



基因组分析揭示现代家马环境适应的遗传基础

李 聪¹, 李 娜¹, 秦旭勇¹, 高 峰¹, 韩家乐¹, 赵舒悦¹, 格明古丽·木哈台², 党瑞华^{1*}

(1. 西北农林科技大学动物科技学院, 杨凌 712100; 2. 塔里木大学动物科学与技术学院, 阿拉尔 843300)

摘 要: 为阐明现代家马在驯化过程中适应环境变化的遗传基础, 本研究收集国内外 19 个马品种的 317 个个体基因组数据进行分析, 解析马的遗传多样性及种群结构。结合 25 种环境因子, 通过全基因组关联分析与选择清除分析(F_{ST} 和 π ratio), 筛选与环境适应性显著相关的候选基因。结果表明, 两种分析方法共鉴定了 107 个与环境适应性显著相关的共享候选基因, 通路富集分析显示这些基因主要涉及嗅觉传导、血红素结合、氧结合及低氧细胞应答等通路。其中, 7 号染色体上 *OR52A1J* 基因区域存在强烈的选择信号。单倍型分析表明该区域能明显区分高海拔与低海拔马品种, 揭示其可能是藏马高海拔适应的关键候选基因。本研究为揭示现代家马环境适应性提供了新视角, 也为动物遗传资源的管理与保护提供了参考依据。

关键词: 现代家马; 环境适应性; 全基因组关联分析; 选择清除

中图分类号: S821.2

文献标志码: A

文章编号: 0366-6964(2026)01-0194-12

Genomic Analysis Reveals the Genetic Basis of Environmental Adaptation in Modern Domestic Horses

LI Cong¹, LI Na¹, QIN Xuyong¹, GAO Feng¹, HAN Jiale¹, ZHAO Shuyue¹, MUHATAI·Gemingguli²,
DANG Ruihua^{1*}

(1. College of Animal Science and Technology, Northwest A&F University, Yangling 712100, China;
2. College of Animal Science, Tarim University, Alaer 843300, China)

Abstract: To clarify the genetic basis of modern domestic horses' adaptation to environmental changes during domestication, this study collected genomic data from 317 individuals of 19 horse breeds at home and abroad for analysis, and explored the genetic diversity and population structure of horses. Combined with 25 environmental factors, genome-wide association study and selective sweep analysis (F_{ST} and π ratio) were used to screen candidate genes significantly associated with environmental adaptability. The results showed that a total of 107 shared candidate genes significantly associated with environmental adaptability were identified by the two analytical methods. Pathway enrichment analysis indicated that these genes were mainly involved in pathways such as olfactory transduction, heme binding, oxygen binding, and cellular response to hypoxia. Among them, the *OR52A1J* gene region on chromosome 7 showed a strong selection signal. Haplotype analysis revealed that this region could clearly distinguish horse breeds at high and low altitudes, suggesting that it might be a key candidate gene for the adaptation of Tibetan horses to high altitudes. This study provides a new

收稿日期: 2025-07-22

基金项目: 国家重点研发计划(2024YFD1300500); 佳米驴泌乳及肉质性状的精准鉴定(K4050425388); 陕西省重点研发计划项目(2025NC-YBXM-110)

作者简介: 李 聪, 博士, 主要从事动物遗传资源与育种研究, E-mail: lc863749946@163.com

* 通信作者: 党瑞华, 主要从事草食家畜遗传育种与繁殖研究, E-mail: dangruihua@nwsuaf.edu.cn

perspective for revealing the environmental adaptability of modern domestic horses and offers a reference for the management and protection of animal genetic resources.

Keywords: modern domestic horse; environmental adaptability; GWAS; selective sweep

* **Corresponding author:** DNAG Ruihua, E-mail: dangruihua@nwsuaf.edu.cn

现代家马主要分布在北美和欧亚大陆,其分布区域广阔,地形多样且气候差异显著^[1]。它们面临着强烈的选择压力,并能适应各种当地生态环境,包括不同的温度、降水和海拔条件。这些压力导致个体表型发生变异,以及等位基因频率发生改变。由于个体表型可塑性和遗传适应潜力存在差异,它们应对气候变化的能力可能也有所不同^[2]。迄今为止,通过对马基因组的分析,已鉴定出许多与人工选择和重要经济性状相关的基因座,例如毛色^[3]、速度^[4]和步态^[5]。进一步分析证实,马的有效种群规模变化与气候及栖息地可利用性之间存在密切关联^[6]。然而,与马的环境适应和自然选择相关的基因组位点相关研究较少。

环境适应是物种生存和生物多样性保护的基础。近年来,已有研究尝试通过对多个物种进行选择清除分析,以及对极端环境适应性表型进行全基因组关联分析,来鉴定适应性位点,涉及牦牛^[7]、山羊^[8]和鸡^[9]等。此前有一项研究通过选择扫描鉴定了与马驯化和选择过程中高海拔适应相关的基因组区域和基因^[10]。但该研究仅包含中国的马品种,对于全球范围内现代家马的整体选择历史尚不明确。此外,全基因组关联分析通过鉴定与环境变量高度相关的SNP位点,探究环境适应模式。环境关联分析已成功应用于预测物种适应性相关的研究,是对选择信号分析的重要补充^[11]。环境适应性是一种受多个遗传位点影响的复杂性状。了解适应极端环境品种的遗传基础,对于动物遗传资源的管理和保护至关重要。迄今为止,对马基因组的适应性变异及分子机制仍未完全明晰。因此,本研究收集了目前规模较大的马基因组数据集,以挖掘与适应性相关的基因组区域。

本研究收集了分布于全球的317匹马,涵盖19个具有不同遗传背景和地理起源的地方品种。基于基因组数据,对马的遗传多样性和种群遗传分化进行了综合分析。通过对比极端环境下马的基因组特征,旨在鉴定与环境适应性相关的候选基因和信号通路,并揭示选择作用和环境压力对马遗传资源的影响。这有助于对马品种遗传资源进行科学保护。

1 材料与方法

1.1 试验材料

本研究中19个马品种的317个全基因组测序数据下载于NCBI和ENA数据库(表1)。其中,源自欧洲的3个品种(弗赖伯格马、弗里斯兰马、雅库特马)共56匹马;源自北美的28匹夸特马;源自亚洲北部的6个品种(鄂伦春马、蒙古马、宁强马、焉耆马、伊犁马、济州马)共94匹马;源自亚洲南部的3个品种(德保马、建昌马、晋江马)共66匹马;源自亚洲西南部的18匹阿拉伯马;源自亚洲青藏高原地区的4个品种(柴达木马、大通马、河曲马、西藏马)共52匹马,以及源自亚洲蒙古边境的3匹普氏野马。

从新疆马业协会(<https://horse.xjau.edu.cn/>)获取样本的产地信息,确定马品种的地理坐标。环境数据来源于国家青藏高原科学数据中心(<https://data.tpdc.ac.cn/>)、美国国家海洋和大气管理局物理科学实验室(<https://psl.noaa.gov/>)。此外,使用elevatr0.4.2软件包获取海拔数据。本研究共收集了1960—2010年的25个环境因素(地理与地表覆盖参数:海拔、植被覆盖率;温度参数:2 m温度、2 m最高温度、2 m最低温度、地表温度、潜在温度、昼夜温差、霜冻期;水分与降水参数:降水量、降水率、降水量>0.1 mm的天数、可降水量、蒸气压、蒸散量、潜在蒸发率、土壤湿度、帕默尔干旱指数、积雪覆盖;大气与气象参数:相对湿度、2 m比湿、地面气压、风速、云量、向外长波辐射)。

