

分类号

密级

UDC



北京林业大学

博士学位论文

基于景观基因组学的高山栎组植物的生态适应研究

Landscape Genomics of Ecological Adaptation in *Quercus* section *Heterobalanus*

罗艳君

指导教师 杜芳教授
学 院 生态与自然保护学院
学 科 专 业 生态学
研 究 方 向 修复生态学

二〇二六年 六 月 五 日

本论文研究工作受国家自然科学基金—联合基金项目（项目批准号：U2571202；项目名称：典型珍稀濒危植物对迁地生境的适应与调节机制）和面上项目（项目批准号：42071060；项目名称：栎类生态适应的分子机理研究）的资助，在此表示感谢！

独创性声明

本人声明所呈交的论文是本人在导师指导下独立进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得北京林业大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

签名： 罗艳君 日期： 2026年6月5日

关于学位论文使用授权的声明

本人完全了解北京林业大学有关保留、使用学位论文的规定，即：研究生在校攻读学位期间论文工作的知识产权单位属北京林业大学；学校有权保留并向国家有关部门或机构送交论文的纸质版和电子版，允许学位论文被查阅、借阅和复印；学校可以将学位论文的全部或部分内容公开或编入有关数据库进行检索，可以允许采用影印、缩印或其它复制手段保存、汇编学位论文。

签名： 罗艳君 导师签名： 栾 日期： 2026年6月5日

答 辩 委 员 会 成 员 信 息

	姓 名	职 称	工 作 单 位
主 席	桑卫国	教授	中央民族大学
委 员	廖万金	教授	北京师范大学
	李景文	教授	北京林业大学
	刘艳红	教授	北京林业大学
	侯继华	教授	北京林业大学

摘要

青藏高原 (Qinghai-Tibet Plateau) 以其独特的环境条件和丰富的物种多样性成为研究物种演化和生物多样性形成的理想自然实验室。研究发现在气候变化驱动下, 许多生物类群出现了走出青藏高原, 并扩散到亚洲、欧洲以及北美洲等地, 对生物多样性形成与演化产生了重要影响。然而, 目前依旧缺乏对高山植物类群的遗传变异格局及其局部环境适应分子机制及其从青藏高原扩散的研究。本研究以青藏高原及其邻近的地区的壳斗科栎属常绿硬叶高山栎组植物 (*Quercus sect. Heterobalanus*) 七个种为研究对象, 通过全基因组重测序, 结合群体基因组学和景观基因组学分析方法对该组内的遗传多样性、谱系分化、进化历史、选择性清除模式以及局地适应进行研究, 并挖掘在环境适应过程中的关键适应性候选基因。主要结果如下:

(1) 高山栎组植物未按照传统形态物种界限进行聚类, 而是呈现出空间地理分化格局。群体遗传结构结果显示 $K=4$ 是高山栎组植物的最佳聚类值, 分为了以主要分布在青藏高原的西藏地区的谱系 A, 在横断山脉地区的谱系 B 和 C, 在秦岭至华东地区的谱系 D。谱系间遗传分化差异明显, 横断山脉地区的谱系 B 和 C 遗传多样性最高, 西藏地区的谱系 A 处于中等水平, 秦岭至华东地区的谱系 D 的最低; 谱系间最大的遗传分化发生在谱系 C 与 D 间 ($F_{ST}=0.2$), 谱系 A 与 B 间最小 ($F_{ST}=0.04$) 且存在明显的谱系 A 向 B 的基因流。种群历史动态分析结果表明, 四个谱系在第四纪期间经历了反复的种群扩张与收缩增加了二次接触可能性。曼特尔检验 (Mantel test) 结果表明高山栎组植物的遗传分化主要受到了地理隔离的驱动 (Mantel $r = 0.39$, $P = 0.0001$); 同时生态位重叠 (Niche overlap) 分析表明谱系间的重叠度偏低甚至无重叠, 各谱系存在着一定的生态环境偏好。

(2) 基因组扫描进行选择性清除分析揭示不同谱系在正选择高分化区域上的明显差异, 谱系分化过程中可能涉及蛋白编码序列的改变和调控变异影响基因表达。谱系 A 与 B 间检测到最多数量的正选择高分化区域。所有谱系对中正选择高分化区域位于基因间区的单核苷酸多态性 (Single nucleotide polymorphism, SNP) 较多, 而位于转录本内的 SNP 所占比例相对较低。不同谱系对之间共鉴定到 910 个候选基因, 功能富集分析显示这些基因富集与发育调控、非生物胁迫响应及生殖过程相关。在

多个谱系对的共有同源基因中鉴定到若干个在不同谱系分化区域中反复出现，并与温度和盐胁迫响应相关的基因。此外仅在青藏高原地区分布的谱系中发现了与叶片大小及形态发育有关的同源基因。顺式调控元件分析显示，11%-38%的候选 SNP 位于基因上下游调控区域，且多与光响应和叶片发育相关元件相关。这些结果表明，高山栎组植物的谱系分化不仅依赖功能保守基因，还通过调控变异实现谱系特异性的适应性演化。

(3) 高山栎组植物各遗传谱系通过不同的环境响应策略实现对山地生态系统环境梯度变化的局地适应。基于遗传分化 (F_{ST}) 和基因型-环境关联 (Genotype-environment associations) 鉴定的推定受选择位点结果显示，谱系 A 鉴定出最多的适应性位点。冗余分析 (Redundancy analysis) 表明，气候变量对高山栎组植物中各谱系的遗传变异贡献度最大。偏冗余分析 (Partial redundancy analysis) 显示谱系 A 和 C 的适应性遗传变异同时受到地理与气候驱动，而谱系 B 和 D 则更多受到气候因素影响。梯度森林 (Gradient analysis) 表明太阳辐射是驱动高山栎组植物适应性遗传变异的重要环境因子，且各谱系对环境变化的响应存在差异。功能富集分析显示适应性候选基因主要富集于光响应、非生物胁迫响应、代谢调控及生长发育等过程，可能与高山栎组植物对不同环境的适应策略密切相关。此外，通过整合各谱系的正选择高分化区域和推定受选择位点，各谱系均有潜在适应干旱胁迫环境的相关基因，促进了高山栎组植物对青藏高原及其邻近地区的干旱生态环境的适应。因此，高山栎组植物的耐旱基因与内在生殖隔离机制可能协同作用，共同驱动其谱系形成和适应性演化。

本研究为理解青藏高原地区木本植物在复杂地质历史和气候变化背景下的物种界定、谱系分化及其适应性进化机制提供了基因组证据，也为在全球气候变化背景下的高山森林生态系统保护与管理提供重要的理论支撑。

关键词：栎属，青藏高原，谱系分化，生态适应，景观基因组学

Landscape Genomics of Ecological Adaptation in *Quercus* section *Heterobalanus*

Doctor Candidate: Yanjun Luo

(Ecology)

Directed by Fang Du

Abstract

The Qinghai-Tibet Plateau (QTP), characterized by its unique environmental conditions and rich species diversity, serves as an ideal natural laboratory for investigating species evolution and diversification. A number of studies found that many species migrated out of the QTP and spread to other regions under climate change, such as Asia, North America and Europe, which played an important role in the formation of biodiversity. However, the genetic variation patterns and the genetic mechanisms of local adaptation are still a lack of alpine plants in QTP region. This study focuses on seven sclerophyllous oak species of *Quercus* sect. *Heterobalanus* in Fagaceae family (*Quercus*), which mainly distributed in the QTP and its adjacent high-altitude areas. This study combined with population genomics and landscape genomics analysis to investigate the genetic diversity, lineage divergence, evolutionary history, selective sweep and local adaptation using whole-genome resequencing, and explore adaptive candidate genes in local adaptation. The main results are as follows:

(1) Genetic cluster of seven sclerophyllous oak species is not consistent with the traditional morphological species boundaries, but showed different geographical regions. The population genetic structure showed that $K = 4$ was the optimal value of *Quercus* sect. *Heterobalanus*, which could be divided into four genetic lineages: lineage A in Tibet region; lineages B and C in the Hengduan Mountains; and lineage D ranging from the Qinling Mountains to East China region. Lineages B and C had higher genetic diversity, lineage A exhibited a moderate level of genetic diversity, while lineage D had the lowest genetic diversity. The relative divergence (F_{ST}) between lineages C and D was the highest ($F_{ST} = 0.2$), while F_{ST} between lineages A and B was the lowest ($F_{ST} = 0.04$). An asymmetric gene flow was observed between different lineages, and the direction was mainly from lineage A to lineage B. Demographic history showed that all lineages was subjected to repeated population expansion and contraction during the Quaternary glaciations, increasing the possibility of secondary contact. The Mantel test showed that geographical isolation may be the main factor driving the genetic differentiation of *Quercus* sect. *Heterobalanus* (Mantel $r = 0.39$, $P = 0.0001$), and the PCA-env plots showed that low or even no overlap was observed among different lineage pairs, suggesting that there is an ecological preference in each lineage.

(2) The genome scan revealed significant differences in highly diverged regions (HDRs) between pairs of different lineages through selective sweep. The process of lineage differentiation may not only the changes in protein-coding sequences but also regulatory variations play a crucial role in the adaptive evolution of lineage divergence. The most HDRs were detected between lineage A and B, which had experienced positive natural selection. More SNPs in positive selection HDRs were located in intergenic regions in all lineage pairs, while the lower proportion of SNPs were located within the transcripts. A total

of 910 candidate genes were identified in positive selection HDRs in different lineages pairs, GO enrichment analysis of the candidate positive selection HDRs were enriched in developmental regulation, abiotic stress response, and reproductive-related processes. Stress response-related gene, such as temperature and salt-stress tolerance, were identified between a lineage and all other lineages. Additionally, homologous genes related to leaf size and morphological development were found only in three lineages distributed in QTP regions. 11% - 38% of candidate SNPs were located in the upstream and downstream regulatory regions of the genes, the cis-regulatory elements revealed that a large proportion of candidate SNPs can be matched with elements related to light response and leaf development. These results indicated that the lineage differentiation of *Quercus* sect. *Heterobalanus* not only depends on functional conserved genes, but also achieves lineage-specific adaptive evolution by regulating variation.

(3) Each genetic lineage of *Quercus* sect. *Heterobalanus* adapted to environmental gradient variations in mountain ecosystems through lineage-specific environmental response strategies to achieve local adaptation. F_{ST} outliers test and genotype-environment association showed that the largest number of putative selected SNPs was identified in lineage A. The redundancy analysis indicated that genetic variation was found to be mainly by climatic variables in each lineage. The partial redundancy analysis showed that the adaptive genetic variation of lineages A and C was driven by both geographical and climatic factors, while lineages B and D was mainly explained by climatic factors. The gradient forest indicated that solar radiation was an important environmental factor driving the adaptive genetic variation of the alpine oak species, and each lineage responded differently to climatic factors. GO enrichment of candidate adaptive genes were enriched in the category of response to light, abiotic stress response, metabolic regulation, and regulation of growth, development. Finally, this study found that each lineage had potential drought-stressed related genes in response to environmental changes through identifying common genes among positive selection HDRs and putative selected SNPs, which promote the local adaptation of alpine plants to the arid ecological environment in QTP and its adjacent areas. Therefore, the drought-tolerant genes and the internal reproductive isolation mechanism of *Quercus* sect. *Heterobalanus* may act synergistically to jointly drive lineage differentiation and adaptive evolution.

This study provides genomic evidence for understanding the species definition, lineage divergence, and the ecological adaptation mechanisms of complex evolutionary history of the QTP alpine plants. It also provides important theoretical foundation for the protection and management of alpine forest ecosystems under the global climate change.

