基因组数据收集

本研究对多达 42 个基因组进行了测序，其中包括 24 只元宝鸡，8 只大围山鸡，1 只红色丛林鸡，6 只广州本地鸡和 3 只云南当地鸡。使用酚-氯仿提取法提取 DNA，并通过电泳和 NanoDrop 分光光度计 2000 测量质量。根据 Illumina 标准基因组文库制备管道，仅使用高质量 DNA 构建基因组测序文库。测序在 Illumina Hiseq 2000 平台上进行，读取长度为 101bp。来自我们之前研究的 33 只鸡的基因组和 Fan 等人的研究中的 2 只鸡和来自 Yi 等人的研究中的 12 只鸡，被纳入我们的研究。总体而言，获得了 7 只 Red Junglefowls 和 82 只家养鸡的 89 个基因组。

基因组序列比对，SNP 调用和注释

使用 cutadapt 和 Btrim 软件通过过去去除接头和低质量碱基，过滤原始序列读数。使用具有默认设置的 BWA-MEM 将合格的读数与鸡参考基因组比对。然后采用一系列后处理来处理对齐 BAM 格式文件，包括排序，重复标记，局部重新校准和基本质量重新校准，这些都是使用 Picards 中的 SortSam 和

MarkDuplicates 函数进行包，以及 Genome Analysis Toolkit 中的

RealignerTargetCreator，IndelRealigner 和 BaseRecalibrator 工具。使用 GATK 中的 UnifiedGenotyper 和 VariantFiltration 命令调用和过滤 SNP 和 indel。RMS 映射质量<25 且基因型质量<40 的基因座被删除，其中具有零映射质量的读数占该站点所有读数的> 10%。移除等位基因> 2 个并且在簇内的基因座。使用基于 ENSEMBL 鸡注释的 ANNOVAR 将所有 SNP 分配到特定基因组区域和基因。评估缺失频率<10%的 SNP，每个染色体的单倍型由 BEAGLE 推导。

种群变异和种群遗传分析

使用 VCF 工具使用 50kb 滑动窗口以 25kb 逐步递增计算元宝鸡、红原鸡、大围山鸡和其他鸡群的全基因组遗传多样性。采用几种方法推断元宝鸡的种群结

构。首先，我们使用软件 PHYLIP 基于所有 SNP 位点的简单匹配距离导出的成对距离矩阵构建了邻接树。使用 MEGA5 观察树。其次，为了最大限度地减少由广泛的强连锁不平衡区域贡献的 SNP 的影响，我们根据使用 PLINK 观察到的样本相关系数来修剪 SNP，并且使用 GCTA 进行 PCA。第三基于修剪的数据，使用 ADMIXTURE 进行混合物分析以观察群体结构，其祖先群体大小为 2 至 5。第四，我们使用基于单倍型的方法 ChromoPainter 和 fineSTRUCTURE 来推断种群结构。

全基因组选择性扫描分析

我们采用三项试验来研究元宝鸡阳性选择痕迹的基因组区域。如其他地方所述，估计每个 SNP 的 F_{ST} 值在元宝和其他鸡之间。基于三组之间的 F_{ST} 值计算每个 SNP 的 LSBL 统计量。由于大围山鸡具有与红色丛林鸟类相似的特征，因此我们将元宝鸡定义为 A 组，红原鸡和大围山鸡为 B 组。其他家养鸡品系被指定为 C 组。使用 LSBL 公式计算每种变体的 LSBL 统计数，对于每个 50kb 窗口中的 F_{ST} 和 LSBL，以 25kb 逐步递增进行滑动窗口分析。此外，我们计算了元宝与其他鸡之间每个 SNP 的绝对等位基因频率差异，以确认阳性选择的信号。通过以上方法检测到的推导候选选择性扫描使用

<http://asia.ensembl.org/info/docs/tools/index.html> 上提供的变异预测进行注释。使用 g: Profiler 分析蛋白质编码基因的功能富集，包括基因本体论类别，KEGG 途径和人类表型预测。