1.2 提取基因组变异集合

使用SRAToolkit2.8.1软件^[12]将基因组原始数据转换为FASTQ文件,随后利用Trimmomatic0.36软件^[13]过滤接头序列和低质量序列,并利用Samtools0.1.18软件^[14]构建索引文件。之后,采用Burrows-Wheeler Aligner0.7.17软件^[15]将过滤后的文件比对到马参考基因组(EquCab3.0);随后使用Picard2.18.17软件(<http://broadinstitute.github.io/picard/>)去除重复序列。使用GATK3.8软件^[16]以默认参数进行变异检测。最后,利用VCFtools0.1.16软件^[17]通过“--min-alleles 2--max-alleles 2--min-

表1 试验样本信息统计表

Table 1 Statistics of experimental samples information

组别 Group	品种 Breed	样本量/匹 Sample size	数据来源 Data source
亚洲北部 Asia North	鄂伦春马 Erlunchun horse	6	PRJNA515160
	济州马 Jeju horse	18	PRJNA168142
	蒙古马 Mongolian horse	35	PRJEB10098
	宁强马 Ningqiang pony	13	PRJNA780752
	焉耆马 Yanqi horse	6	PRJNA780752
	伊犁马 Yili horse	16	PRJNA780752
亚洲青藏高原 Asia Qinghai-Tibet plateau	柴达木马 Chaidamu horse	8	PRJNA515160
	大通马 Datong horse	9	PRJNA780752
	河曲马 Hequ horse	6	PRJNA515160
	西藏马 Tibet horse	29	PRJNA515160
亚洲南部 Asia South	德保矮马 Debao pony	52	PRJNA416233
	建昌马 Jianchang horse	7	PRJNA780752
	晋江马 Jinjiang horse	7	PRJNA780752
亚洲西南部 Asia Southwest	阿拉伯马 Arabian horse	18	PRJEB14779
欧洲 Europe	弗赖伯格马 Freiberg horse	31	PRJEB10098
	弗里斯兰马 Friesian horse	17	PRJEB13863
	雅库特马 Yakutian horse	8	PRJEB10854
北美洲 North America	夸特马 Quarter horse	28	PRJEB14779
野马 Wild	普氏野马 Przewalski's horse	3	PRJEB10098

meanDP 5--max-missing 0.75"参数对 SNP 文件进行过滤。通过 PLINK1.9 软件^[18]排除位点检出率 < 90%、次要等位基因频率(MAF) < 0.05 的 SNP 以及性染色体 SNP 位点。质量控制后的数据集包含 317 个个体和 7 399 516 个 SNPs 位点。

1.3 群体遗传分析

使用 PLINK 软件的"--indep-pairwise 1000 50.5"参数对 SNP 数据集进行筛选,最终获得 1 199 266 个独立 SNPs 位点用于后续分析。通过 PLINK 软件对上述 SNP 数据集进行主成分分析。采用 ADMIXTURE1.3.0 软件^[19]进行群体结构分析。使用 PopLDdecay3.42 软件^[20]计算常染色体 SNP 位点间的连锁不平衡程度。利用 PLINK 软件的"--homozyg"模块识别个体的长纯合片段(runs of homozygosity, ROH)。通过 VCF2Dis1.47 和 FastME2.0 软件构建用于系统发育树的距离矩阵^[21],并使用 iTOL 和 R4.0.5 软件进行可视化^[22]。

1.4 全基因组关联分析和选择清除分析

本研究采用 GEMMA0.98.5 软件^[23]的混合线性模型,进行全基因组关联分析。以 25 个环境因素

作为评估环境适应性的表型指标,同时使用前 5 个主成分作为协变量以矫正种群分层。全基因组显著阈值设定为 0.05 除以分析使用的 SNP 位点数量 ($-\log_{10}(P)=7.94$)。候选区域定义为显著 SNP 位点上、下游各延长 50 kb 的基因组区域。

通过对 19 个马品种所处的环境因素进行排序,将环境因素极端组的阈值定义为均值加减一个标准差。筛选出每个环境因素的极端低值组和极端高值组。使用 VCFtools 软件计算两组马群之间的 F_{ST} 和 π ratio 值(采用 50 kb 滑动窗口,10 kb 步长)。通过排序确定 F_{ST} 和 π ratio 的显著阈值,选取前 1% 的窗口作为受选择区域。

1.5 候选基因鉴定与富集分析

通过 BioMart 数据库 (<http://asia.ensembl.org/>)对显著基因组区域进行基因注释。候选基因为全基因组关联分析和选择清除分析共同检出的基因。利用 KOBAS3.0^[24]进行候选基因功能注释和通路富集分析,显著富集的 GO 条目/KEGG 通路以 $P < 0.05$ 为阈值。

2 结 果

2.1 现代家马的基因组群体结构

对 SNP 数据进行主成分(图 1A)和系统发育树分析(图 1B),其中普氏野马作为外群。结果显示,亚洲马群与其他马群存在明显分化,且可进一步划分为 4 个群体(亚洲北部、南部、青藏高原及西南地区群体),这与它们的地理分布保持一致。通过连锁不平衡衰减分析发现,亚洲西南、北美及欧洲群体的衰减速率最慢,其次是亚洲青藏高原、南部和北部马群体(图 1C)。基于长纯合片段统计结果,亚洲西南、欧洲及北美群体的平均 ROH 长度最长,而其他 3 个亚洲群体较短(图 1D)。群体结构分析与该结果一致,显示亚洲群体的祖先组成具有均一性,且与其他群体存在差异(图 1E)。结合系统发育树和群体结构分析发现,北美马可能与欧洲马存在