Key Words: *Quercus*, Qinghai-Tibet Plateau (QTP), Lineage divergence, Ecological adaptation, Landscape genomic

目录

1 前言	1
1.1 植物的生态适应	1
1.1.1 植物对环境变化的适应	1
1.1.2 景观基因组学在植物生态适应研究中的应用	3
1.2 青藏高原植物多样性及其形成机制	4
1.2.1 青藏高原植物多样性的驱动因素	5
1.2.2 青藏高原植物多样性的演化模式	6
1.3 栎属植物的生态适应研究	6
1.3.1 栎属植物的生态分布	6
1.3.2 栎属植物的景观基因组学研究背景	7
1.3.3 栎属高山栎组植物的生态适应研究现状	9
1.5 研究目的与内容	11
1.5.1 研究目的	11
1.5.2 研究内容	11
1.5.3 技术路线	12
2 高山栎组植物的遗传结构形成机制	13
2.1 引言	13
2.2 材料与方法	13
2.2.1 样品采集	13
2.2.2 重测序	16
2.2.3 分析方法	17
2.3 结果	21
2.3.1 重测序 SNP 位点统计	21
2.3.2 遗传多样性与遗传结构	22
2.3.3 基因流与种群动态历史	26
2.3.4 地理隔离分析	29
2.3.5 生态位重叠	31
2.4 讨论	33
2.4.1 形态物种与遗传谱系的不一致性	33
2.4.2 谱系间的遗传分化与种群动态历史	35
2.4.3 地理隔离与环境异质性驱动的遗传结构与生态位分化	36
2.5 小结	38
3 高山栎组植物谱系间自然选择信号	39
3.1 引言	39
3.2 分析方法	40
3.2.1 基因组水平分化	40
3.2.2 基因组分化岛检测	40
3.2.3 富集分析	40

3.2.4 顺式调控元件分析	41
3.3 结果	41
3.3.1 基因组分化水平	41
3.3.2 高分化区域	42
3.3.3 选择性信号检测	44
3.3.4 正选择基因	46
3.3.5 顺式作用调控元件	48
3.4 结论	50
3.4.1 正选择的高分化区域	50
3.4.2 选择信号揭示的谱系分化机制	50
3.5 小结	53
4 高山栎组植物谱系的局地适应研究	54
4.1 引言	54
4.2 分析方法	54
4.2.1 环境因子	54
4.2.2 适应性位点检测	54
4.2.3 冗余分析	55
4.2.4 梯度森林	56
4.2.5 适应性候选基因鉴定	56
4.2.6 基因功能注释与富集分析	56
4.3 结果	57
4.3.1 适应性位点鉴定	57
4.3.2 基因型-环境关联分析	59
4.3.3 GO 富集与顺式调控元件	67
4.3.4 适应性候选基因	69
4.4 讨论	71
4.4.1 环境因子驱动适应性遗传分化	71
4.4.2 局地适应的遗传基础	72
4.5 小结	74
5 结论与展望	75
5.1 主要结论	75
5.2 创新点	76
5.3 展望	76
参考文献	78
附录	93
导师简介	108
获得成果目录清单	109

1 前言

1.1 植物的生态适应

人类活动导致全球变暖和更频繁的极端气候事件不断出现，如极端高温、低温与强降雨事件等，对自然种群的生存构成了严重威胁，使得生物多样性资源面临严重丧失的风险，以至于严重的生境破碎化已成为威胁种群续存并导致物种灭绝的主要因素^[1]。目前全球 1/3 的森林已丧失，而现存的森林破碎化严重^[2]。面对持续的环境变化对植物和其他生物构成的日益增长的威胁，如果有足够长的时间，物种可能会适应不断变化的环境。然而，快速的环境变化对世代周期很长的多年生植物具有重大挑战，因此，理解植物如何应对快速变化的环境对于预测和缓解持续气候危机带来的后果至关重要。

1.1.1 植物对环境变化的适应

自然生态系统中，固着生长的植物主要通过三种策略应对快速的气候和环境变化：调整其表型以适应变化的环境，扩大分布范围迁移到更适宜的栖息地或者通过遗传变异适应新的环境^[3-5]。在短期的时间内，固着的生活方式会严重限制植物向新环境的迁移，逃离不适宜的环境效果可能有限。因此，植物应对气候变化的直接和间接的适应能力都将决定物种的灭绝以及生物多样性的维持。

首先，植物对快速气候与环境变化的最直接且响应最快的适应策略为通过调节形态结构、生理过程及资源分配方式，从而呈现出不同表型特征变化。理论上，在资源丰富或水分充足的环境中，植物往往表现出较高的生长速率和较大的叶面积，以提高光合效率和资源获取能力^[6]；而在资源受限或环境胁迫较强的条件下，植物则可能形成较厚的叶片或较低的比叶面积，以减少资源消耗并提高抗逆性^[5]。尽管不同物种在不同资源条件和环境压力下能够通过调整其生长模式和资源分配方式从而提高自身的生存和繁殖能力，但是越来越多的研究发现同一物种在不同生境条件下的表型特征也可能表现出显著的形态差异^[7, 8]。例如，Kieltyk^[9]对中欧塔特拉山的米氏假雏菊（*Bellidiastrum michelii* Cass.）研究发现其形态特征在不同海拔梯度上存在显著差异，如植株高度、叶重和花重等形态特征随着海拔升高而降低或减少；Yang 等^[10]通过对中国亚热带区域同域分布的马尾松（*Pinus massoniana* Lamb.）和黄山松（*P. hwangshanensis* W. Y. Hsia）的表型变异研究发现两个物种的针叶性状（叶厚、叶面积、比叶面积等）的种内变异程度均大于种间变异。此外，植物还能通过调节繁殖性状和物候特征来应对环境变化。如研究表明较大的种子虽然在幼苗建立和资源竞争方面具有优势，但由于其质量较大，往往只能实现较短距离的扩散，而较小的种子则更容易通过风或其他媒介实现远距离传播^[11-13]。种子还与植物在干燥时期的生存和生长有关，提高植物对恶劣环境的适应^[14-16]，如种子较大的物种通常

分布于环境胁迫更强、能量投入需求更高的区域（如高海拔），且在幼苗期的抗旱能力等方面具有优势^[17]。在物候方面，植物通过调控开花时间等关键物候过程，使其生长发育阶段与外界环境条件相协调，从而提高生存和繁殖成功率^[18]。同时，全球气候变化正在显著影响植物的物候过程，大量研究表明全球气候变暖正在改变开花植物的物候期，并进一步影响其地理分布格局^[19, 20]。如 Stinchcombe 等^[21]通过同质园实验（Common garden experiments）对来自 70 种北欧和地中海生态型的拟南芥（*Arabidopsis thaliana*）的研究发现，种群的开花时间在纬度变化上存在显著差异，反映了植物对气候条件变化的适应方式。因此，植物表型特征的变化能够在不发生遗传改变的前提下快速响应环境变化。然而，该策略的适应范围依旧会受到遗传背景和环境差异的限制，当环境变化超过可塑性调节能力时，仅仅依赖植物的可塑性变化是难以维持种群的续存。

其次，分布范围的迁移是植物在空间尺度上通过种子扩散或者种群扩张将分布范围转移至适宜的生境。如欧洲地区的大部分温带森林物种，如栎属植物的夏栎（*Quercus robur* L.）和无梗花栎（*Q. petraea* (Matt.) Liebl.），在第四纪冰期期间被迫退缩至欧洲南部的冰期避难所，冰期结束后的间冰期，栎属植物逐渐向北扩张，形成了当前的分布格局及遗传结构^[22, 23]。类似地，杨属（*Populus*）植物凭借较强的种子扩散能力和生态适应性，在冰后期迅速向高纬度或高海拔地区扩张^[24, 25]。在当前全球气候变暖的背景下，大量研究发现许多温带物种正被迫向更高纬度或高海拔的寒冷地区迁徙^[19, 26]。然而，在短时间内，植物固着性生长的特点使其迁移速度往往滞后于气候变化速度，难以实现有效迁移。如对欧洲阿尔卑斯山区的高山植物研究发现，随着气候变暖，高山植物的适宜生境向高海拔地区迁移，但由于高山地区地形复杂、生境破碎化严重，且许多高山植物的种子扩散能力较弱（主要依靠重力扩散或小型昆虫扩散），导致其分布范围迁移速率远低于气候变暖导致的适宜生境迁移速率，许多物种的分布范围不断缩小，种群数量持续下降^[27-29]。此外，分布迁移策略还受到许多因素的限制，如种子传播能力不足、栖息地破碎化以及人类活动干扰等多种因素均可能阻碍植物通过迁移实现对环境变化的适应。

第三，植物还可以通过遗传变异适应新的环境，即在长期的自然选择作用下，种群对其所处的环境形成局地适应（Local adaptation），一般来说种群在其原生环境中的适合度高于来自其他地区的种群^[30, 31]。局地适应不仅能够反映种群对不同环境选择压力的响应，还在维持物种的适应潜力和促进种群的适应性分化发挥重要的作用^[32]。理论上，许多植物在不同程度上都可能呈现出局部适应特征。短生命周期且自交亲和的物种会在较小的时空尺度上可能形成明显的适应性遗传分化，通常被认为表现出更强的局地适应^[33]。相比之下，对广泛分布且处于环境异质性较强的长寿植物而言，有限的表型可塑性可能难以应对复杂多变的环境条件，特别是在快速变化的气候条件背景下，种群内的遗传变异成为了其种群维持适应能力和进化潜力的

重要基础^[34]。局地适应的形成通常受到多个因素的共同影响，其中最关键的两个因素为基因流和遗传变异。频繁的基因流可能会抵消局部自然选择的作用，从而阻碍局地适应的形成；而遗传变异的不足则会限制种群对环境选择压力的响应能力^[35]。然而，基因流在局地适应过程中也可能具有双重作用。当种群内部遗传变异较为有限时，适当的种间基因交流通过引入新的等位基因，从而提高种群的遗传多样性，并增强其对自然选择的响应能力^[36, 37]。此外，随着种群间距离的增加而增加，种群间的基因流减少，不同种群间的遗传分化增强，提高局地适应的可能性^[37, 38]。

1.1.2 景观基因组学在植物生态适应研究中的应用

尽管植物可以通过不同的方式实现对环境变化的适应，但局地适应在维持生物多样性及应对全球气候变化过程中具有重要作用，不仅决定了种群在不同环境中的适合度差异，也是驱动种群遗传分化和物种进化的重要基础^[32, 39]。植物种群中局地适应的遗传基础研究往往采用数量遗传学方法，即将来自不同种源的种群中的植物栽培于相同环境条件下，通过交互移植实验（Reciprocal transplant experiments）或同质园实验比较种群在不同自然环境中的适应性表型和遗传差异，从而评估种群的局地适应能力^[32, 40]。该方法已在许多森林种源研究中得到了应用，如中欧的主要本土树种-欧洲水青冈（*Fagus sylvatica* L.）^[41]，北美的红云杉（*Picea rubens* Sarg.）^[42]等，为理解林木种群的适应性分化提供了重要证据^[43, 44]。尽管交互移植实验或同质园实验成为了研究植物适应性的有力工具，但是随着基因组学技术与分析工具的快速发展和应用，利用基因组学数据研究局地适应的方法越来越多。尤其是在难以开展或缺乏条件进行实验研究的情况下，整合环境空间数据和基因组数据对物种进行景观基因组学（Landscape genomics）的研究为探索自然种群的局地适应提供了新的途径^[45, 46]。

新兴的景观基因组学综合了群体遗传学、景观生态学和空间分析方法等多学科方法，通过整合遗传和空间分布数据研究自然种群的遗传变异与景观环境因子之间的相互作用关系，揭示环境因子对遗传结构及适应性遗传变异的影响，以及评估物种在气候变化背景下的适应潜力和脆弱性，并且还能够识别参与局地适应的遗传位点^[47, 48]。景观基因组学方法通常基于基因型-环境关联分析（Genotype-environment associations, GEAs）来识别可能参与局地适应的遗传位点以及揭示遗传变异与环境因子之间的关系，从而解析物种生态适应的遗传基础^[49, 50]。目前常用的方法主要包括基于广义线性混合模型的 Bayenv 和潜在因素混合模型（Latent factor mixed modelling, LFMM）等方法，常用于鉴定与环境变量相关的异常位点；多变量统计分析，如蒙特尔检验（Mantel test）和冗余分析（Redundancy analysis, RDA）等方法，常被用于评估遗传距离与环境或地理距离之间的关系；以及一些非线性模型，如梯度森林（Gradients forest, GF）和广义相异建模（Generalized dissimilarity modelling, GDM）等方法，通常被广泛用于评估个体或种群的等位基因频率沿着环境梯度的变

化趋势以及预测遗传变异在环境梯度上的空间分布格局^[50, 51]。

近年来, 景观基因组学方法已被成功用于探究物种局地适应的研究并且极大的促进了对许多非模式物种的局地适应性遗传基础的研究^[47, 50]。如科学家利用简化基因组测序 (Restriction-site associated DNA sequencing, RAD-seq) 获得的单核苷酸多态性 (SNPs) 数据对台湾杉 (*Taiwania cryptomerioides* Hayata) 进行景观基因组学研究发现, 该物种在中国的分布范围内存在明显的遗传分化, 且环境隔离对其空间遗传变异的塑造作用强于地理隔离, 同时鉴定出了多个与高温、干旱和盐胁迫响应的相关的候选位点, 并发现降水对驱动该物种的适应性方面有着重要作用^[52]。类似地, 对日本柳杉 (*Cryptomeria japonica* (Thunb. ex L. f.) D. Don) 的研究发现, 该物种的遗传变异分布模式呈现地理分布格局, 并鉴定出 239 个与气候相关的位点, 其中温度和降水被认为是驱动适应性遗传变异的重要环境因子^[53]。随着高通量测序技术的持续发展, 基于全基因组重测序 (Whole-genome resequencing, WGS) 对粗皮桉 (*Eucalyptus pellita* F. Muell.) 的研究表明, 该物种的遗传结构同样表现出明显的地理分化, 并鉴定出 198 个与气候相关的适应性位点, 涉及与温度、水分和盐响应相关的 45 个功能基因^[54]。因此, 景观基因组学为揭示自然种群适应性遗传变异及其环境驱动机制提供了重要手段, 并在解析物种局地适应及预测其对气候变化响应方面发挥着越来越重要的作用, 而对处于高度环境异质性的物种进行局地适应研究将有助于进一步理解植物适应性遗传变异的形成及其环境驱动机制。