SNP 验证和网络建设

在 Applied Biosystems ABI 3730XL 遗传分析仪上进行 Sanger 重测序，用于验证 BMP10 基因上游区域和第一外显子区域的 SNP。共有 54 只鸡，包括元宝鸡和当地家养土鸡。用于扩增和测序的引物列于补充表 S7 中。使用 PHASE 程序分阶段进行单倍型。使用网络构建了中间连接网络。

基因分型和关联分析

为了进一步确定 BMP10 的变异如何影响体型，我们对来自五个鸡品系的 301 只鸡进行了 SNP 的基因分型。使用 Sanger 重测序法获得体重信息和基因型对应。解释的重量变化比例使用 PLINK 与线性模型估算。

RNA 提取和实时定量 PCR 检测

使用 TRNzol-A +试剂从成年鸡的心脏，肝脏，脾脏，肺，肌肉，肾脏和脑组织中分离总 RNA，并使用 RNeasy Micro Kit 纯化。使用电泳和 NanoDrop 分光光度计 2000 测量 RNA 的浓度和完整性。根据制造商的说明，使用 PrimeScript RT-PCR 试剂盒以 25 盒以的终体积使用总 RNA 合成单链 cDNA。使用实时定量 PCR 和相对标准曲线方法以及对持家基因 GAPDH 的归一化来测量鸡心脏中 BMP10 的相对 mRNA 表达水平。

RNA-seq 分析

对于 RNA-seq 分析，我们使用 Hiseq 平台生成了来自元宝鸡、家养土鸡和红原鸡的肺，心脏，肌肉，脾脏，肝脏和肾脏组织的 47 个样本转录组数据。首先，使用 Trimmomatic 参数设置为 LEADING，过滤掉质量差的读数。其次，使用 HISAT2 将干净的读数与鸡参考基因组比对，参数设置 Cuffcompare 用于测量红原鸡和元宝鸡之间以及家鸡和元宝鸡之间每种组织的基因表达差异的显著性。基于泊松离散分散模型计算 P 值。

荧光素酶报告分析

为了推断元宝鸡 BMP10 上游高度分化的 SNP 是否增加了 BMP10 启动子活性，我们进行了荧光素酶报告基因检测。分别通过 PCR 产生 BMP10 基因的两个上游片段，长 2449bp 片段和短 949bp 片段。并克隆到 pGL3 Basic 载体中。Kpn I 和 XhoI 酶位点用于构建载体。使用 Lipofectamine TM 2000 在每个孔中用 pGL3 报告质粒和 80ng pRL-TK Renilla 萤光素酶构建体以 60%-70% 汇合转染 24 孔中的人胚肾 293T 细胞。萤光素酶活性使用 GloMax 转染后在第 24 个小时测定 96 微孔板发光。以 Basic 载体为参照计算萤火虫发光/海肾发光的比率。进行了三次技术重复。t 检验用于衡量元宝与其他鸡之间的统计学意义。

电泳迁移率变动分析

从 BMP10 的起始位点起 2249-bp 上游区域内总共 21 个 SNP 通过 Sanger 测序验证的基因，使用 EMSA 的功能测定以揭示 DNA-蛋白质相互作用的潜在差异。合成了总共 20 对 5'标记的探针和 chr22: 276331 突变共享一个探针。在将鸡处死并在 -70° 下储存直至进一步分析后，收集来自 4 只鸡的心脏组织。使用 NucBuster 蛋白质提取试剂盒从心脏组织提取核物质。将核提取物加入到结合反应中，然后在冰上预孵育 20 分钟。将 20pmol 未标记的双链寡聚物加入到反应

中。预温育后，将 20fmol 生物素化的寡聚物加入到反应中并在室温下温育 20 分钟。通过在 0.5% 非变性聚丙烯酰胺凝胶上在 120V 下在 0.5V 聚丙烯运行缓冲液中电泳分离 DNA-蛋白质复合物 90 分钟。将分离的复合物在冷的 0.5 将分离的中以 390mA 转移至结合膜 40 分钟。使用 Stratalinker UV 交联剂交联 DNA-蛋白质复合物，并使用 Lighten 检测生物素化的探针 将分离的复合物在冷的 0.5 的复合物中以 390mA 转移至结合膜 40 分钟。使用 Stratalinker UV 交联剂交联 DNA-蛋白质复合物，并使用 Lighten 检测生物素化的探针 将分离的复合物在冷的 0.5 的复合物中以 390mA 转移至结合膜 40 分钟。使用 Stratalinker UV 交联剂交联 DNA-蛋白质复合物，并使用 Lighten 检测生物素化的探针® HRP-B 底物溶液 A 和 B 连接。