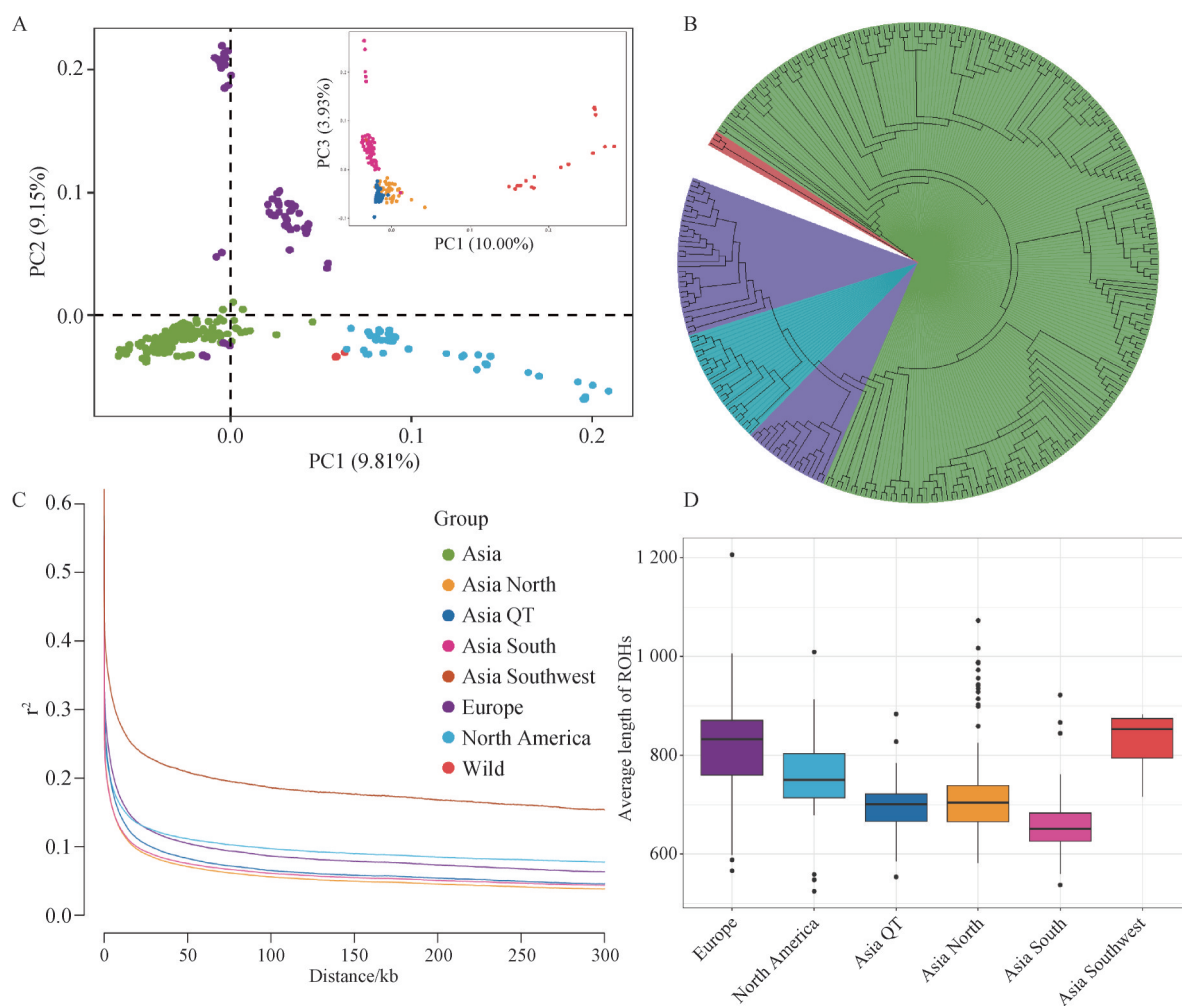
基因流。

2.2 环境因素的统计分析

本研究共收集了与环境适应性相关的 25 个环境因素,对环境因素数据进行统计分析,结果表明各因素符合正态分布(图 2)。通过对这些环境因素进行相关性分析,结果发现大多数环境因素间存在高度相关(图 3),说明环境因素可能通过协同作用影响群体的环境适应性。

2.3 全基因组关联分析和选择信号分析

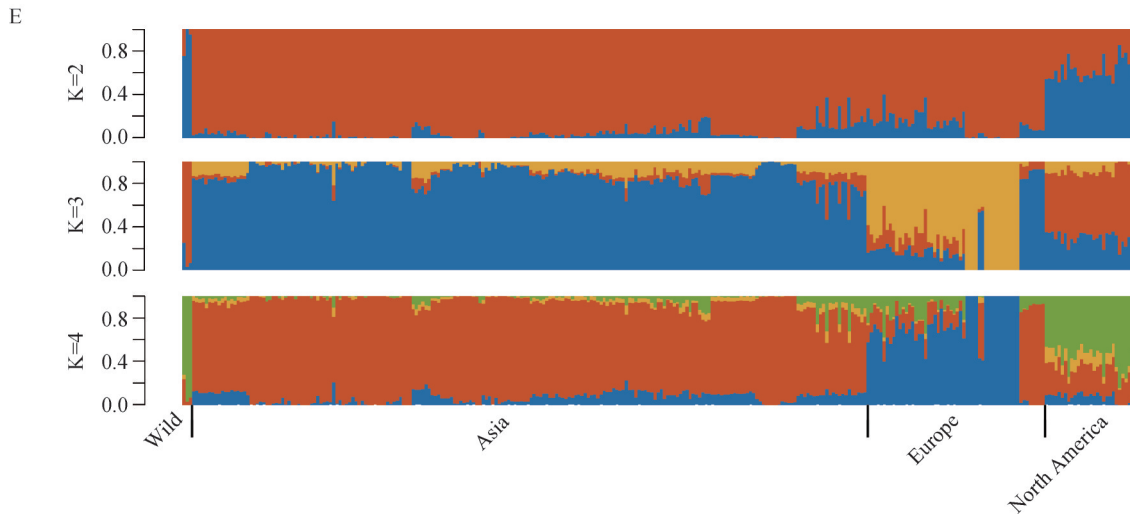
全基因组关联分析共检测到 2 294 个显著 SNPs 位点(图 4),并在候选基因组区域内鉴定出 418 个候选基因。其中,7 号染色体区域(74.93~76.65 Mb)存在一个候选基因簇(包括 *HBB*、*TRIM6*、*TRIM34* 以及 *OR52A1J* 等多个嗅觉受体家族基因),该基因簇与地面气压、潜在温度、霜冻期和帕默尔干旱指数显著相关。此外,*TBX3* 和 *TBX5* 这两个候选基因



(图 1 Fig.1)

(转下页 Carried forward)

(续图1 Continued)