1.2 青藏高原植物多样性及其形成机制

青藏高原 (Qinghai-Tibet Plateau, QTP) 是全球海拔最高和面积最大的高山系统, 面积达 250 万平方公里, 平均海拔超过 4500 米, 有着“世界屋脊”与“地球第三极”的称号^[55, 56]。青藏高原的形成主要源于新生代时期的印度板块与欧亚板块的持续碰撞和隆起, 这一地质事件不仅塑造了其独特的地貌形态, 也深刻影响了亚洲季风系统以及周边地区的气候环境^[57, 58]。从地理分布范围来看, 青藏高原的南起喜马拉雅山, 北至昆仑山和祁连山脉, 西起帕米尔高原和喀喇昆仑山, 东至横断山脉^[59]。现在普遍认为青藏高原包括横断山脉、喜马拉雅山和高原台地 (包括其中的帕米尔高原)^[60], 在中国分布范围涵盖了整个西藏自治区和青海省、四川省西部、新疆维吾尔自治区南部, 以及甘肃省和云南省的部分区域。虽然青藏高原拥有独特且复杂的自然环境, 如强辐射、气候寒冷、低氧和干旱等极端环境, 但是大量研究表明地质活动引发的地貌改变和气候波动导致的复杂气候等因素促进了该地区的高度环境异质性, 为高山植物的物种分化和适应性演化提供了多样化的生态位, 使其成为了世界上重要的高山生物多样性热点地区之一^[61]。因此, 青藏高原不仅是研究物种进化与生物多样性研究的重要天然实验室, 也是探索植物多样性形成与高海拔环境适应机制的重要区域。对青藏高原地区植物多样性格局及其形成机制的研究, 有助于深

入理解高山生态系统的演化历史，并为全球气候变化背景下的生物多样性保护提供重要科学依据。

1.2.1 青藏高原植物多样性的驱动因素

青藏高原复杂的地质历史和高度异质性的生态环境为植物多样性的形成提供了重要条件。研究表明，青藏高原及其邻近地区有着 20000 余种的种子植物，其中青藏高原东南部地区，尤其是在横断山脉地区具有丰富的物种多样性，并存在大量的特有高山植物^[62-64]。因此，青藏高原及其邻近地区拥有北温带最高的物种多样性，其中高山植物多样性尤为明显^[65]。

首先，青藏高原的隆起过程是驱动植物多样性形成的关键因素之一。造山运动不仅造成了青藏高原及其邻近地区的海拔差异和栖息地破碎，还促进了地理隔离和异域物种形成，为许多植物的谱系分化和适应性演化提供了重要的条件，特别是在中新世早期至第四纪期间，青藏高原发生的多次隆起事件已经被证实促进了植物的物种形成与多样化^[65, 66]。一些植物谱系的起源可能早于青藏高原的形成或早期隆升阶段，但大量分子系统学研究表明，许多高山植物类群的物种分化主要发生在晚中新世至第四纪期间^[65, 67, 68]。例如，Wang 等^[69]对蓼科 (Polygonaceae) 大黄属 (*Rheum* L.) 的 26 个物种的叶绿体 DNA 片段进行系统发育分析表明该属在上新世至第四纪约 7 Ma 期间经历了显著的多样化。菊科 (Asteraceae) 华蟹甲属 (*Sinacalia* H. Rob. & Brettell) 和橐吾属 (*Ligularia* Cass.) 等类群也在这一时期经历了显著的物种辐射或谱系分化^[67]。

第二，第四纪气候振荡是植物多样性形成和维持的重要驱动因素之一，在促进青藏高原及其邻近地区的物种形成与谱系分化过程中具有重要作用。频繁的冰期-间冰期的气候变化不仅改变了物种的分布格局，还促进了该地区许多特有物种发生适应性辐射分化，从而显著地影响并重塑了青藏高原地区的物种分布格局和多样化^[66, 70]。多次冰期的发生让青藏高原大部分地区被冰川覆盖，植物种群不得已向未被冰川覆盖的山地避难所或局地微避难所收缩，如青藏高原东南缘和横断山脉地区等重要的“避难所”^[71]。因此，在冰期期间的青藏高原保存了大量的古老或孑遗植物，如西藏红豆杉 (*Taxus wallichiana* Zucc.)、水青树 (*Tetracentron sinense* Oliv.) 和巨柏 (*Cupressus gigantea* W. C. Cheng & L. K. Fu) 等^[72]。伴随着冰期的结束和间冰期的气候回暖，植物种群开始从避难所向高海拔或高纬度地区进行重新扩张，来自不同避难所的种群扩散过程中发生二次接触，部分物种间可能发生杂交与渐渗，从而促进种群分化甚至推动物种形成^[73]。例如，对青藏高原的四数獐牙菜 (*Swertia tetraptera* Maxim.) 的谱系地理学研究表明，该物种在更新世冰期期间可能存在多个微型避难所 (microrefugia)，这些避难所为种群遗传多样性的保存提供了条件，并在气候回暖时期促进了种群扩张^[74]。类似的研究在金露梅 (*Potentilla fruticose* (L.) Rydb.) 中也得到了结论支持，其叶绿体 DNA 分析显示，该物种在冰期与间冰期之

间经历了显著的分布范围变化，并形成了复杂的种群遗传结构^[75]。总体而言，第四纪气候振荡通过驱动物种分布的周期性收缩与扩张，维持避难所物种的遗传多样性，并促进种群之间的二次接触与基因交流，在塑造青藏高原及其邻近地区植物的遗传结构和物种多样性方面发挥了重要作用。

1.2.2 青藏高原植物多样性的演化模式

许多分子系统学和谱系地理学研究表明，青藏高原的许多植物类群从欧亚大陆其他地区迁入后在高山环境中发生了适应性辐射。如 Ebersbach 等^[76]基于约 60% 的虎耳草科 (*Saxifragaceae*) 虎耳草属 (*Saxifraga* Tourn. ex L.) 植物研究表明该属可能起源于北美洲地区，并在早期分化期间向南美洲和北亚扩散，然后在晚始新世时在欧洲和青藏高原地区定殖，随着青藏高原的隆升发生快速分化并从西藏地区向外扩散。Lai 等^[77]通过对兰科 (*Orchidaceae*) 兰亚族 (*Orchidinae*) 物种的系统发育研究表明，该属物种的多样性是由于原地分化、迁入和本地繁殖等多种方式共同作用而形成的，并且地理分析揭示存在 29 次从亚洲、非洲和欧洲向青藏高原的扩散事件，以及 15 次从青藏高原向外的扩散事件，而且该属的大多数特有物种是在过去的 600 万年里进化而来的。一些植物类群则可能起源于青藏高原，并在该地区进行了快速分化，形成大量特有种，后续甚至扩散至其他地区。如 Zhang 等^[78]对景天科 (*Crassulaceae*) 红景天属 (*Rhodiola* L.) 55 个物种的系统发育研究发现，该属多个物种在青藏高原隆升期间发生了快速分化，祖先重建表明红景天科起源于青藏高原，然后扩散至北半球的其他地区。如类似的研究还包括胡颓子科 (*Elaeagnaceae*) 沙棘属 (*Hippophae* L.)^[79]、十字花科 (*Brassicaceae*) 双果芥属 (*Megadenia* Maxim.)^[80]、龙胆科 (*Gentianaceae*) 龙胆属 (*Gentiana* (Tourn.) L.) 部分谱系^[81]、忍冬科 (*Caprifoliaceae*) 忍冬属 (*Lonicera* L.)^[82]等研究物种也得到了相似的结论。科学家根据以上这些研究，提出了“走出青藏高原”假说 (*Out-of-Tibet Hypothesis*)^[83, 84]，指出青藏高原不仅是高山植物多样性形成的摇篮，也可能是部分欧亚高山植物类群的起源地，这些类群随后通过高山走廊向周边地区扩散^[60, 61]。因此，青藏高原植物多样性的形成并非单一来源，而是由原地分化、迁入以及向外扩散等多种过程共同塑造，其中原地适应性辐射和“走出青藏高原”模式在植物多样性的形成中发挥了重要作用。

1.3 栎属植物的生态适应研究

1.3.1 栎属植物的生态分布

壳斗科 (*Fagaceae*) 栎属 (*Quercus* L.) 是北半球温带和亚热带森林生态系统中的重要森林树种，全球约有 400-600 种^[85]。最新的栎属分类系统认为栎属分为两个亚属和八个组^[86]。其中，栎亚属 (*Subgenus Quercus*) 包括广泛分布于北美洲和欧亚大陆的白栎组 (*Section Quercus*)，分布于北美洲和欧洲的本都栎组 (*Section*

Ponticae), 仅分布于北美洲的中间栎组 (Section *Protobalanus*)、红栎组 (Section *lobatae*) 和活栎组 (Section *Virentes*); 土耳其栎亚属 (Subgenus *Cerris*) 包括仅分布于东亚的青冈组 (Section *Cyclobalanopsis*), 欧亚大陆的冬青栎组 (Section *Ilex*) 和土耳其栎组 (Section *Cerris*) [86]。

美洲栎属植物种类最多, 估计约 255 种, 占全球栎属物种总数的近 60%, 主要分布于北美洲, 中美洲和南美洲仅零星分布, 其中墨西哥 (约 170 种) 和美国 (91 种) 是栎属物种多样性最丰富的两个国家, 墨西哥更因占全球栎属物种 38% 的占比, 成为全球栎树物种多样性最多的国家 [87, 88]。亚洲也是全球栎属植物的物种多样性分布地区之一, 约有 150 余种, 占全球总数的 35% 以上, 主要集中分布于东亚及东南亚部分区域, 中国是栎属植物多样性最丰富的国家之一 [86]。欧洲栎属植物有 29 余种, 主要分布于中欧平原、西欧大西洋沿岸及南欧地中海沿岸山地 [89]。

1.3.2 栎属植物的景观基因组学研究背景

栎属植物往往存在较高的遗传多样性、较强的迁移扩散能力以及频繁的种间杂交与渐渗现象, 使其能够在不同环境条件下实现快速扩散和适应, 从而形成广泛的生态分布格局 [87, 90, 91]。已有研究表明栎属不同物种在温度、降水、土壤及海拔等环境梯度上表现出明显的生态分化, 使其能够成为研究环境变化与局地适应的重要植物类群 [47, 92]。此外, 栎属植物普遍存在的种间杂交与渐渗现象为解析自然选择与基因流共同作用下的适应性进化提供了理想研究材料 [93]。因此, 将景观遗传学方法与基因组学相结合开展景观基因组学研究, 有助于揭示环境异质性对栎属植物的遗传结构及适应性进化的影响。

随着高通量测序技术的发展和應用, 栎属的基因组学研究越来越受到关注, 自 Sork 等 [94] 发表北美地区的加州白栎 (*Q. lobata* Née) 基因组以来, 标志着栎属进入了基因组时代, 自此不断有新的栎属基因组资源完成开发并释放。近年来, 亚洲地区的蒙古栎 (*Q. mongolica* Fisch. ex Ledeb.)、槲树 (*Q. dentata* Thunb.) 和长穗高山栎 (*Q. longispica* (Hand.-Mazz.) A. Camus) 等, 以及北美洲地区的加州白栎 (*Q. lobata*)、北美红栎 (*Q. rubra* L.) 和岛栎 (*Q. tomentella* Engelm. & A.Gray) 等的基因组均被释放 (表 1.1)。这些基因组资源不仅为栎属植物的系统发育关系和进化历史提供了重要数据, 还为局地适应、适应性渐渗及挖掘与功能形态相关的基因奠定了基础, 从而为研究栎属植物对生态环境变化响应的适应性进化机制提供理论基础。

表 1.1 栎树植物的全基因组测序信息

Table 1.1 Genomic resources are available for *Quercus*

物种 Species	组别 Section	地点 Distribution	基因组大小 Genome size (Mb)	参考 Reference
蒙古栎 <i>Q. mongolica</i>	白栎组	东亚	809.84	Ai et al., 2022 [95]
加州白栎 <i>Q. lobata</i>	白栎组	北美	810	Sork et al., 2022 [96]

物种 Species	组别 Section	地点 Distribution	基因组大小 Genome size (Mb)	参考 Reference
夏栎 <i>Q. robur</i>	白栎组	欧洲	790	Plomion et al., 2018 ^[90]
北美红栎 <i>Q. rubra</i>	白栎组	北美	739	Kapoor et al., 2023 ^[97]
槲树 <i>Q. dentata</i>	白栎组	东亚	893.54	Wang et al., 2023 ^[98]
欧洲栓皮栎 <i>Q. suber</i>	冬青栎组	欧洲	953.3	Ramos et al., 2018 ^[99]
冬青栎 <i>Q. ilex</i>	冬青栎组	欧洲	842.2	Rey et al., 2023 ^[100]
长穗高山栎 <i>Q. longispica</i>	冬青栎组	东亚	901.31	Ren et al., 2025 ^[101]
赤皮青冈 <i>Q. gilva</i>	青冈组	东亚	889.71	Zhou et al., 2022 ^[102]
西畴青冈 <i>Q. sichouensis</i>	青冈组	东亚	848.75	Liu et al., 2024 ^[103]
大果青冈 <i>Q. rex</i>	青冈组	东亚	883.46	Liu et al., 2024 ^[103]
麻栎 <i>Q. acutissima</i>	麻栎组	东亚	758	Fu et al., 2022 ^[104]
栓皮栎 <i>Q. variables</i>	麻栎组	东亚	791.89	Chang et al., 2023 ^[105]
岛栎 <i>Q. tomentella</i>	中间栎组	北美	781	Mead et al., 2024 ^[106]
白栎 <i>Q. alba</i>	白栎组	北美	794	Larson et al., 2025 ^[107]