斑马鱼的饲养和保存

成年斑马鱼在 14 小时光照/ 10 小时黑暗周期中维持在 28.5 周。每对放置四到五对斑马鱼进行自然交配。平均而言，产生了 100-200 个胚胎。将胚胎在鱼水中保持在 28.5 将。根据 Kimmel 洗涤并分级胚胎。fli1a-EGFP 转基因系的建立和表征。上海生物模型生物科技发展有限公司的斑马鱼设施获得了国际实验动物管理评估与认证协会的认可。

斑马鱼显微注射

对于过表达测定，受精的单细胞胚胎每个胚胎注射 200pg 野生型斑马鱼 BMP10 mRNA。

斑马鱼血管生成研究

为了评估斑马鱼中的血管形成，向受精的单细胞 fli1a-EGFP 转基因胚胎每个胚胎注射 200pg 野生型 BMP10 mRNA。在 32h 时，用 0.016%MS-222 麻醉胚胎，并计数将 DA 连接到 DLAV 的完整 ISV 的数目。使用公式测定抗血管生成作用

图像采集

用 Nikon SMZ 1500 荧光显微镜检查胚胎和幼虫，然后用数码相机拍照。使用 Adobe Photoshop 7.0 软件调整图像子集的亮度，对比度，色调和饱和度，以最佳地可视化表达模式。使用基于图像的形态测量分析处理定量图像分析。每种处理定量 10 只动物。

统计分析

所有数据均以平均值±有数据表示。使用 GraphPad Prism 5.0 进行统计分析和数据的图形表示。使用学生进行统计显著性测试，ANOVA 或卡方检验进行。统计显著性用*表示。

三、研究的内容及可行性分析

研究内容：

家鸡的羽色在驯化以及选择过程中获得了较大的表型变异，但背后的遗传学基础难以探究。群体遗传学以及高通量测序技术的发展带来了希望。在本次研究中，我们尝试利用 NGS（下一代测序技术）使用群体基因组来识别潜在的決定家鸡羽毛颜色变化的遗传学基础。

我们采取的主要物种-元宝鸡，是一种著名的观赏鸡种，它的人工培育历史可以追溯到唐朝，其纯净的毛色以及元宝大小的小巧体型受到了人们的喜爱，因此也被广泛用作生物模型，进行进化的研究。为了更好地解析元宝鸡为代表的家鸡在驯化过程中导致的遗传突变，我们基于 NGS 和比较群体基因组学研究了元宝鸡的重测序数据。我们将会确定可能控制颜色变化的新基因座区域，将提供有关基础色素沉着的新的生物学见解。