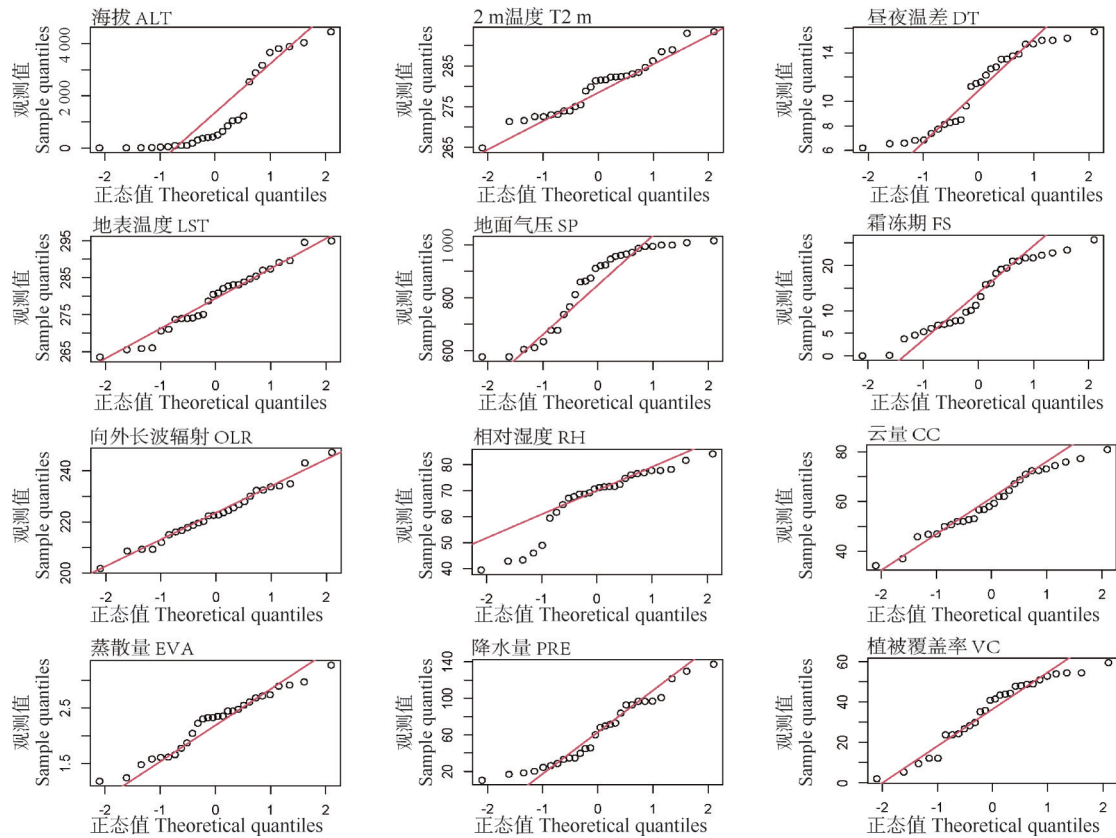


A. 群体的主成分分析; B. 邻接系统发育树; C. 群体的连锁不平衡衰减模式, 横坐标为遗传标记间的物理距离, 纵坐标为连锁不平衡强度; D. 长纯合片段统计; E. 马的群体结构分析(K=2-4)

A. Principal component analysis of population; B. NJ phylogenetic tree; C. LD decay patterns of population, X-axis indicate physical distance between genetic markers, Y-axis indicate linkage disequilibrium strength; D. ROH statistics; E. Admixture analysis for horse populations for K values between 2 and 4

图1 马的遗传多样性和群体结构

Fig. 1 Genetic diversity and population structure of horses



(图2 Fig.2)

(转下页 Carried forward)

(续图 2 Continued)

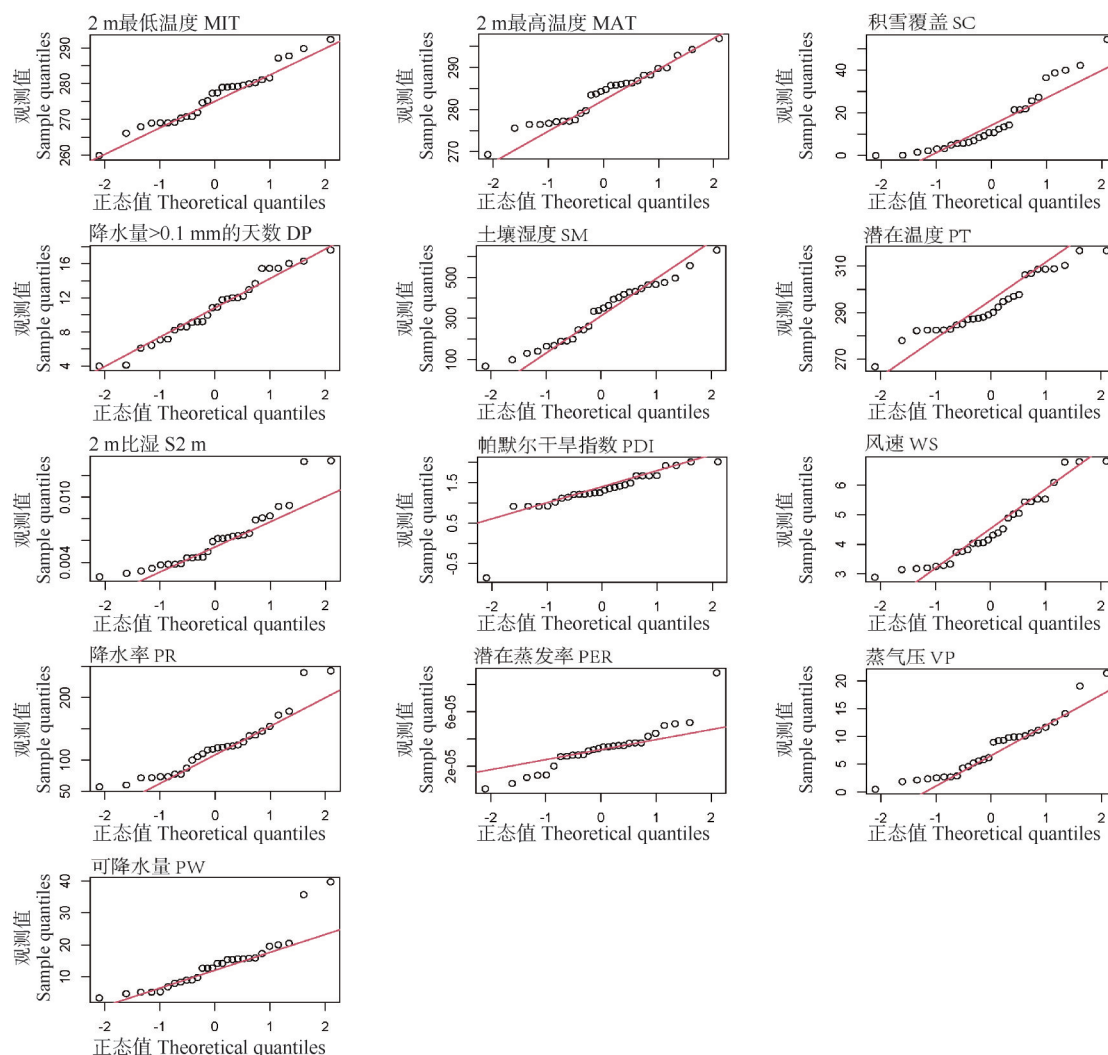


图 2 环境因素的 Q-Q 图

Fig. 2 Quantile-Quantile plot of environmental factors

在可降水量、霜冻期、2 m 温度和 2 m 比湿中存在重叠信号。14 号染色体区域 (39.40~39.50 Mb) 注释到 5 个候选基因 (*TGFBI*、*LECT2*、*FBXL21*、*IL9* 和 *SLC25A48*)，这些基因与地表温度、降水量、蒸气压、霜冻期、土壤湿度和可降水量均显著相关。在基于 F_{ST} 和 π ratio 的选择信号分析中，通过比较极端群体与单一环境因素的关联，共筛选出 6 890 个受选择窗口，这些窗口分布在 7、8、14 和 18 号染色体上，从中注释到 2 172 个候选基因。

2.4 候选区域内基因功能分析

选取全基因组关联分析和选择清除分析结果的交集基因作为环境适应性候选基因，共获得 107 个候选基因。其中包括 19 个嗅觉受体家族的基因，候

选基因 *OR52A1J* 与海拔、地面气压和潜在温度显著相关。通过对该候选基因区域进行单倍型分析 (图 5A)，发现高海拔群体 (亚洲青藏高原地区) 出现了明显的单倍型差异。进一步比较了高海拔和低海拔群体中 *OR52A1J* 基因错义突变 rs1140337342 (位于 7 号染色体的 76 036 185 bp) 的突变频率 (图 5B)，发现海拔 < 1 km 群体中的 C 等位基因频率极低 (13.5%)；而随着海拔的升高，马群体中衍生的 C 等位基因频率逐渐增加。候选基因的功能注释和通路富集分析结果显示，共有 19 个显著条目，包括嗅觉转导 (ecb04740)、血红素结合 (GO:0020037)、氧结合 (GO:0019825) 和细胞对缺氧的应答 (GO:0071456) 等 (图 5C)。