结合参考基因组，运用景观基因组学方法可以评估物种或谱系的遗传多样性、局地适应性、景观连通性以及生态脆弱性等。栎属最早进行景观基因组学研究的是 Martins 等^[108]以加州白栎为参考基因组，应用简化基因组对墨西哥地区白栎组的皱纹栎 (*Q. rugosa* Née) 进行研究表明该物种存在显著的遗传分化，揭示了季节性降水是驱动空间遗传变异的重要环境因子，并鉴定出了一些与当地环境适应相关的功能基因。随后运用同样的参考基因组，Gugger 等^[109]对加州白栎进行研究表明在种间基因流较强的背景下，依旧发现多空间尺度下气候驱动的自然选择影响了加州白栎的适应性遗传分化，并鉴定出大量参与气候适应的候选基因。以北美红栎为参考基因组，Sun 等^[110]对长果青冈 (*Q. longinux* Hayata) 进行研究表明环境隔离是驱动其适应性遗传变异的重要因素，并进行了遗传偏移和非适应性风险分析评估对未来气候变化的适应能力。Quijano 等^[111]对北美红栎进行研究表明温度是驱动其局地适应的适应性遗传分化的重要气候因子。在全基因组重测序的背景下，Mead 等^[112]以加州白栎为参考基因组，对北美分布的岛屿栎进行研究表明地理和气候因子共同影响了其遗传变异，其中气候因子对遗传变异的影响更大，并进行了遗传偏移的评估。Ren 等^[101]以长穗高山栎为参考基因组，对川滇高山栎 (*Q. aquifolioides* Rehder & E. H. Wilson) 及其近缘种进行研究，揭示了同域分布的近缘种间发生的渐渗是驱动物种内适应性遗传变异的重要因素。此外，Meger 等^[113]以夏栎为参考基因组，采用序列捕获 (Sequence capture) 对夏栎进行景观基因组学研究，筛选出了与地理、气候和表型变量相关的受选择位点，揭示了其潜在的局地适应能力，并评估了分布范围内的气候变化脆弱性区域。

景观基因组学的研究也可以不依赖于参考基因组，极大的促进了非模式物种的局地适应研究，能够有效解析环境因子驱动物种的适应性遗传变异的机制。因此，在无参考基因组条件下，Vanhove 等^[114]对地中海地区的欧洲栓皮栎 (*Q. suber* L.) 进

行景观基因组研究表明温度和降水是影响欧洲栓皮栎局地适应的重要气候因子，并预测了未来气候变化情境下的遗传脆弱性。

尽管近年来栎属的景观基因组学研究取得了重要进展，但目前的研究仍主要集中于欧洲和北美地区的栎属物种，而关于东亚地区栎属的景观基因组学研究仍然相对较少。因此，开展东亚栎属植物的景观基因组学研究，对于揭示环境异质性下的适应性进化机制具有重要意义。

1.3.3 栎属高山栎组植物的生态适应研究现状

(1) 高山栎组植物基本特征及生态意义

栎属高山栎组植物隶属于栎属冬青栎组 (*Quercus sect. Ilex* Loudon) 中的一个单支系，并在中新世中期经历了快速分化^[115-117]。对高山栎组植物的早期研究主要依据化石证据、地理分布及叶片结构等形态特征进行鉴别与初步分类^[118]。周等^[119]进一步结合分子生物学和形态学的证据对高山栎组植物复杂的分类问题进行了梳理，明确现在被广泛接受的 7 至 11 个物种。《中国植物志》在第 1998 年版的 22 卷将该组植物划分为 11 个物种，分别为川滇高山栎、高山栎 (*Q. semecarpifolia* Sm.)、灰背栎 (*Q. senescens* Hand.-Mazz.)、帽斗栎 (*Q. guyavifolia* H. Lév.)、矮高山栎 (*Q. monimotricha* (Hand.-Mazz.) Hand.-Mazz.)、毛脉高山栎 (*Q. rehderiana* Hand.-Mazz.)、黄背栎 (*Q. pannosa* Hand.-Mazz.)、长苞高山栎 (*Q. fimbriata* Chun et Huang)、长穗高山栎 (*Q. longispica* (Hand.-Mazz.) A. Camus)、光叶高山栎 (*Q. pseudosemecarpifolia* A. Camus) 和刺叶高山栎 (*Q. spinosa* David ex Franch.)。随后在 1999 年的《中国植物志》(英文修订版, FOC) 将该组植物划分为 7 个物种，分别为川滇高山栎、高山栎、毛脉高山栎、灰背栎、帽斗栎、矮高山栎和刺叶高山栎^[120]。

高山栎组植物分布范围为南起泰国，北至秦岭，东达台湾，西止于喜马拉雅地区并上延至阿富汗地区；在中国，高山栎组植物分布于西南的青藏高原地区及其邻近地区，其中绝大多数物种向东止于四川盆地的西部边缘，只有刺叶高山栎从横断山脉地区一直零星分布到台湾地区^[119]。高山栎组植物的海拔分布从 1700 m 到 4800 m，以 2400 m 到 3600 m 的海拔最为集中^[119]。基本形态特征为叶片均为常绿、厚革质，叶片坚硬，叶面常具光泽、皱褶不平或平坦，叶柄粗短；叶片形状通常为椭圆形、长椭圆形、倒卵形、卵状披针形；先端多钝圆、短渐尖、少急尖；基部楔形、圆形或浅心形；叶缘全缘或有锯齿，叶背普遍具棕色至黄褐色星状毛、腺毛或糠秕状鳞秕，老叶的叶背毛有时会脱落，叶面多无毛或幼时疏被星状毛；壳斗呈现杯形或浅杯形、帽斗状、浅碗形或碟形；坚果为卵形、长卵形、近球形或椭圆形^[119, 120]。因此，高山栎组植物同域分布、形态特征相似度高以及频繁的杂交渐渗等特点，使其在分类界定中长期存在争议，为一个典型的物种复合体。

高山栎组植物通常是高山及山地生态系统中的优势种或建群种，在维持生态系统的结构、功能以及生物多样性方面发挥着重要作用。在生态系统结构方面，高山

栎组植物对低温、干旱、贫瘠土壤及强紫外辐射具有较强耐受性，能够在林线以上或陡峭峡谷等极端环境中生长。其发达的根系和稳定的生长形态有助于固定土壤、减少水土流失，并维持陡坡及碎石坡等不稳定生境中的群落结构完整性^[121]。在生态系统稳定性方面，高山栎组植物作为木本植物，通常具有较高生物量和较长生命周期，是高山生态系统重要的碳汇组成部分，对区域碳循环具有重要调节作用；同时，其凋落物为土壤提供稳定的有机质来源，促进养分积累和微生物群落形成，从而维持土壤生态系统功能稳定^[122]。在生物多样性维持方面，高山栎组植物形成的林下及灌丛微生境为多种植物、真菌和动物提供栖息环境，其叶片、果实及枝干为昆虫、鸟类和小型哺乳动物提供重要食物资源^[123]。此外，高山栎组植物与菌根真菌的共生关系有助于其适应贫瘠土壤，并促进地下生物多样性和养分循环。同时，高山栎组植物长期以来也是当地居民重要的生活资源，如饲料、建材及薪炭材等。

(2) 高山栎组植物的形态特征对环境变化的适应策略

高山栎组植物表现出了对极端环境的显著适应能力，其种内叶形态变异尤为显著，同一物种、同一种群甚至同一植株的叶片也呈现明显的多样性，可以为椭圆形、卵形或倒卵形，叶基部圆形或浅心形，叶缘有齿或无齿以实现环境的适应^[124]。Sun 等^[125]研究发现帽斗栎随着海拔升高，其叶面积和叶长/宽比显著降低均与年均降水量呈显著正相关，而叶长/宽比和叶柄长度则与生长季日均最大 UV-B 辐射强度呈显著负相关。Bai 等^[126]通过对石漠化林地中的毛脉高山栎的叶片性状研究表明石漠化林地中具有更大的叶片厚度和更小的叶面积，呈现“资源保守性策略”以适应石漠化生境。针对川滇高山栎的研究表明，其叶片形态和气孔性状的变异是遗传、地理分布及气候条件共同作用的结果，其中年均降水量可能是驱动气孔性状变化的关键气候因子^[127, 128]；尤其是与刺叶高山栎的叶形态比较分析进一步表明川滇高山栎对严酷的生境具有较强适应性^[129]。通过对乌蒙山的刺叶高山栎叶功能性状对环境的响应研究发现，不同海拔梯度的刺叶高山栎采取了不同的资源策略以适应变化的环境^[130]。这些研究反映了高山栎组植物对水分、温度及辐射等环境因子变化的响应，也体现了其在复杂环境梯度变化下的适应策略与生存权衡。

(3) 高山栎组植物的群体遗传学和基因组学研究

近年来已有大量研究利用群体遗传学或群体基因组学探讨高山栎组植物的遗传结构及历史动态。Meng 等^[131]利用 9 个微卫星分子标记和 4 个叶绿体片段对分布于青藏高原及其邻近地区的高山栎组植物的七个种研究表明，生活在相似环境中的种群在遗传结构上更相似。Ju 等^[132]基于特异性位点扩增片段测序 (SLAF-seq) 技术对刺叶高山栎进行研究，发现其种群存在两个明显的地理隔离的遗传谱系。Du 等^[133, 134]基于微卫星分子标记、叶绿体片段和简化基因组测序数据研究表明，川滇高山栎存在显著的种内遗传谱系分化并表现出明显的局地适应特征。Liu 等^[135]通过微卫星分子标记技术对川滇高山栎和刺叶高山栎进行研究发现两个物种均存在明显的种内遗

传分化，并进一步通过基因型-环境关联分析表明温度和降水分别是两个栎属植物不同遗传谱系遗传变异的关键环境因子。随着基因组学技术的发展，Ma 等^[136]运用全基因组重测序技术对 11 个高山栎组植物的研究表明位于青藏高原地区的所有种群形成了一个杂合体 (Syngameon)。Ju 等^[137]基于 SLAF-seq 技术对高山栎组植物的七个种研究表明，个体间的遗传聚类同时受到地理距离和祖先遗传成分的影响。Ren 等^[101]运用全基因组重测序技术研究发现川滇高山栎与其同域近缘种之间存在双向的历史基因流，揭示了高山栎组植物物种间复杂的遗传结构关系。

尽管当前关于高山栎组植物的研究已从传统的形态分类学发展到分子系统学、群体遗传学和基因组学等多个研究层面，但在景观尺度上探究环境因子驱动的遗传分化及适应性进化机制的研究仍相对缺乏。结合群体基因组学、景观基因组学等多种方法进一步探讨高山栎组植物的遗传分化机制、基因流及其对环境变化的适应机制，为理解高山栎组植物的物种形成过程、遗传分化机制及其对高山环境的适应策略提供重要理论依据。

1.5 研究目的与内容

1.5.1 研究目的

中国栎属高山栎组植物主要分布于青藏高原及其邻近区域，是该地区高山森林、灌丛生态系统中的优势种或建群种。高山栎组植物可能受到青藏高原地区复杂的地质历史和第四纪气候波动的影响，是研究栎属植物遗传结构形成机制与环境适应性的理想研究材料。本研究以高山栎组植物的七个种为研究对象，在其分布范围内进行大范围的取样后，利用全基因组重测序技术，采用群体基因组学与景观基因组学的分析方法开展高山栎组植物的遗传结构、进化历史和适应性进化机制研究，系统探讨驱动高山栎组植物的遗传分化及环境适应的关键机制。本研究不仅为理解复杂的地质历史和气候振荡影响下的森林树种的物种界定、谱系分化和局地适应机制提供新的理论依据。

1.5.2 研究内容

本研究通过全基因组重测序对高山栎组植物的遗传结构和环境适应性进行了系统研究，研究内容具体如下：

(1) 利用群体基因组学开展高山栎组植物的遗传多样性、遗传结构和种群历史动态等分析，阐明历史演化过程对高山栎组植物的物种形成和谱系分化历程的影响，进一步为高山栎组植物的物种界限划分提供参考依据；