可行性分析：

本实验原理上完全可行，并且当前学校实验室所提供的条件完全可以满足实验要求，所以实验可行性很高。

四、论文拟解决的关键问题及难点

关键问题：

选择具有较高质量的元宝鸡基因组二代测序数据集

遗传育种实验室高性能服务器的环境配置与维护

根据测序数据选择合适的统计学数论方法以及生物信息学分析流程

难点：

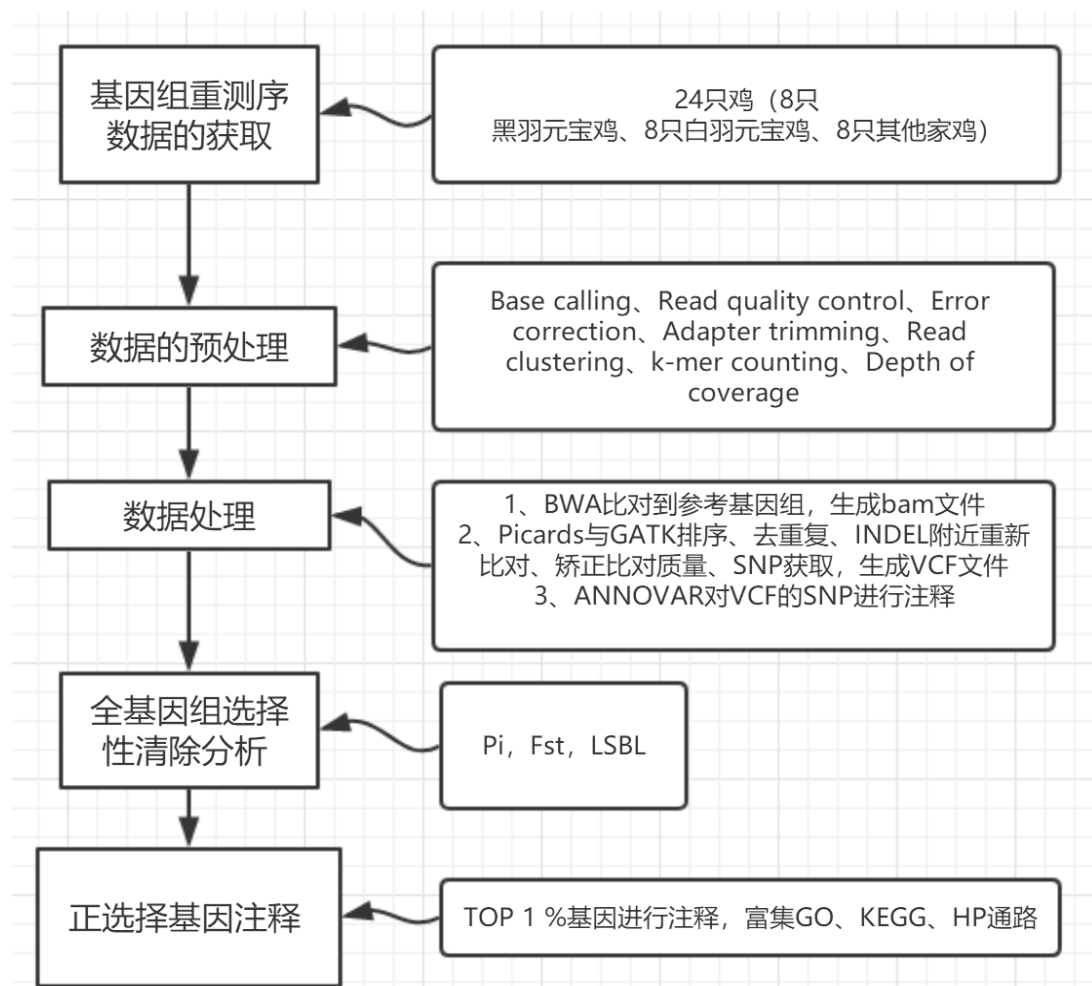
作为二代基因组测序数据集，数据量非常大，对服务器的性能要求以及使用人员的合理规划提出了考验

根据数据的实际情况进行质量控制以及选择合适分析软件

生物信息学的相关操作对实验人员的统计学以及计算机技能提出了较高的考验

五、 研究方法与技术路线(重点论述技术方案)

流程图如下：



1、基因组测序数据的获取

分析数据所有 24 只鸡基因组均来自文章 “Comparative population genomics reveals genetic basis underlying body size of domestic chickens” 中的部分数据集，该研究由 8 只黑色元宝鸡，8 只白色元宝鸡以及 8 只其他家养鸡组成。通过文末提供的 SRA 号，SRAtoolkit 工具批量下载 NCBI 的数据集，存储于高性能服务器。