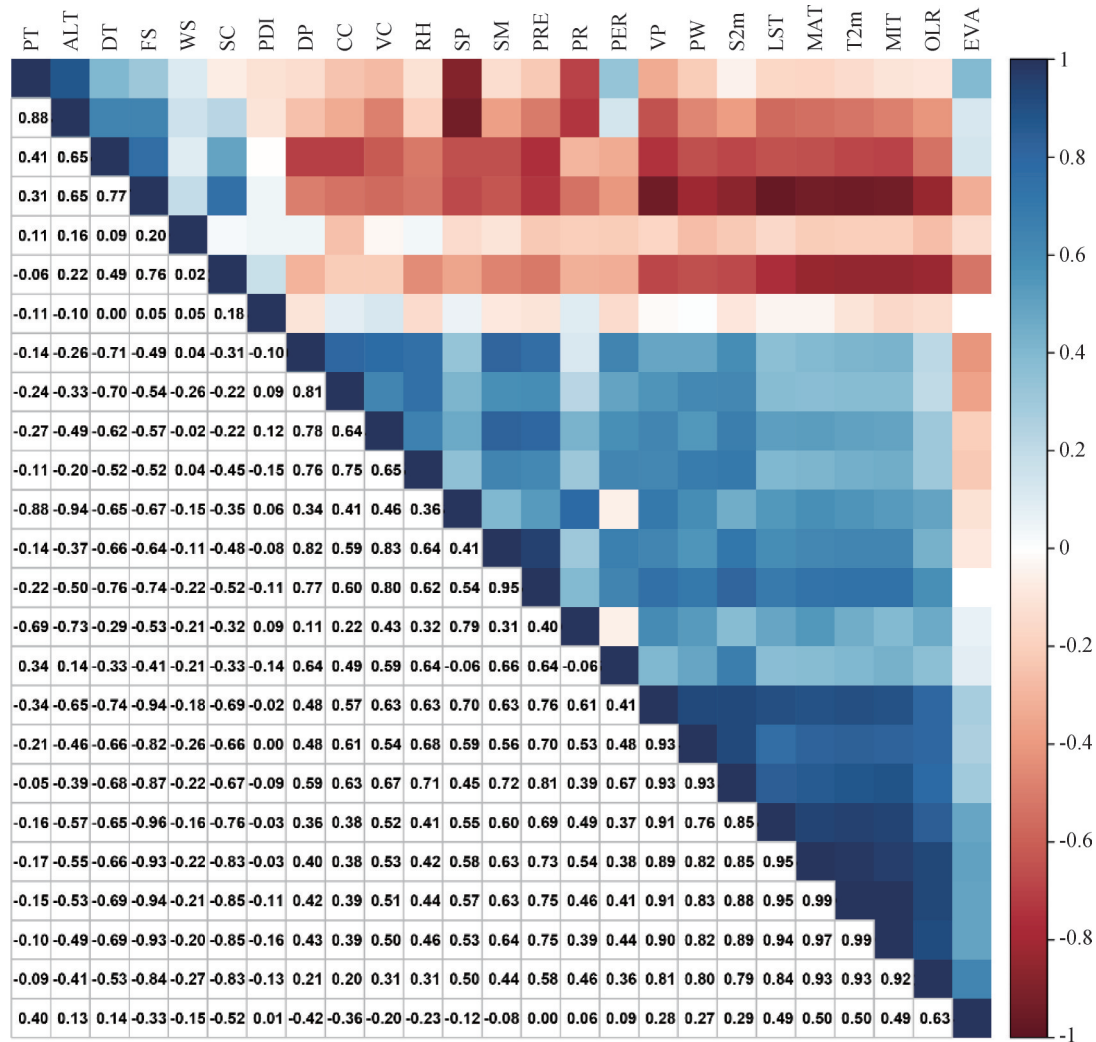


图3 环境因素的相关性热图

Fig. 3 Correlation heat map of environmental factors

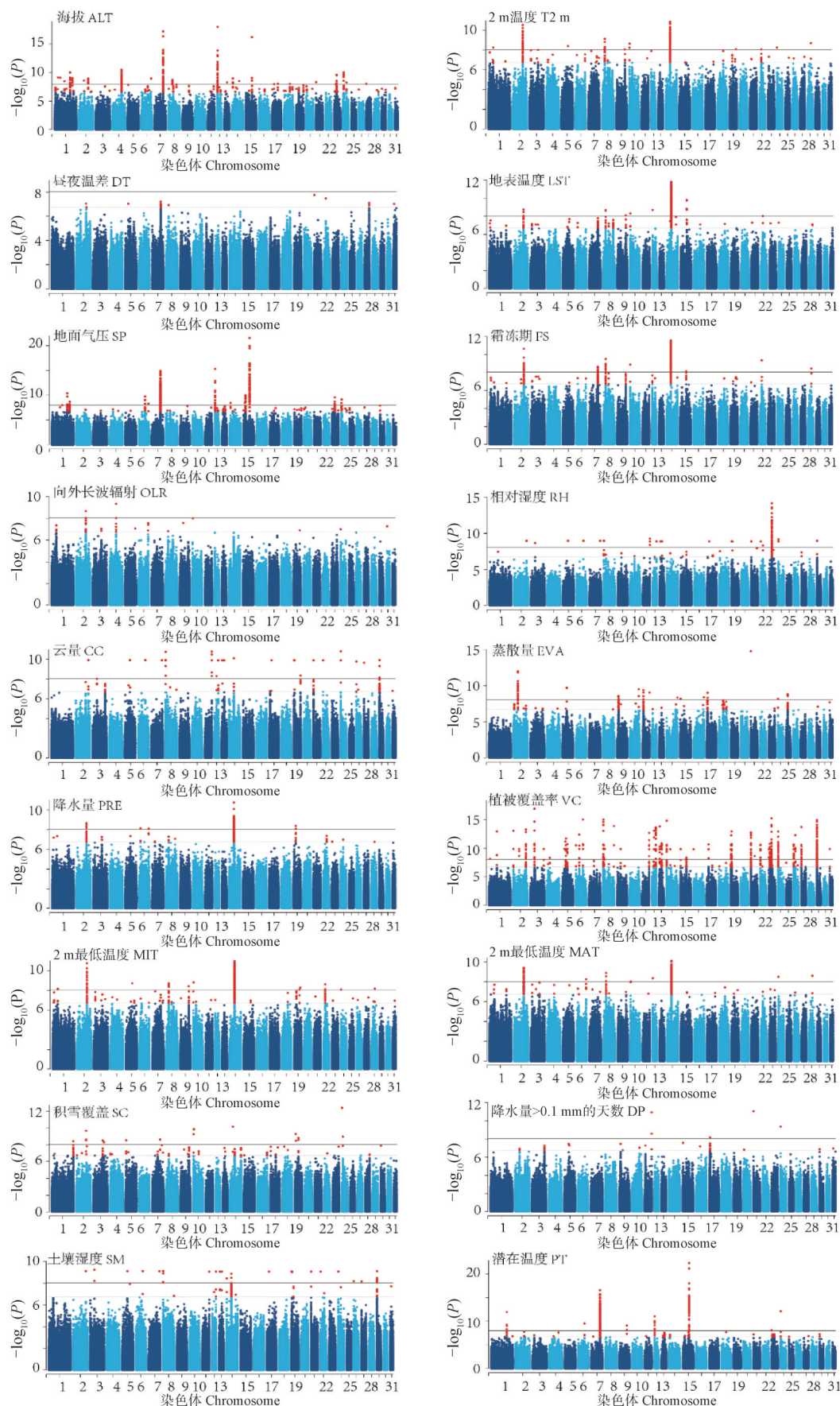
3 讨论

本研究探究了19个马品种的遗传多样性,并通过全基因组关联分析和选择清除分析鉴定了与环境适应性显著相关的候选基因。亚洲马品种的群体遗传结构与其他马品种呈现明显的差异,印证了自然和人工选择对遗传特征的影响。此外,揭示马环境适应性的遗传基础,有助于理解其对环境多样性的进化适应机制。