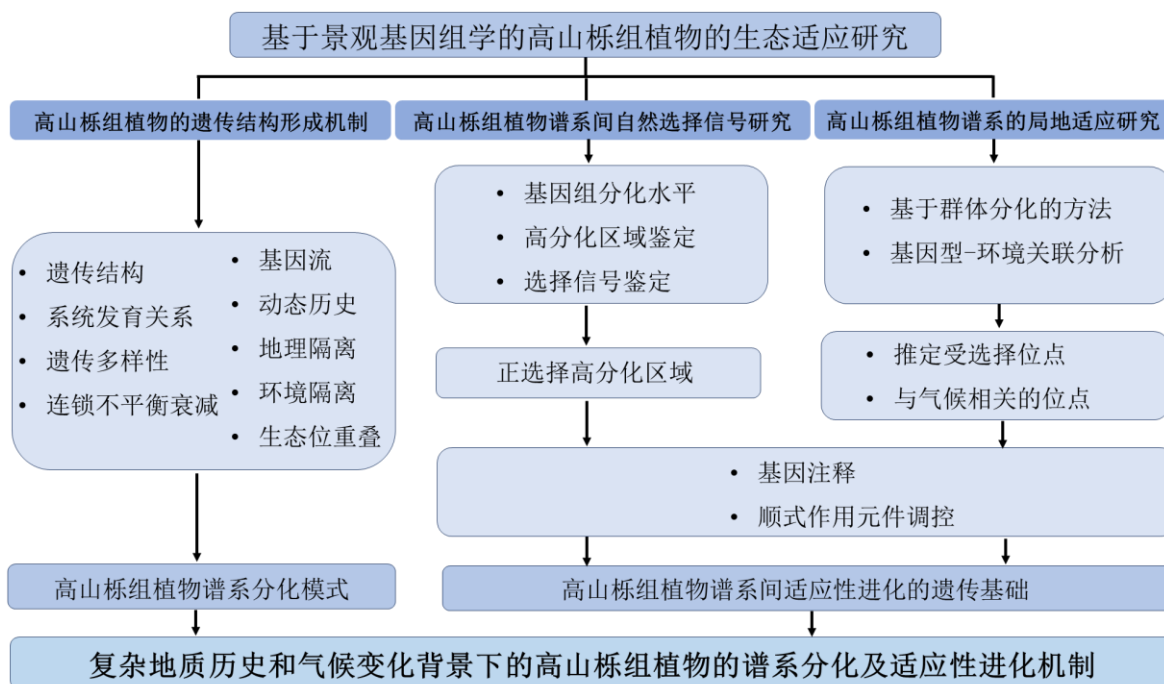
(2) 结合地理和气候因子，利用曼特尔检验评估地理与气候因子对遗传变异的影响，通过生态位重叠分析解析高山栎组植物中不同谱系的生态位分化程度；

(3) 通过基因组扫描比较不同谱系间的高分化区域与选择性清除信号，结合 GO 富集、功能注释和模式植物拟南芥同源基因比对，筛选潜在的受选择基因，评估

自然选择在谱系分化中是否具有谱系特异性；

(4) 运用景观基因组学方法识别驱动不同谱系的适应性遗传变异的关键环境因子，并鉴定驱动不同谱系产生环境适应分化的关键适应性候选基因，揭示高山栎组植物的局地适应的遗传基础。

1.5.3 技术路线



2 高山栎组植物的遗传结构形成机制

2.1 引言

遗传结构是指遗传变异在不同种群之间的分布格局^[138, 139]，其形成与维持受到多种进化过程的共同作用，包括基因流、遗传漂变、自然选择^[140]。在自然种群中，遗传结构不仅能够反映物种或谱系的进化历史，也能揭示环境变化对种群遗传分化和适应性进化的影响。因此，解析物种遗传结构的形成机制，对于理解物种的形成过程、局地适应以及生物多样性的维持具有重要意义。

复杂的地理格局和多样化的气候环境通常会对植物种群的遗传结构产生影响。地理隔离阻碍植物种群间的基因交流，种群间产生遗传分化^[141, 142]。气候环境的变化能够通过自然选择驱动植物种群产生适应性遗传分化，进而促进局地适应；并通过限制基因流促进种群遗传分化^[143]。大量研究发现许多植物的遗传结构同时受到了地理格局和气候变化的影响，特别是在青藏高原及其邻近地区，剧烈的地质历史活动和第四纪气候波动对植物区系的演化产生了重要影响，被认为是全球山地生物多样性形成的重要驱动因素之一^[61, 65]。因此，许多主要分布在青藏高原地区的高山植物均表现出明显的遗传分化现象^[65, 144]。近年来，随着高通量测序技术的发展，基因组学信息被广泛应用于揭示自然种群遗传结构；同时，通过结合遗传数据解析与环境因子之间的关系为深入理解遗传分化机制提供了研究思路。

栎属冬青栎组中的高山栎组植物主要分布于青藏高原及其周边地区，通常生长在海拔较高以及环境条件复杂的山地地区，不同种群之间往往被复杂的地形屏障和显著的环境梯度所隔离，从而为研究地理隔离和环境差异对遗传结构形成的影响提供了理想的研究群体。因此本章节的研究运用以川滇高山栎为参考基因组进行全基因组重测序生成的高质量单核苷酸多态性（Single nucleotide polymorphism, SNP）数据，对高山栎组植物中的七个种进行遗传结构和系统发育分析，解析其谱系分化模式，并进行进化历史、空间遗传结构与生态位分化的研究。

2.2 材料与方法

2.2.1 样品采集

共收集 68 个自然种群的 109 个冬青栎组的高山栎组植物个体，覆盖了高山栎组植物的七个种在中国地区的主要分布范围（图 2.1）。采集时，种群内个体之间的距离保持在至少 100 米，以尽可能避免克隆繁殖植株的影响。每个种群需要平均 2 株个体的新鲜叶片，将采集的个体进行对应种群及个体的编号，擦拭干水分后放入对应编号的自封袋中并放入硅胶进行干燥。并记录每个采样点的纬度、经度和海拔信息。此外，每个种群采集制作 1 份标本，并详细记录采集者、地点、日期和生境等基本信息。参照《中国植物志》中描述的物种形态特征对采集的样本进行初步物种

鉴定，样本均保存于实验室。其中川滇高山栎 18 个种群 24 株个体，刺叶高山栎 18 个种群 25 株个体，帽斗栎 6 个种群 12 株个体，矮高山栎 7 个种群 12 株个体，毛脉高山栎 7 个种群 14 株个体，高山栎 5 个种群 10 株个体，灰背栎 7 个种群 12 株个体（表 2.1）。在经费有限的情况下，本研究中尽量选择多个种群，每个种群的个体数相对较少，这是因为已有研究表明具备足够空间代表性的种群数量更能提高景观因子与遗传学结合分析的准确性及对遗传多样性的可靠估计^[145, 146]。

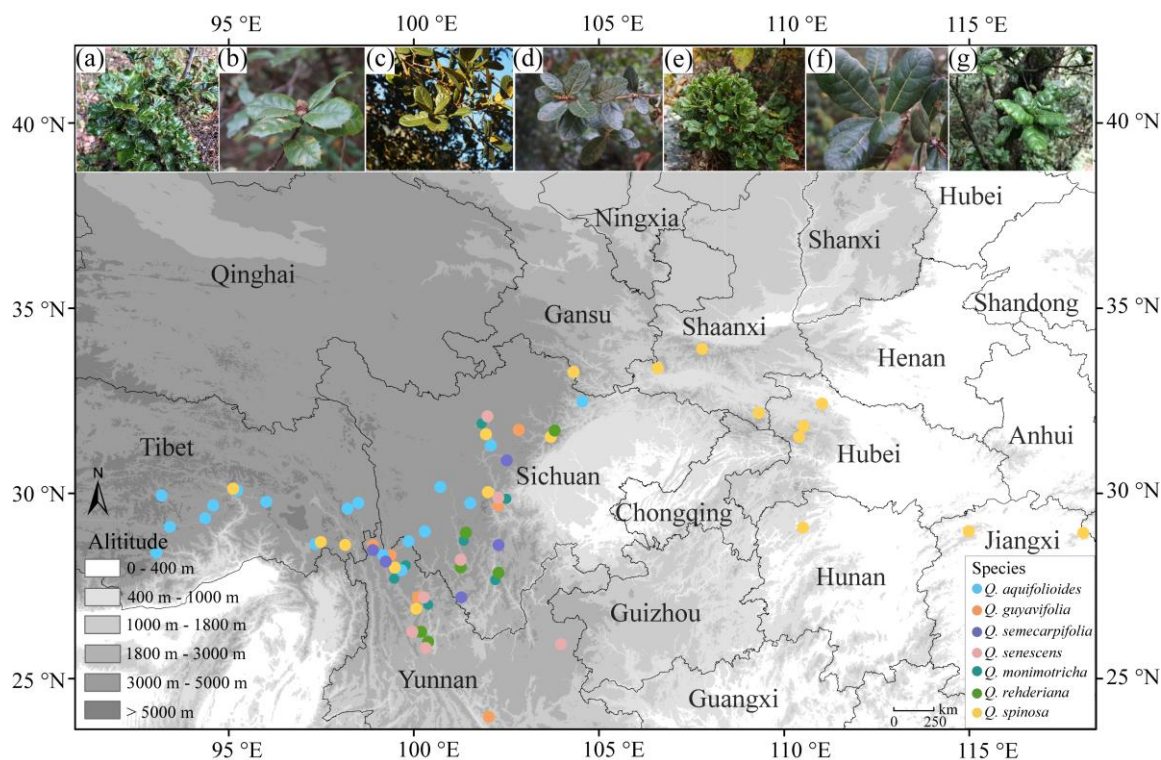


图 2.1 7 个高山栎组物种的采样点分布及叶片形态图。(a) 川滇高山栎；(b) 帽斗栎；(c) 高山栎；(d) 灰背栎；(e) 矮高山栎；(f) 毛脉高山栎；(g) 刺叶高山栎

Figure 2.1 Sample sites and leaf morphology in natural populations of seven species of *Quercus* sect. *Heterobalanus*. (a) *Quercus aquifolioides*, (b) *Q. guyavifolia*, (c) *Q. semecarpifolia*, (d) *Q. senescens*, (e) *Q. monimotricha*, (f) *Q. rehderiana*, (g) *Q. spinosa*

表 2.1 高山栎组植物采样信息表

Table 2.1 Sampling information of seven species of *Quercus* sect. *Heterobalanus*

物种 Species	种群 Code	经度 Longitude	纬度 Latitude	位置 Location
川滇高山栎 (<i>Q. aquifolioides</i>)	PW	104.53	32.46	四川县平武县
	BMS	95.21	30.06	西藏自治区波密县
	GBX	93.16	29.92	西藏自治区工布江达县
	LZA	94.34	29.30	西藏自治区米林市
	MLJ	93.38	29.07	西藏自治区米林市
	LZ	94.52	29.55	西藏自治区米林市
	LZD	93.02	28.37	西藏自治区隆子县
	BMR	95.99	29.75	西藏自治区波密县

物种 Species	种群 Code	经度 Longitude	纬度 Latitude	位置 Location
	CYX	97.32	28.61	西藏自治区察隅县
	MKD	98.18	29.57	西藏自治区芒康县
	MKR	98.47	29.72	西藏自治区芒康县
	BZL	99.15	28.31	云南省德钦县
	ZD	99.64	27.90	云南省香格里拉市
	XCR	99.84	28.68	四川省乡城县
	DCE	100.28	28.95	四川省稻城县
	YJH	100.69	30.15	四川省理塘县
	DMX	101.50	29.71	四川省康定县
	JCK	102.06	31.27	四川省金川县
帽斗栎 (<i>Q. guyavifolia</i>)	BMXS	99.15	28.31	云南省德钦县
	ELS	102.26	29.85	四川省泸定县
	MLXS	98.87	28.45	云南省德钦县
	MNP	100.25	27.17	云南省玉龙纳西族自治县
	MPS	102.00	23.94	云南省红河县
	MYL	102.81	31.66	四川省理县
高山栎 (<i>Q. semecarpifolia</i>)	BMXS	99.15	28.31	云南省德钦县
	FMC	102.26	28.58	四川省木里藏族自治县
	GSL	101.26	27.17	四川省盐边县
	MEX	102.48	30.87	四川省马尔康县
	MLXS	98.87	28.45	云南省德钦县
灰背栎 (<i>Q. senescens</i>)	ELS	102.26	29.85	四川省泸定县
	FMC	101.24	27.98	四川省木里藏族自治县
	JZS	100.36	25.97	云南省宾川县
	MNP	100.25	27.17	云南省玉龙纳西族自治县
	PMS	101.96	32.06	四川省康定县
	XYZ	100.17	26.24	云南省鹤庆县
	ZJY	103.95	25.91	云南省曲靖市
矮高山栎 (<i>Q. monimotricha</i>)	ELS	102.26	29.85	四川省泸定县
	GZX	99.78	28.06	云南省香格里拉市
	LS	102.26	27.83	四川省西昌市
	MNP	100.25	27.17	云南省玉龙纳西族自治县
	PMS	101.96	32.06	四川省康定县
	XZ	99.64	27.90	云南省香格里拉市
	ZR	101.38	28.93	四川省九龙县
毛脉高山栎 (<i>Q. rehderiana</i>)	ELS	102.26	29.85	四川省泸定县
	FMC	101.24	27.98	四川省木里藏族自治县
	JZS	100.36	25.97	云南省宾川县
	KKZ	103.78	31.67	四川省茂县
	LS	102.26	27.83	四川省西昌市
	XYZ	100.17	26.24	云南省鹤庆县
	ZR	101.38	28.92	四川省九龙县
刺叶高山栎 (<i>Q. spinosa</i>)	XGJQ	100.04	26.86	云南省玉龙纳西族自治县
	GZ	99.63	27.91	云南省香格里拉市
	ZJH	106.57	33.36	陕西省勉县
	GDS	101.98	30.01	四川省康定县