yuanbao_black	140401_Ypt581-700_AGTCAA_L005
yuanbao_black	140401_Ypt582-700_CCGTCC_L005
yuanbao_black	140403_Ypt588-650_GCCAAT_L006
yuanbao_black	140430_Ypt589-650_CAGATC_L004
yuanbao_black	140403_Ypt595-700_GCCAAT_L003
yuanbao_black	140403_Ypt596-700_CAGATC_L004
yuanbao_black	140403_Ypt597-700_ACTTGA_L004
yuanbao_black	140403_Ypt598-700_GATCAG_L004
yuanbao_white	140401_Ypt578-700_GCCAAT_L004
yuanbao_white	140401_Ypt580-700_CTTGTA_L005
yuanbao_white	140403_Ypt583-650_ATCACG_L005
yuanbao_white	140403_Ypt584-650_CGATGT_L005
yuanbao_white	140403_Ypt591-700_CGATGT_L003
yuanbao_white	140403_Ypt593-700_TGACCA_L003
yuanbao_white	140403_Ypt594-700_ACAGTG_L003
yuanbao_white	140430_Ypt601-700_CTTGTA_L001
yuanbao_other	140403_Ypt585-650_TTAGGC_L005
yuanbao_other	140403_Ypt586-650_TGACCA_L005
yuanbao_other	140416_Ypt590-650_ATCACG_L004
yuanbao_other	140401_Ypt579-700_CAGATC_L004
yuanbao_other	140403_Ypt587-650_ACAGTG_L005
yuanbao_other	140403_Ypt592-700_TTAGGC_L003
yuanbao_other	140403_Ypt599-700_TAGCTT_L004
yuanbao_other	140403_Ypt600-700_GGCTAC_L004

2、基因组比对、变异位点获取以及注释

原始的基因组测序数据去掉接头序列后用 Btrim (<http://graphics.med.yale.edu/trim/>) 去掉低质量的碱基和序列。利用 BWA (<https://github.com/lh3/bwa>) 的 MEM 算法将过滤后的双端测序 reads 比对到家鸡的参考基因组上 (Galgal4)，其余参数均为默认。接着，用 Picards

(<http://picard.sourceforge.net>) 和 Genome Analysis Toolkit (GATK) 软件包中相应的工具对比对后的 bam 文件进行一系列的处理, 包括排序、去除其重复序列、INDEL 附近的 reads 重新比对以及矫正其比对质量等。其次利用 GATK 中的 UnifiedGenotyper 工具获得群体的 SNP, 并使用 VariantFiltration 对 SNP 进行过滤。最终得到 SNP 采用 ANNOVAR 软件对其进行注释。

3、全基因组选择性清除分析

为了全面挖掘家鸡在驯化过程中受人工选择作用的基因, 用基于不同算法的 F_{ST} 、 P_i 、LSBL 三种方法比较家鸡和元宝鸡的群体基因组数据进行比较。首先, 根据 F_{ST} 算法衡量家鸡和元宝鸡的基因组分化程度, 并计算了家鸡和元宝鸡中每一个 SNP 的 F_{ST} 值。接着, 以 50Kb 的窗口、25Kb 的步长分别计算了家鸡和元宝鸡的核苷酸多态性, 根据 $\Delta \pi$ 计算了 v_c 计算了二者的差值。其次, 根据公式进行 LSBL 的计算。计算公式为: $LSBL=(F_{ST}(AB)+F_{ST}(AC)-F_{ST}(BC))/2$ 。

4、正选择基因注释

选取 F_{ST} 、 P_i 、LSBL、 $\Delta \pi$ 分析中的前 1% 的值所相对应的基因组区域为候选的选择性清除的基因座位, 通过在线工具 Variant Effect Predicto

(<http://asia.ensembl.org/info/docs/tools/index.html>) 进行注释。获得的蛋白质编码基因利用 g: Profiler 进行功能注释, 获得富集显著的 GO、KEGG 和 HP 功能通路。

六、论文的进度安排

时间	实验项目	备注
2018.9	查阅文献, 咨询导师, 撰写开题报告	已完成
2018.9-10	学生学习实验的基本技能, 参与课题组实验	已完成
2018.10-11	阅读文献, 参与相关实验, 掌握技能	已完成
2018.10-12	进行实验 分析实验内容。	进行中
2019.1-6	数据分析, 查阅相关文献, 撰写论文, 参与答辩	