基于主成分、系统发育树和群体结构分析结果,欧洲与北美马品种表现出遗传混合的特征,表明欧洲和北美马品种在历史上存在持续的基因交流。由于地理分布的显著差异,亚洲马品种进一步划分为亚洲北部、南部、西南部以及青藏高原群体,这些地区具有复杂的地理和气候条件。根据连锁

不平衡和长纯合片段的统计结果:衰减速率越慢,群体的驯化程度越高,选择强度也越大。同时,本研究发现,亚洲北部、南部、西南部马群体的遗传多样性高于欧洲和北美马群体,这与国内家马种群的遗传多样性研究结果一致^[25],也与欧美现代家马经历高强度人工选择导致遗传多样性降低的历史背景相符^[26]。

通过全基因组关联分析,筛选出了与马环境适应性相关的候选基因。对这些候选基因进行功能注释后发现,它们可能直接参与高海拔环境诱导的多种应激反应。其中,*OR52A1J*和*TMEM247*基因与血红蛋白浓度相关^[27-28],*PRKCE*则在葡萄糖稳态和脂质代谢中发挥作用^[29]。*HBB*、*EPAS1*、*CAT*这3个已知基因此前已被报道与人类和绵羊的高海拔适应性以及雅库特马的耐寒性相关^[30-32]。此外,



(图4 Fig.4)

(转下页 Carried forward)

(续图 4 Continued)

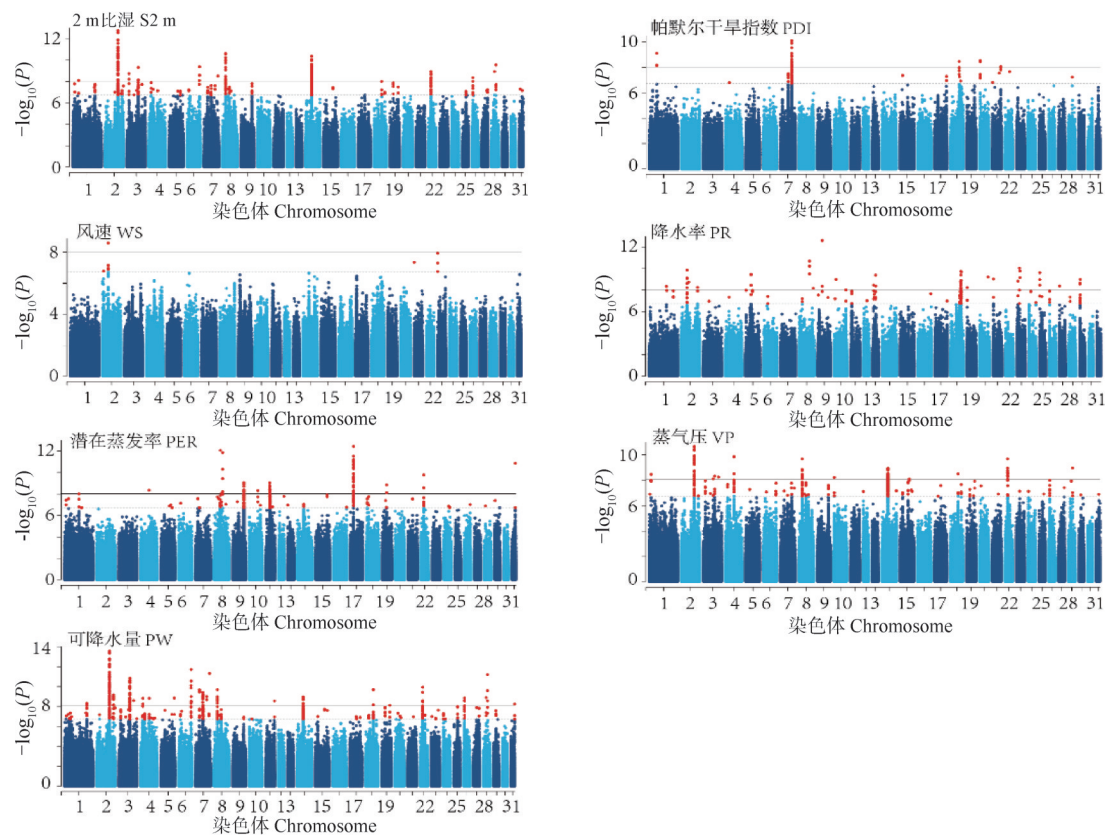
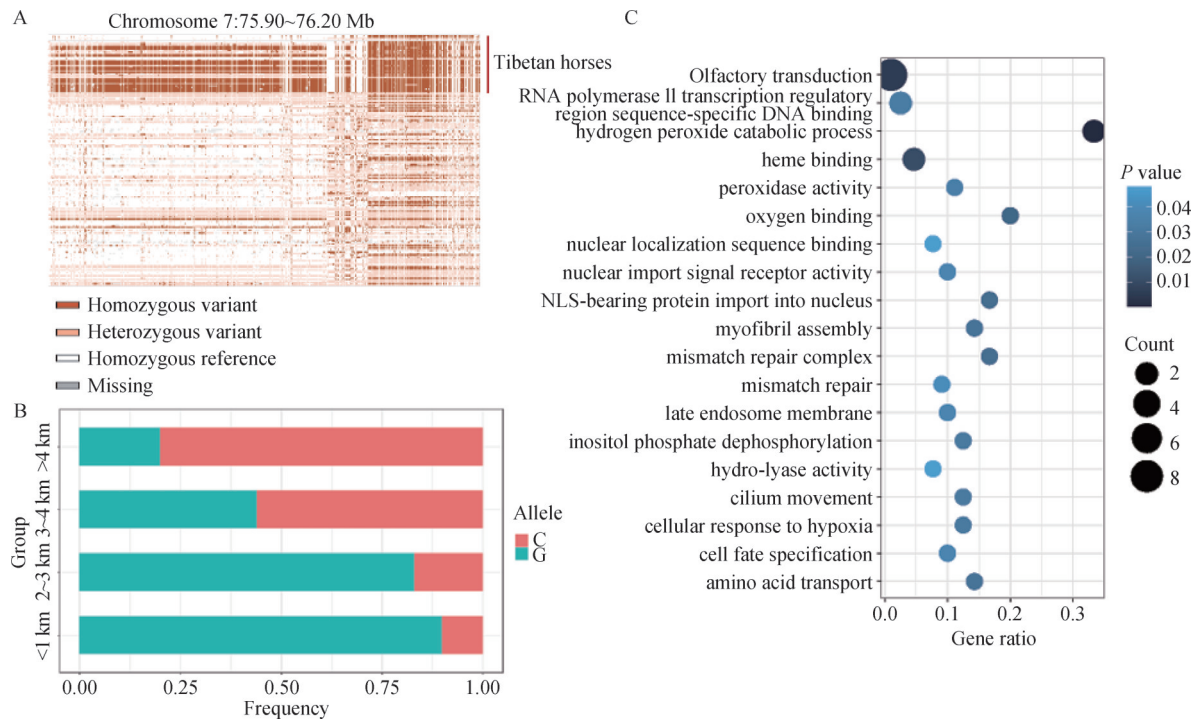