物种 Species	种群 Code	经度 Longitude	纬度 Latitude	位置 Location
	SWJ	101.92	31.57	四川省金川县
	LBZ	103.68	31.50	四川省汶川县
	JZG	104.29	33.25	四川省南坪县
	TM	95.10	30.10	西藏自治区波密县
	TJS	107.76	33.87	陕西省周至县
	BDX	109.29	32.15	陕西省平利县
	TDY	110.50	31.79	湖北省神农架林区
	DJP	110.38	31.50	湖北省神农架林区
	WDS	111.00	32.40	湖北省丹江口市
	TMS	110.48	29.05	湖南省永定区
	WKS	114.96	28.96	江西省武宁县
	SQS	118.06	28.91	江西省玉山县
	CYXC	97.46	28.66	西藏自治区察隅县
	BLLK	98.12	28.58	西藏自治区察隅县

2.2.2 重测序

(1) DNA 提取、文库构建与测序

使用改进的 CTAB 法从变色硅胶干燥的叶片中提取基因组 DNA^[147]。提取的 DNA 使用琼脂糖凝胶电泳法进行质量检测，合格后的 DNA 样品保存于该实验室-20°C 的冰箱。按照 BGISEQ-500 平台标准文库构建流程，制备插入长度为 150 bp 的双端测序文库，并对文库进行高通量测序，预期测序深度达到 30×。然后本研究使用由华大基因 (BGI) 开发的 SOAPnuke 1.5.4^[148] 软件对高通量测序获得的原始数据 (Raw data) (FASTQ 格式) 进行质量控制和预处理。首先删除包含接头的 reads；然后去除低质量碱基比例过高的 reads (如质量值 ≤ 5 的碱基占比超过 50%)；最后去除含未知碱基 (N) 比例超过 1% 的 reads，获得初步质量控制后的有效数据 (Clean data)。

(2) 数据比对和变异检测

使用 BWA-MEM 算法^[149]，参数设置为“-t 4 -M -R”，将所有有效数据比对至川滇高山栎参考基因组，在 Burrow-Wheeler aligner (BWA) v.0.7.15 软件^[150] 上进行序列索引和比对。比对生成 SAM 文件后，首先利用 SAMtools v.1.3.1 的 sort 命令将其转换为按基因组坐标排序的 BAM 文件，这是进行所有下游分析的基础。接着，为获得高质量的变异检测结果，必须对 BAM 文件进行严格预处理：其一，标记或移除 PCR 扩增引入的重复 reads，防止过度扩增带来的假阳性变异；其二，使用 samtools view -q 30 参数严格过滤，剔除比对质量值 (MAPQ) 低于 30 的低可信度比对结果。运用这两步处理显著提升数据的信噪比，确保变异检测准确性。然后应用 GATK MarkDuplicates 工具标记并移除重复及多重比对 reads，从根本上降低后续变异检测的假阳性率。随后本研究使用 GATK 的 UnifiedGenotyper 对群体样本进行系统性变

异检测，通过设置高置信度阈值等参数，生成一个整合样本变异信息的.vcf文件。

(3) SNP 数据过滤

本研究使用 GATK^[151]的 SelectVariants 工具对初步鉴定得到的.vcf文件中提取出 SNP 数据集，然后运用 VariantFiltration 工具对 SNP 数据集进行严格的质量过滤。过滤的参数设置为：`gatk VariantFiltration -R Reference Genome.fasta -V SNP.vcf.gz -O SNPfiltered.vcf.gz --filter-name "ExcessHetFilter" --filter-expression "ExcessHet > 13.0" --filter-name "MissingRateFilter" --filter-expression "AN < (10 * 2 * 0.9)" --genotype-filter-name "DPFilter" --genotype-filter-expression "DP < 10.0 || DP > 200.0" --genotype-filter-name "AF04Filter" --genotype-filter-expression "isHet = 1 & (AD[1] / (AD[0]+AD[1]+0.001)) < 0.4" --genotype-filter-name "AF06Filter" --genotype-filter-expression "isHet = 1 && (AD[1] / (AD[0]+AD[1]+0.001)) > 0.6" --genotype-filter-name "HomoRefFilter" --genotype-filter-expression "isHomRef = 1 & (AD[0] / (AD[0]+AD[1]+0.001)) < 0.99" --genotype-filter-name "HomoAltFilter" --genotype-filter-expression "isHomVar = 1 && (AD[1] / (AD[0]+AD[1]+0.001)) < 0.99" --genotype-filter-name "ZeroRead" --genotype-filter-expression "AD[0]+AD[1]=0"`。最后运用 PLINK2 命令对初步过滤的 SNP 数据集再次进行严格的质量控制，主要包括：①过滤掉次等位基因频率 (Minor allele frequency, MAF) 低于 0.05 的位点；②过滤掉在所有样本中基因型缺失率高于 5% 的位点；③进行连锁不平衡 (Linkage disequilibrium, LD) 从数据集中筛选出相互独立的变异位点，参数设置为：`--indep-pairwise 50 10 0.2`。

(4) SNP 变异的注释与统计

利用 SNPEff 软件^[152]对变异位点进行注释，首先使用 SNPEff 自带的命令行工具，根据参考基因组及基因组注释文件构建数据库。之后根据构建好的数据库进一步对需要注释的 VCF 文件进行变异位点注释分析，得到每个 SNP 的突变类型、变异位点所在的区域类型以及基因 ID 等信息，在 EXCEL 中绘制统计结果图。

2.2.3 分析方法

(1) 遗传结构与遗传多样性分析

运用 Admixture v.1.4.0^[153]对高山栎组植物的七个种进行群体遗传结构分析。该软件假设所有样本来自 K 个祖先群体，然后通过最大似然 (Maximum likelihood, ML) 估计祖先比例，运用最小交叉验证错误值 (Cross validation error, CV error) 进行确定最佳 K 值。本研究的 K 值范围设置为 1~15，每个 K 值重复 20 次，然后根据最小交叉验证 (CV) 误差选择最可能的 K 值，将对应的 .P 和 .Q 文件在 R 4.3.2^[154]中进行个体遗传比例可视化。

主成分分析 (Principal component analysis, PCA) 是群体遗传学中揭示群体结构的一种高效且直观的数据降维算法，以提取其中最主要变异模式。主要是将全基因组上数以万计的 SNPs 所包含的基因型数据，通过线性变换降维到少数几个能代表最主要遗传变异的主成分上。然后根据 SNPs 的变异程度，将个体聚类为不同群体，从

而评估种群结构。本研究利用 PLINK v1.90^[155]进行 PCA 分析, 生成 vec 和 eval 文件在 R 4.3.2^[154]中进行结果可视化。

运用 VCFtools v.0.1.15^[156]工具的--window-pi 选项计算高山栎组植物的 4 个谱系的核苷酸多态性 (Nucleotide diversity, π), 以 10 kb 为固定窗口大小、10 kb 为步长 (即采用非重叠连续窗口) 在全基因组范围内计算每个窗口内的平均核苷酸多样性值, 用以有效评估群体的遗传多样性。此外还利用 VCFtools v.0.1.15 软件计算了 4 个谱系的期望杂合度 (Expected heterozygosity, H_E)、观测杂合度 (Observed heterozygosity, H_O)、近交系数 (Inbreeding coefficient, F_{IS}) 和 Tajima's D 值。Tajima's D 值可以指示种群处于平衡选择还是定向选择、扩张或收缩状态。Tajima's D 值是一个经典的中性检验统计量, 当 Tajima's D 大于 0 时, 表明存在大量中等频率变异, 群体内稀有等位基因频率较低, 这可能是由种群收缩或平衡选择所造成的。而当 Tajima's D 小于 0 时, 表明存在大量低频率变异, 稀有等位基因频率增加, 可能是由于种群扩张或定向选择导致。此外本研究运用 Pixy v.0.95.0^[157]软件基于非滑动窗口 10 kb 大小计算相对分化 (F_{ST}) 和绝对分化 (d_{xy}) 指数评估遗传分化程度。

(2) 系统发育树构建

基于质控后的全基因组数据集, 采用 PLINK v1.9^[155]软件计算所有个体间基于状态同源 (Identity By State, IBS) 的遗传距离矩阵以评估个体间的亲缘关系。本研究采用 VCF2Dis v.1.53e^[158]软件进行系统发育树构建, 该工具可高效处理大规模基因组分型数据, 直接读取标准 VCF 文件, 计算任意两个样本在全基因组范围内有效 SNP 位点上的等位基因频率差异, 并将其平方后累加, 最终生成一个对称的个体间距离矩阵。因此, 本研究基于过滤后的高质量 SNP 数据集进行计算, 首先使用 bgzip 工具对 VCF 文件进行压缩处理, 然后执行命令为 VCF2Dis -InPut SNP.vcf.gz -OutPut all_dis.mat, 最终生成用于下游分析的成对距离矩阵文件 all_dis.mat。最后, 使用 iTOL v.6 (Interactive Tree Of Life)^[159]在线工具将生成的 Newick 格式树文件 (.nwk) 上传至 iTOL v.6 (Interactive Tree Of Life) 在线平台进行系统发育树的可视化与美化。

(3) 基因流

基因流在种群遗传分化中起着至关重要的作用。使用 TreeMix v1.12^[160]来推断高山栎组植物谱系间基因流的方向和强度, 通过从多个种群中获取等位基因频率来构建物种的最大似然树。然后, 通过种群在系统发育树上的关系提供估计值。最后, 根据估计值和实际值之间的大小来推断是否存在基因流动事件。TreeMix 运行时设置迁移边数 (Migration edges, M) 为 10, 使用 R 语言中的 *OptM* 包^[161]估计 TreeMix 结果的最优迁移边数, 从而估计高山栎组植物谱系间的最佳基因流动事件。

运用 DSuite v0.3^[162]进行了 ABBA-BABA 分析 (D-statistic) 评估谱系间的基因流情况。D 统计量是为 P1、P2、P3 和外群 (O) 的 4 个分类单元的固定系统发育而设计的。当 O、P1、P2 和 P3 之间没有基因导入时, D 统计量为 0; 当 P2 和 P3 之间

或 P1 和 P3 之间存在基因导入，则 D 统计量将为正或负。

(4) 连锁不平衡

基于 MaxDist 参数为 100 kb 运用 PopLDdecay V3.26^[163]软件计算不同谱系任意两个位点之间的相关系数 r^2 值，最后使用 PopLDdecay 软件提供的 Plot_MultiPop.pl 脚本绘制 LD 衰减曲线图。

(5) 种群动态历史分析

首先运用成对序列马尔可夫共祖模型 (pairwise sequentially Markovian coalescent, PSMC)^[164]估计各谱系在不同时期的有效种群大小 (Effective population size, N_e) 的变化。从每个谱系中选取测序深度大于 $30\times$ 的个体进行了有效种群大小的变化趋势运算。流程如下：(1) 利用 BWA-MEM 算法对选定个体的重测序数据分别比对至参考基因组生成 .bam 文件；(2) 将 .bam 文件生成二倍体一致序列，为 .fq 格式文件；(3) 运用 PSMC v.0.6.5-r67 工具包中的 fq2psmcfa 脚本将 .fq 文件转化为 PSMC 的输入文件的 .fa 格式；(4) 运行 PSMC 程序：/psmc-master/psmc -N25 -t15 -r5 -p "4+25*2+4+6" -o diploid.psmc diploid.psmcfa；(5) 运用 PSMC v.0.6.5-r67 工具包中的 psmc_plot.pl 脚本进行结果可视化处理。其中模拟动态种群历史参数设置的突变率设置为 $1e-8$ ，世代时间为 30 年。

运用序列马尔可夫共祖模型 (Sequential Markov coalescent, SMC++) 对高山栎组植物各谱系的有效种群大小 (N_e) 的历史变化进行分析，SMC++ 可以结合每个物种的多个个体进行位点频谱信息计算，能够提高有效种群大小的精确度^[165]。分析流程如下：(1) 建立循环脚本，运用 vcf2smc 脚本将含有每个物种的遗传数据文件 .vcf 文件按照染色体拆分转换成符合 SMC++ 输入文件的 .smc.gz 格式，命令为 for i in 1..12 do smc++ vcf2smc SNP.vcf.gz chr\$ i.smc.gz \$i S: Sample1,Sample2, ...,Sample* done；(2) 进行估算范围为 10^3 - 10^7 代之间的历史有效种群大小，smc++ estimate --spline cubic --knots 15 --timepoints 1000 1000000 --cores 20 -o Quercus 1e-8 *.smc.gz；(3) 将上一步的数值结果进行可视化处理，smc++ plot ./Quercus.pdf ./Quercus/model.final.json -g 30 --ylim 0 100000 -c。

(6) 环境数据收集及处理

从 WorldClim version 2.1 (<https://worldclim.org/data/worldclim21.html>) 下载空间分辨率为 30 角秒的 19 个生物气候变量、太阳辐射的月度数据、蒸散的月度数据和风速数据的月度数据^[166]，其中对太阳辐射、风速和蒸散的月度数据进行平均后获得每个地点的太阳辐射、风速和蒸散数据，合计共 22 个环境变量 (不含经度、纬度和海拔)。在 ArcGIS 10.8 中的 GCS_WGS_1984 下将环境变量重新采样为 1 公里的分辨率，然后提取每个地点的环境变量。为减少环境变量共线性问题造成的模型过度拟合，通过 usdm v.2.1.7 包^[167]使用方差膨胀因子分析 (Variance Inflation Factor, VIF)