七、毕业设计研制报告或毕业论文撰写提纲（初步）

摘要

引 言

第一部分 文献综述

第二部分 研究内容

1 材料与方法

2 数据与分析

3 讨论

4 小结

结 论

参考文献

八、主要参考文献

- [1]中国科学院中国动物志编辑委员会主编.中国动物志 鸟纲 第4卷 鸡形目[M].1978
- [2]Carl-Johan Rubin;Michael C. Zody;Jonas Eriksson;Jennifer R. S. Meadows;Ellen Sherwood;Matthew T. Webster;Lin Jiang;Max Ingman;Ted Sharpe;Sojeong Ka;Finn Hallb  k;Francois Besnier;  rjan Carlbord;Bertrand Bed'Hom;Mich  le Tixier-Boichard;Per Jensen;Paul Siegel;Kerstin Lindblad-Toh;Leif Andersson.Whole-genome resequencing reveals loci under selection during chicken domestication[J].Nature,2010,Vol.464(25): 587-593
- [3]刘益平,朱庆,曾凡同,邱祥聘.原鸡线粒体 DNA 部分序列多态性分析[J].畜牧兽医学报,2004,第 35 卷(2): 134-140
- [4]陆俊贤,贾晓旭,唐修君,樊艳凤,唐梦君,高玉时,苏一军.2 个云南原始鸡种遗传多样性及其与红色原鸡的亲缘关系[J].浙江大学学报(农业与生命科学版),2016,第 42 卷(3): 385-390
- [5]包文斌,束婧婷,王存波,张红霞,Steffen Weigend,陈国宏,BAO Wen-bin, SHU Jing-ting, WANG Cun-bo, ZHANG Hong-xia, Steffen Weigend, CHEN Guo-hong.中国家鸡和红色原鸡 mtDNA 控制区遗传多态性及系统进化分析[J].畜牧兽医学报,2008,(11): 1449-1459
- [6]胡日查,满初日嘎,赵建国,王学梅,李笑春,吴科榜.红色原鸡及其研究进展[J].中国家禽,2010,(1)
- [7]霍海龙,霍金龙,李大林,苗永旺,伍革民,李月体.红色原鸡群体遗传多样性[J].动物学杂志,2007,第 42 卷(5): 131-135
- [8]王晓峰,钱勇.活禽市场规范经营的有效举措[J].中国禽业导刊,2010,(16): 5-11
- [9]Chunyan Mou;Frederique Pitel;David Gourichon;Florence Vignoles;Athanasia Tzika;Patricia