图 4 环境因素的全基因组关联分析

Fig. 4 Genome-wide association study of environmental factors



A. 单倍型分析; B. 等位基因频率分析; C. 候选基因富集气泡图

A. Haplotype analysis; B. Allele frequency analysis; C. Candidate genes enrichment bubble plot

图 5 候选基因功能分析

Fig. 5 Functional analysis of candidate genes

SLC13A1、*TBX3*、*TBX5*和*PGM5*也与环境因子显著相关,且有报道表明它们在生长发育中发挥关键作用^[33-35]。在选择清除分析鉴定的候选基因包括参与牦牛群体热应激反应的*SLC25A48*^[36],以及参与免疫和炎症反应的*TRIM34*、*TGFBI*、*LECT2*、*FBXL21*、*IL9*和*DOCK8*^[37-39]。

极端环境条件(如高海拔、低氧、低温和强紫外线辐射)对动物的生存构成了挑战。其中,极端温度会触发内分泌反应。本研究中,*GPR26*和*LECT2*基因与胰岛素分泌及葡萄糖代谢相关^[40-41],这可能反映了马在调节血糖方面的适应性特征,而血糖调节与葡萄糖产热及抗冻能力相关。此外,部分候选基因可能在生长发育中发挥作用。例如,*SLC13A1*、*TBX3*、*TBX5*、*SPRYD7*、*MSTN*和*DMRT2*对骨骼发育^[33-34]、骨骼肌^[42]和软骨形成^[43]具有影响。另外,*TGFBI*参与脂质代谢^[44]; *MYO3B*和*ORMDL1*与脂肪沉积相关^[45-46]; *DMRT3*、*LDLRAD4*、*MDGA2*、*OSGEPL1*则可能参与运动协调、突触间神经递质传递及神经发育^[47]。上述基因的功能特征共同揭示了马对能量代谢、骨骼发育及神经系统的综合适应。通过对7号染色体强选择区域的单倍型分析进一步发现,青藏高原马群体中存在大片段单倍型的固定现象,为解析现代家马的高海拔适应性提供了新线索。此外,相关研究在甘孜马中鉴定到*OR52A1J*基因携带的一个错义突变可正向调控马匹对高海拔环境的适应性^[48]。该基因的功能特征与此区域的强选择信号相互印证,共同揭示了一个典型的适应模式:极端环境作为选择驱动力,首先筛选出携带适应性突变的基因,进而驱动其功能特化以响应环境需求,最终形成了马种与高海拔环境长期共存的遗传适配模式,从而阐释了现代家马海拔适应的遗传机制。

4 结 论

本研究基于全基因组测序系统解析了现代家马种群的遗传多样性特征及其环境适应性进化的基因组基础,发现亚洲马品种具有较高的遗传多样性水平。通过全基因组关联分析和选择清除联合分析,鉴定出多个与嗅觉转导、缺氧适应、血红素代谢及氧结合相关的候选基因。其中,*OR52A1J*基因表现出显著的高海拔适应选择信号,这一结果为嗅觉系统进化与环境适应性之间的关联提供了有力证据。本研究揭示了环境适应性进化在马基因组

和品种特征上留下的遗传印记,为深入理解基因组变异与马适应性性状之间的复杂互作机制提供了重要理论依据。

参考文献(References):

- [1] ORLANDO L. Equids[J]. *Curr Biol*, 2015, 25(20): R973-R978.
- [2] BRADSHAW W E, HOLZAPFEL C M. Genetic shift in photoperiodic response correlated with global warming[J]. *Proc Natl Acad Sci U S A*, 2001, 98(25):14509-14511.
- [3] LUDWIG A, PRUVOST M, REISSMANN M, et al. Coat color variation at the beginning of horse domestication[J]. *Science*, 2009, 324(5926):485.
- [4] BOWER M A, MCGIVNEY B A, CAMPANA M G, et al. The genetic origin and history of speed in the thoroughbred racehorse [J]. *Nat Commun*, 2012, 3:643.
- [5] WUTKE S, ANDERSSON L, BENECKE N, et al. The origin of ambling horses [J]. *Curr Biol*, 2016, 26(15):R697-R699.
- [6] ORLANDO L, GINOLHAC A, ZHANG G, et al. Recalibrating equus evolution using the genome sequence of an early Middle Pleistocene horse [J]. *Nature*, 2013, 499(7456):74-78.
- [7] 杨柏高, 郝海生, 杜卫华, 等. 牦牛高原适应研究进展[J]. 畜牧兽医学报, 2023, 54(1):12-23.
YANG B G, HAO H S, DU W H, et al. Advances in research on plateau adaptation of yak [J]. *Acta Veterinaria et Zootechnica Sinica*, 2023, 54(1):12-23. (in Chinese)
- [8] 张政凯, 李业芳, 叶绍辉, 等. 山羊环境适应性的研究进展[J]. 畜牧兽医学报, 2022, 53(7):2035-2046.
ZHANG Z K, LI Y F, YE S H, et al. Research progress of environmental adaptability in goats [J]. *Acta Veterinaria et Zootechnica Sinica*, 2022, 53(7):2035-2046. (in Chinese)
- [9] 黄晓隆, 盛熙晖, 袁经纬. 基于组学分析的鸡环境适应性研究进展[J]. 畜牧兽医学报, 2024, 55(7):2809-2824.
HUANG X L, SHENG X H, YUAN J W. Research progress of environmental adaptability in chickens from perspective of omics analysis [J]. *Acta Veterinaria et Zootechnica Sinica*, 2024, 55(7):2809-2824. (in Chinese)
- [10] LIU X, ZHANG Y, LI Y, et al. EPAS1 gain-of-