将 VIF 值大于 5 的环境变量剔除。运用克鲁斯卡尔-沃利斯检验 (Kruskal-Wallis test) 对每个环境变量进行显著性检验。

(7) 地理隔离分析

本研究采用曼特尔检验 (Mantel test) 分析地理与环境因素在塑造高山栎组植物的空间遗传分化中的作用。遗传距离矩阵首先运用 PGDSpider_2.1.1.0^[168] 软件在 Linux 系统中将 .vcf 文件转换成 GENEPOP 格式的文件, 然后使用 genepop 程序计算成对种群间遗传分化矩阵 (F_{ST} 矩阵), 并进一步计算出成对种群间 $F_{ST}/(1 - F_{ST})$ 值, 最后将遗传距离数据进行标准化后转换为距离矩阵。地理距离使用“geosphere”包^[169] 计算基于采样点的经纬度坐标, 使用 Haversine 公式获取地理间隔距离。气候数据使用“ecodist”^[170] 包通过 Bray-Curtis 相异性指数转化为环境距离矩阵。所有统计分析在 R 4.3.2 版本中完成, 使用 vegan^[171] 包进行曼特尔检验, 置换的次数设置为 9999 次, 显著性水平设定为 $P = 0.05$ 。

(8) 生态位重叠

运用 Broennimann 等^[172] 提出在环境空间中进行生态位相似性测试 (Niche Similarity Tests) 分析探究不同遗传单元间的生态位分化与相似性模式。该方法使用重叠指数 (Schoener's D) 作为环境生态位重叠, 并包括一个统计框架, 以测试生态位相似性: Warren 等^[173] 认为当两个种群在环境空间中没有重叠时, Schoener's D 值为 0, 当两个种群共享相同的环境空间时, Schoener's D 值为 1。使用主成分分析的前两个轴创建一个多变量环境网格, 主要用 ade4^[174] 包的 dudi.pca 函数对标准化后的环境变量进行降维, 保留前两个主成分轴作为生态位空间的主要维度, 该主成分分析总结了先前选择的所有环境变量 (PCA-env)。随后主要为应用高斯核密度来估计每个区域的占有率 (Z_{ij}), 并且基于获得的不同占有率值来计算 Schoener's D 值。采用 ecospat^[175] 包的 grid.clim.dyn 函数进行生态位模型的构建, 以所有样本点在 PCA 空间中的分布作为全局环境背景, 然后针对每个遗传单元分别计算其在环境空间中的生态位占据情况。模型设置分辨率为 100×100 的网格, 通过核密度估计计算每个网格单元的生境适宜性概率, 生成平滑的生态位分布曲面。最后通过比较两个生态位分布曲面的相似性计算得出生态位重叠的 Schoener's D 指数, 该指数范围从 0 (无重叠) 到 1 (完全重叠), Schoener's D 值越高表明生态位重叠越大, 生态需求越相似。

为进一步检验生态位差异的统计学显著性, 本研究进行了生态位等价性检验和生态位相似性检验。生态位等价性检验用于比较两个遗传单元的生态位是否完全相同, 以及两个生态位空间是否可以互换。通过 ecospat.niche.equivalency.test 函数实现, 置换次数设为 1000 次。等价性检验用于评估生态位相似是否超出随机预期, 采用双向检验策略。将每对遗传单元的生态位重叠值 (D) 与 100 个重叠值的零分布进行比较。同时生态位相似性检验在考虑地理范围 (M 区) 周围环境条件的差异的条件下, 评估任何每对遗传单元的生态位是否不像偶然预期的那样相似, 其中 M 区为种群分

布的地理区域。在本研究中，为避免地理隔离和非适生区环境噪音对生态位分析的干扰，同时匹配高山栎组植物的较远的种子扩散能力，基于经纬度数据，构建梯度化缓冲区作为生态位分析的 M 区。首先合并所有谱系的空间点，采用 WGS84 (EPSG:4326) 坐标系构建最小凸多边形，代表高山栎组植物的实际联合分布区；随后基于测地线距离，对该凸包分别做 50 km 的缓冲区扩展，形成梯度化分析范围。这项分析使用 R 4.3.2 中的 Ecospat 程序包的 `ecospat.niche.similarity.test` 函数，置换次数 1000 次。若 $P < 0.05$ 则表明生态位相似性显著高于随机水平，可能暗示生态位趋同或受相同环境过滤影响。双向检验结果若双向均显著，表明生态位高度相似；若双向均不显著，表明强烈生态位分化；若仅单向显著，则表明非对称的生态位关系。该分析在 R 4.3.2^[154]中完成。

2.3 结果

2.3.1 重测序 SNP 位点统计

全基因组重测序的 109 个高山栎组植物样本共获得 4153.49 G 的原始数据 (Raw data)，经过质量控制、接头修剪、低质序列过滤等后获得 4035.98 G 的有效数据 (Clean data)。个体的平均测序深度为 32.27×，与参考基因组的平均比对率达到 97.96%，测序质量良好 (附表 1)。经过严格质量控制后，最终保留 309547 个高质量 SNPs 数据集。对 SNPs 在 12 条染色体上分布进行统计结果显示，SNPs 位点较均匀地分布在 12 条染色体上 (图 2.2)，未出现严重的局部富集和缺失的情况，能够合理地反映所有样品在各染色体区间的变异情况，为接下来的分析提供了可靠的基因组覆盖情况。

使用 SNPeff 软件对 309547 个高质量的 SNPs 位点进行功能注释和统计分析结果显示，位于基因间区的 SNPs 数量最多，达到 182970 个，占全部变异的 59%；其次是内含子区域 (Intronic) 的 SNPs 位点共有 49033 个，占全部变异的 16%；位于基因上下游 2000 bp 的 SNPs 位点，共有 44874 个，占全部变异的 14%；而位于基因编码序列外显子区域 (Exonic) 的 SNPs 数量有 30784 个，占基因组全部变异的 10%。此外，影响剪接位点 (Splicing) 的变异数量最少，共有 1885 个，占比 1% (图 2.3)。



图 2.2 12 条染色体上不同区域的 SNPs 密度图
Figure 2.2 The distribution of SNPs on 12 chromosomes

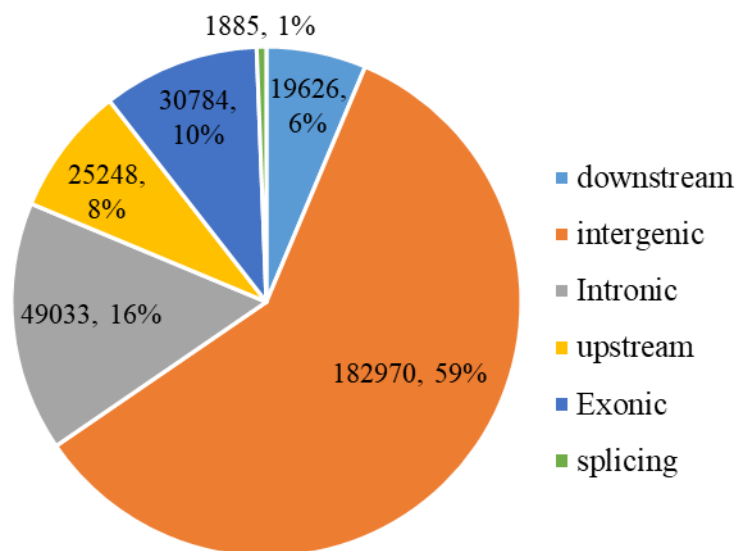


图 2.3 高质量 SNPs 位点在基因组上不同区域的数量及占全部位点的百分比
Figure 2.3 The number and percentage of SNPs in different regions of the genome

2.3.2 遗传多样性与遗传结构

运用 Admixture 分析对高山栎组植物的群体结构分析结果显示 $K = 4$ 时最小交叉验证错误值最小，为最佳的聚类值（图 2.4），表明高山栎组植物可以划分为四个谱

系。K = 2 时，高山栎组植物的谱系分化主要发生在青藏高原地区与秦岭至华东地区的个体之间（附图 1）；K = 3 时，青藏高原东南缘部地区的横断山脉地区的个体进一步分化为两个遗传谱系（附图 2）；在最佳 K = 4 时，高山栎组植物的七个种分为了四个不同的遗传谱系（A-D），谱系 A（lineage A）包括川滇高山栎、高山栎、帽斗栎、刺叶高山栎、灰背栎个体，主要位于青藏高原的西藏区域；谱系 B（lineage B）包括川滇高山栎、帽斗栎、矮高山栎、毛脉高山栎、灰背栎、刺叶高山栎个体，主要位于横断山脉北部地区；谱系 C（lineage C）包括刺叶高山栎、灰背栎、毛脉高山栎个体，主要位于横断山脉南部；谱系 D（lineage D）包括刺叶高山栎、毛脉高山栎的个体，主要位于秦岭至华东地区（图 2.5，附图 3）。

进一步基于全基因组 SNPs 评估高山栎组植物个体的亲缘关系和遗传结构，等位基因共享（IBS）的距离分析、邻接树（Neighbor-Joining tree, NJ）和主成分分析（PCA）的结果均显示位于秦岭至华东地区的个体与青藏高原地区个体的遗传距离较远，分别明显聚为一个支系（图 2.6，2.7）。总体来看，高山栎组植物的等位基因共享的距离分析、系统发育树和主成分分析与 Admixture 聚类的结果一致，高山栎组植物并没有按照传统形态物种水平进行聚类，而是呈现出一定的地理分布的形式聚类。因此，本研究认为高山栎组植物存在四个明显的遗传谱系，这一结论为后续相关研究提供了核心前提与分类依据。

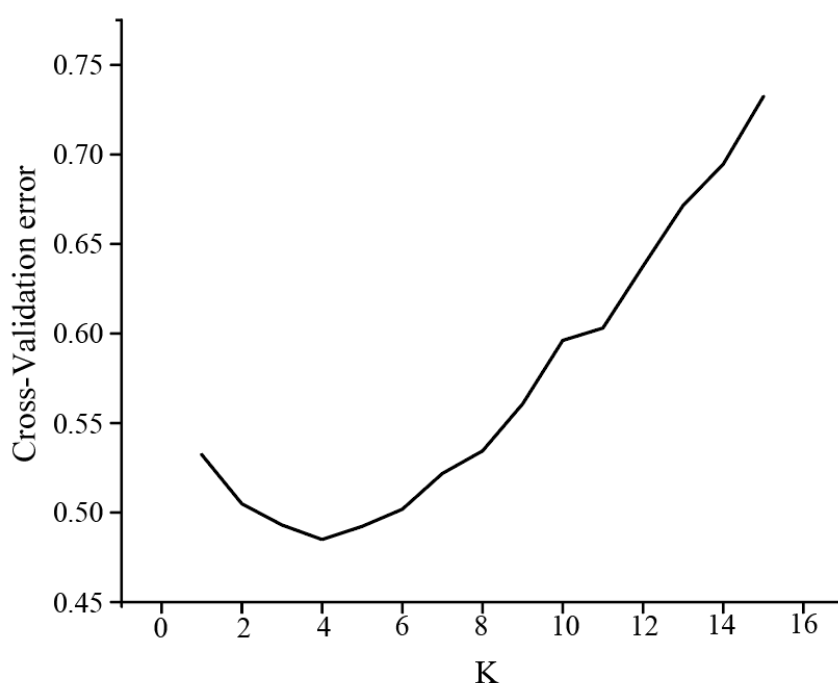


图 2.4 Admixture 分析取不同 K 值的交叉验证错误值

Figure 2.4 The maximum marginal likelihood values and cross-validation (CV) values for different K values are calculated using Admixture analyses