- Tato;Le Yu;Dave W. Burt;Bertrand Bed'hom;Michele Tixier-Boichard;Kevin J. Painter;Denis J. Headon.Cryptic Patterning of Avian Skin Confers a Developmental Facility for Loss of Neck Feathering[J].PLOS Biology,2012,
- [10]Jonas Eriksson;Greger Larson;Ulrika Gunnarsson;Bertrand Bed'hom;Michele Tixier-Boichard;Lina Strömstedt;Dominic Wright;Annemieke Jungerius;Addie Vereijken;Ettore Randi;Per Jensen;Leif Andersson.Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken.[J].PLoS genetics,2008,Vol.4: e1000010
- [11]Chen Siang Ng;Ping Wu;John Foley;Anne Foley;Merry-Lynn McDonald;Wen-Tau Juan;Chih-Jen Huang;Yu-Ting Lai;Wen-Sui Lo;Chih-Feng Chen;Suzanne M. Leal;Huanmin Zhang;Randall B. Wideltz;Pragna I. Patel;Wen-Hsiung Li;Cheng-Ming Chuong.The Chicken Frizzle Feather Is Due to an α -Keratin (KRT75) Mutation That Causes a Defective Rachis[J].PLoS Genetics,2012,Vol.8(7): e1002748
- [12]Hubbard, Joanna K. 1 (Joanna.Hubbard@colorado.edu);Uy, J. Albert C. 2 ;Hauber, Mark E. 3 ;Hoekstra, Hopi E. 4 ;Safran, Rebecca J. 1.Vertebate pigmentation: from underlying genes to adaptive function[J].Trends in Genetics,2010,Vol.26(5): 231-239
- [13]郭军;曲亮;王克华;贺兴龙;.鸡羽色性状基因定位的研究进展[J].中国畜牧兽医,2012,(12): 45-50
- [14]王艳, 舒鼎铭.家禽及哺乳动物类胡萝卜素氧化酶 BCMO1 及 BCO2 研究进展[J].中国家禽,2015,(20): 43-47
- [15]张静, 刘毅, 刘安芳.畜禽羽色候选基因 ASIP 和 TYRP1 的研究进展[J].中国家禽,2015,(1): 55-58
- [16]孟浩浩, 许瑞霞, 代蓉, 李辉, 李良远, 万鹏程, 石国庆.绵羊黑色素合成相关基因的研究进展[J].生物技术通报,2014,(8): 34-39
- [17]徐伟, 封竣淇, 黄兰, 蔡慧芬, 罗卫星.TYR 基因研究进展[J].中国畜牧杂志,2017,第 53 卷 (4): 23-27
- [18]刘小辉, 周荣艳, 张传生, 彭永东, 李祥龙.坝上长尾鸡 TYR 基因核心启动子鉴定与单核苷酸多态性分析[J].农业生物技术学报,2018,第 26 卷(6): 959-969
- [19]刘薇.丝羽乌骨鸡 BAC 文库的构建和黑色素相关基因 TYRP1 和 ID 的研究[D].中国农业大学,2004
- [20]崔丽君, 张桂贤, 王雪娇, 王烨, 刘宇, 刘伟.鸡形目黑素皮质素受体 1 基因多态性研究[J].畜牧与兽医,2014,(7): 10-16
- [21]Sakae Takeuchi;Hideyuki Suzuki;Sayoko Hirose;Masafumi Yabuuchia;Chikara Sato;Hiroaki Yamamoto;Sumio Takahashi.Molecular cloning and sequence analysis of the chick melanocortin 1-receptor gene[J].Biochimica et Biophysica Acta: Gene Structure and Expression,1996,Vol.1306: 122-126
- [22]Takeuchi S;Suzuki H;Yabuuchi M;Takahashi S..A possible involvement of melanocortin 1-receptor in regulating feather color pigmentation in the chicken[J].Biochimica et Biophysica

Acta,1996,Vol.1308(2): 164-168

[23]Marie A Pointer;Nicholas I Mundy.Testing whether macroevolution follows microevolution: Are colour differences among swans (Cygnus) attributable to variation at the MC1R locus?[J].BMC Evolutionary Biology,2008,Vol.8: 249

[24]Eizirik, Eduardo 1,2 ;Yuhki, Naoya 1 ;Johnson, Warren E. 1 ;Menotti-Raymond, Marilyn 1 ;Hannah, Steven S. 3 ;O'Brien, Stephen J. 1 obrien@ncifcrf.gov.Molecular genetics and evolution of melanism in the cat family.[J].Current Biology,2003,Vol.13(5): 448-453

[25]S. Kerje 1 ;J. Lind 1 ;K. Schütz 2 ;P. Jensen 2 ;L. Andersson 1,3.Melanocortin 1-receptor (MC1R) mutations are associated with plumage colour in chicken[J].Animal Genetics,2003,Vol.34(4): 241-248

[26]Maria K. Ling 1 ;Malin C. Lagerström 1 ;Robert Fredriksson 1 ;Ronald Okimoto 2 ;Nicholas I. Mundy 3 ;Sakae Takeuchi 4 ;Helgi B. Schiöth 1.Association of feather colour with constitutively active melanocortin 1 receptors in chicken.[J].European Journal of Biochemistry,2003,Vol.270(7): 1441-1449

[27]翟正晓.基于 RAD 简化基因组测序技术的 13 种中国地方优良鸡品种 SNPs 多态性图谱构建及群体遗传学分析[D].上海交通大学,2014

九、指导教师意见

签名:

20 年 月 日

十、开题审查小组意见

（要求具体意见，对前 7 项进行评价，结论：通过，不通过）

开题审查小组组长签名：

20 年 月 日