- function mutation contributes to high-altitude adaptation in Tibetan horses[J]. *Mol Biol Evol*, 2019, 36(11):2591-2603.
- [11] LI Y, CAO K, LI N, et al. Genomic analyses provide insights into peach local adaptation and responses to climate change [J]. *Genome Res*, 2021, 31 (4) : 592-606.
- [12] KODAMA Y, SHUMWAY M, LEINONEN R. The sequence read archive; explosive growth of sequencing data[J]. *Nucleic Acids Res*, 2012, 40(D1):D54-D56.
- [13] BOLGER A M, LOHSE M, USADEL B. Trimmomatic: a flexible trimmer for Illumina sequence data[J]. *Bioinformatics*, 2014, 30(15):2114-2120.
- [14] LI H, HANDSAKER B, WYSOKER A, et al. The Sequence Alignment/Map format and SAMtools[J]. *Bioinformatics*, 2009, 25(16):2078-2079.
- [15] LI H, DURBIN R. Fast and accurate short read alignment with burrows-wheeler transform [J]. *Bioinformatics*, 2009, 25(14):1754-1760.
- [16] MCKENNA A, HANNA M, BANKS E, et al. The genome analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data [J]. *Genome Res*, 2010, 20(9):1297-1303.
- [17] DANECEK P, AUTON A, ABECASIS G, et al. The variant call format and VCFtools[J]. *Bioinformatics*, 2011, 27(15):2156-2158.
- [18] PURCELL S, NEALE B, TODD-BROWN K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses [J]. *Am J Hum Genet*, 2007, 81(3):559-575.
- [19] ALEXANDER D H, NOVEMBRE J, LANGE K. Fast model-based estimation of ancestry in unrelated individuals [J]. *Genome Res*, 2009, 19(9):1655-1664.
- [20] ZHANG C, DONG S S, XU J Y, et al. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files [J]. *Bioinformatics*, 2019, 35(10):1786-1788.
- [21] GUINDON S, GASCUEL O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood [J]. *Syst Biol*, 2003, 52 (5) : 696-704.
- [22] LETUNIC I, BORK P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation [J]. *Nucleic Acids Research*, 2021, 49 (W1):W293-W296.
- [23] ZHOU X, STEPHENS M. Genome-wide efficient mixed-model analysis for association studies [J]. *Nat Genet*, 2012, 44(7):821-824.
- [24] BU D, LUO H, HUO P, et al. KOBAS-i: intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis [J]. *Nucleic Acids Research*, 2021, 49(W1):W317-W325.
- [25] 虎巧燕. 关中马和建昌马全基因组遗传多样性和选择信号研究[D]. 杨凌:西北农林科技大学, 2025.
- HU Q Y. Whole-genome genetic diversity and selective signals in Guanzhong and Jianchang horses [D]. Yangling: Northwest A&F University, 2025. (in Chinese)
- [26] ORLANDO L. The Evolutionary and historical foundation of the modern horse: Lessons from ancient genomics[J]. *Annu Rev Genet*, 2020, 54:563-581.
- [27] HIMADEWI P, WANG X Q D, FENG F, et al. 3'HS1 CTCF binding site in human β -globin locus regulates fetal hemoglobin expression[J]. *Elife*, 2021, 10:e70557.
- [28] DENG L, ZHANG C, YUAN K, et al. Prioritizing natural-selection signals from the deep-sequencing genomic data suggests multi-variant adaptation in Tibetan highlanders [J]. *Natl Sci Rev*, 2019, 6 (6) : 1201-1222.
- [29] RADDATZ K, TURNER N, FRANGIOUDAKIS G, et al. Time-dependent effects of Prkce deletion on glucose homeostasis and hepatic lipid metabolism on dietary lipid oversupply in mice [J]. *Diabetologia*, 2011, 54(6):1447-1456.
- [30] YI X, LIANG Y, HUERTA-SANCHEZ E, et al. Sequencing of 50 human exomes reveals adaptation to high altitude[J]. *Science*, 2010, 329(5987):75-78.
- [31] LV F H, CAO Y H, LIU G J, et al. Whole-genome resequencing of worldwide wild and domestic sheep elucidates genetic diversity, introgression, and agronomically important loci[J]. *Mol Biol Evol*, 2022, 39(2):msab353.
- [32] LIBRADO P, DER SARKISSIAN C, ERMINI L, et al. Tracking the origins of Yakutian horses and the genetic basis for their fast adaptation to subarctic environments[J]. *Proc Natl Acad Sci U S A*, 2015, 112(50):E6889-E6897.
- [33] BJORNSDOTTIR G, STEFANSDOTTIR L, THORLEIFSSON G, et al. Rare SLC13A1 variants associate with intervertebral disc disorder highlighting role of sulfate in disc pathology [J]. *Nat Commun*, 2022, 13(1):634.
- [34] LIU X, ZHANG Y, LIU W, et al. A single-nucleotide

- mutation within the TBX3 enhancer increased body size in Chinese horses [J]. *Curr Biol*, 2022, 32 (2) : 480-487.
- [35] HANNAN N J, STOCK O, SPENCER R, et al. Circulating mRNAs are differentially expressed in pregnancies with severe placental insufficiency and at high risk of stillbirth [J]. *BMC Med*, 2020, 18 (1) : 145.
- [36] LI G, LUO J, WANG F, et al. Whole-genome resequencing reveals genetic diversity, differentiation, and selection signatures of yak breeds/populations in Qinghai, China [J]. *Front Genet*, 2022, 13: 1034094.
- [37] WANG X, XIONG J, ZHOU D, et al. TRIM34 modulates influenza virus-activated programmed cell death by targeting Z-DNA-binding protein 1 for K63-linked polyubiquitination [J]. *J Biol Chem*, 2022, 298 (3) : 101611.
- [38] ZHOU X J, TSOI L C, HU Y, et al. Exome chip analyses and genetic risk for IgA nephropathy among han Chinese [J]. *Clin J Am Soc Nephrol*, 2021, 16 (2) : 213-224.
- [39] LI Q, ZHANG Z, FAN W, et al. LECT2 protects Nile tilapia (*Oreochromis niloticus*) against streptococcus agalactiae infection [J]. *Front Immunol*, 2021, 12: 667781.
- [40] KICHI Z A, NATARELLI L, SADEGHIAN S, et al. Orphan GPR26 counteracts early phases of hyperglycemia-mediated monocyte activation and is suppressed in diabetic patients [J]. *Biomedicine*, 2022, 10 (7) : 1736.
- [41] ZHU S, BENNETT S, LI Y, et al. The molecular structure and role of LECT2 or CHM-II in arthritis, cancer, and other diseases [J]. *J Cell Physiol*, 2022, 237 (1) : 480-488.
- [42] LEE S J. Targeting the myostatin signaling pathway to treat muscle loss and metabolic dysfunction [J]. *J Clin Invest*, 2021, 131 (9) : e148372.
- [43] ONO K, HATA K, NAKAMURA E, et al. Dmrt2 promotes transition of endochondral bone formation by linking Sox9 and Runx2 [J]. *Commun Biol*, 2021, 4 (1) : 326.
- [44] LEE S G, CHAE J, WOO S M, et al. TGFBI remodels adipose metabolism by regulating the Notch-1 signaling pathway [J]. *Exp Mol Med*, 2023, 55 (3) : 520-531.
- [45] KIM H J, SON H Y, SUNG J, et al. A genome-wide association study on abdominal adiposity-related traits in adult Korean men [J]. *Obes Facts*, 2022, 15 (4) : 590-599.
- [46] BROWN R D R, SPIEGEL S. ORMDL in metabolic health and disease [J]. *Pharmacol Ther*, 2023, 245: 108401.
- [47] KONG H E, LIM J, LINSALATA A, et al. Identification of PSMB5 as a genetic modifier of fragile X-associated tremor/ataxia syndrome [J]. *Proc Natl Acad Sci U S A*, 2022, 119 (22) : e2118124119.
- [48] HAN J L, LU M X, LI C, et al. Whole-genome sequencing of Ganzi horse reveals the genetic diversity and provides unique insights into its plateau adaptation [J]. *Livest Sci*, 2024, 288: 105549.
- (编辑 郭云雁)