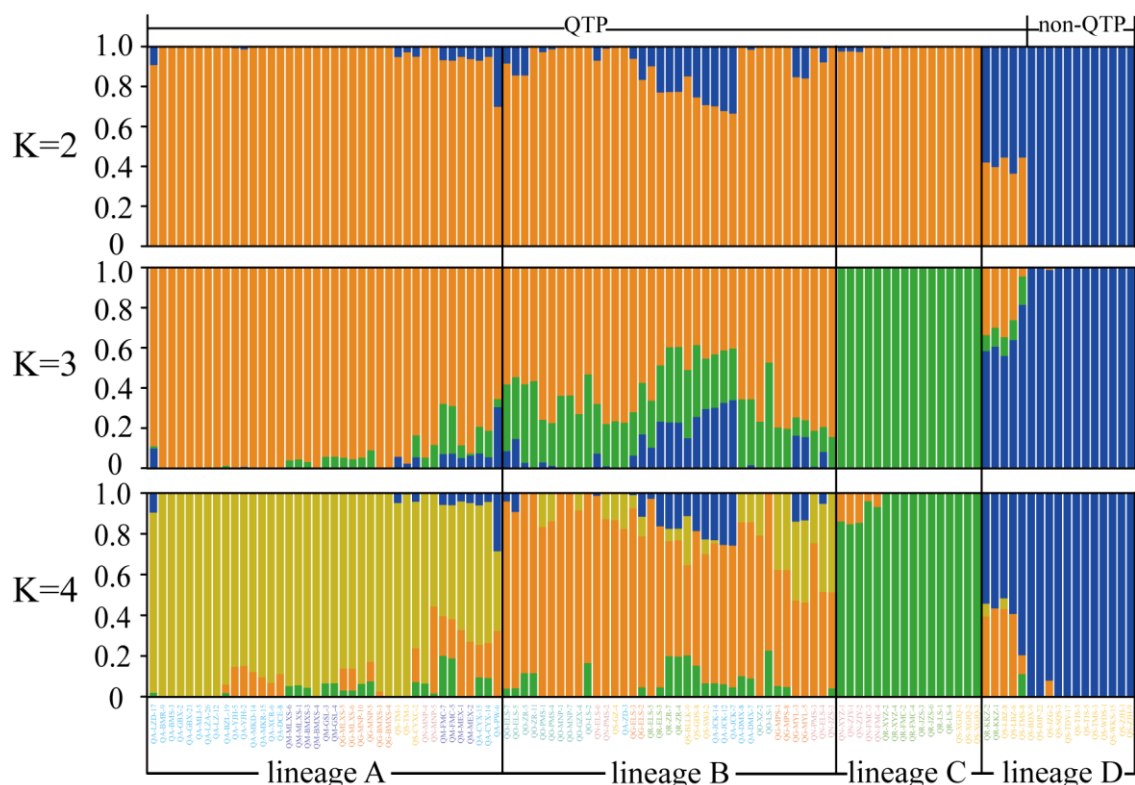


图 2.5 贝叶斯聚类分析高山栎组植物的遗传结构 (K = 2-4)。蓝色、橙色、紫色、粉色、蓝绿色、绿色和黄色的字体颜色分别代表川滇高山栎、帽斗栎、高山栎、灰背栎、矮高山栎、毛脉高山栎和刺叶高山栎

Figure 2.5 Genetic structure of *Quercus* sect. *Heterobalanus* using Bayesian clustering analysis (K = 2-4). The blue, orange, purple, pink, blue-green, green and yellow of the font represent *Q. aquifolioides*, *Q. guyavifolia*, *Q. semecarpifolia*, *Q. senescens*, *Q. monimotricha*, *Q. rehderiana* and *Q. spinosa*, respectively

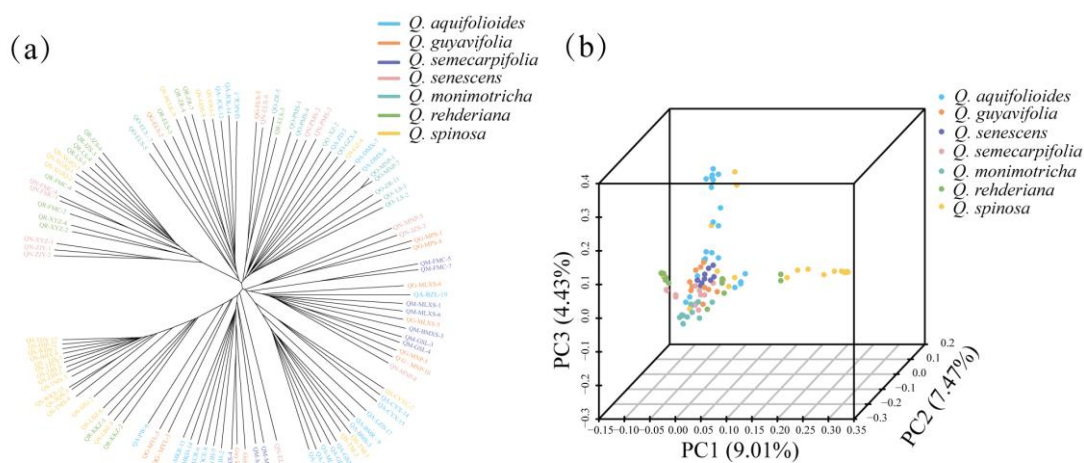


图 2.6 高山栎组植物的系统发育树 (a) 与主成分分析结果 (b)

Figure 2.6 The result of phylogenetic tree (a) and principal components analysis (b) of *Quercus* sect. *Heterobalanus*

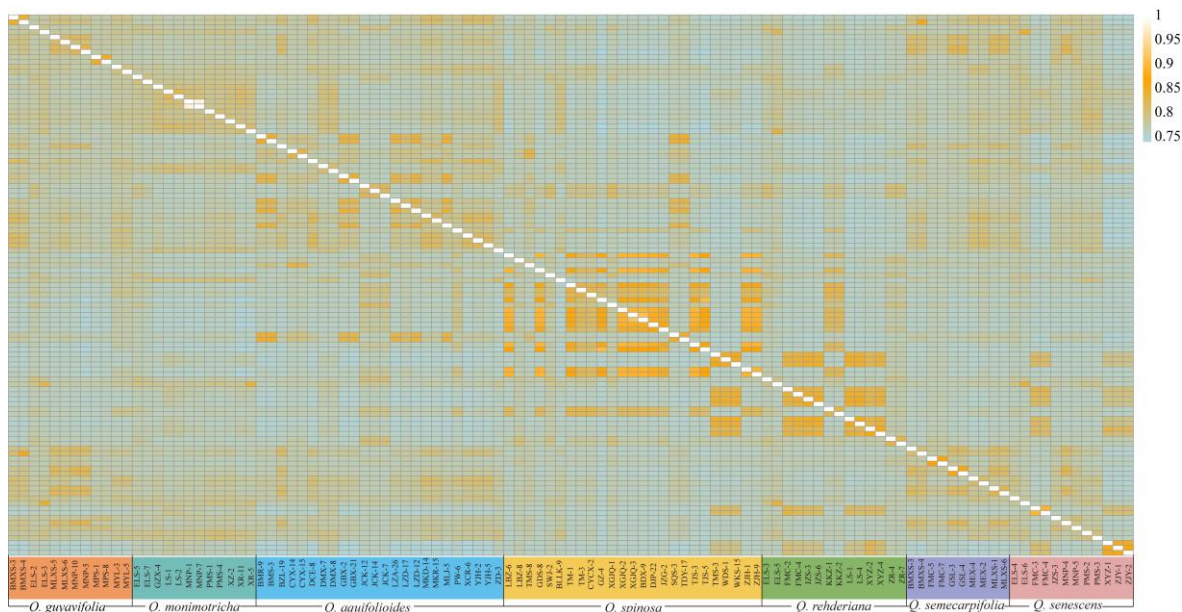


图 2.7 高山栎组植物的等位基因共享分析结果。蓝色、橙色、紫色、粉色、蓝绿色、绿色和黄色分别代表川滇高山栎、帽斗栎、高山栎、灰背栎、矮高山栎、毛脉高山栎和刺叶高山栎
 Figure 2.7 The result of Identity by State (IBS) of *Quercus* sect. *Heterobalanus*. The blue, orange, purple, pink, blue-green, green and yellow represent *Q. aquifolioides*, *Q. guyavifolia*, *Q. semecarpifolia*, *Q. senescens*, *Q. monimotricha*, *Q. rehderiana* and *Q. spinosa*, respectively

在谱系水平，谱系 C 的遗传多样性最高 ($H_E = 0.28$)，其次是谱系 B ($H_E = 0.26$) 和谱系 A ($H_E = 0.25$)，谱系 D 的遗传多样性最低 ($H_E = 0.24$) (附表 1)。核苷酸多样性的结果表明谱系 B 的核苷酸多样性最高 ($\pi = 1.04 \times 10^{-4}$)，谱系 D 的核苷酸多样性最小 ($\pi = 0.77 \times 10^{-4}$) (图 2.8-a)。Tajima's D 的结果显示谱系 A、B 和 C 大于 0，其中谱系 B 的值最大 (Tajima's D = 0.445)，其次为谱系 A (Tajima's D = 0.426) 与谱系 C (Tajima's D = 0.247)，谱系 D 的值最小并且趋近于负值 (Tajima's D = -0.004) (图 2.8-b)。连锁不平衡 (LD) 衰减分析结果表明谱系 D 呈现最慢的衰减速率，谱系 A 表现最快衰减速率 (图 2.8-c)。通过计算绝对分化指数和相对分化指数结果表明四个谱系间存在明显的遗传分化，其中谱系 C 与谱系 D 之间的遗传分化程度最大 ($F_{ST} = 0.2$, $d_{xy} = 0.11$)，谱系 A 与谱系 B 之间的遗传分化程度最小 ($F_{ST} = 0.04$, $d_{xy} = 0.107$) (图 2.8-d)。遗传分化指数 (F_{ST}) 表明不同谱系之间已经发生了相对显著分化差异。

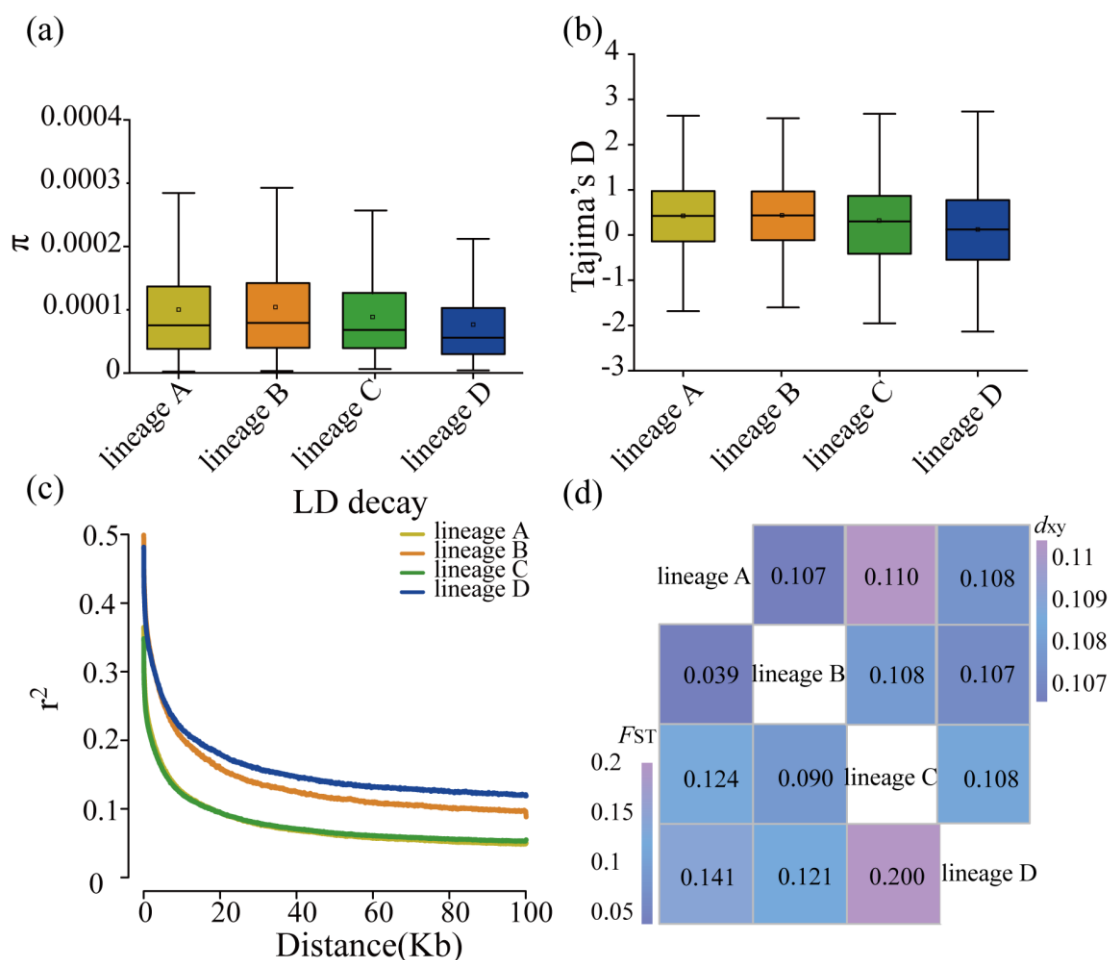


图 2.8 高山栎组植物四个遗传谱系的遗传参数。(a) 核苷酸多样性; (b) Tajima's D; (c) 连锁不平衡衰减情况; (d) 高山栎组植物四个谱系间两两比较的相对遗传分化 (F_{ST}) 和绝对分化 (d_{xy}) 差异的热图。蓝色表示遗传分化程度较低, 紫色表示遗传分化程度较高

Figure 2.8 Genetic parameters of different genetic lineages of *Quercus sect. Heterobalanus*. (a) Nucleotide diversity, (b)Tajima's D, (c) The decay of LD measured by r^2 as a function of physical distance. (d) The heatmap of relative (F_{ST}) and absolute (d_{xy}) divergence for pairwise comparisons among different lineages of *Quercus sect. Heterobalanus*, with blue and purple indicating lower and higher genetic differentiation, respectively

2.3.3 基因流与种群动态历史

通过 TreeMix 分析对高山栎组植物谱系间的基因流结果显示, *OptM* 将 $M = 3$ 作为最佳迁移模型, 解释方差比例达到 99.5%; 谱系间存在明显的不对称基因流, 谱系 A 与谱系 B 间检测到强烈的基因流, 方向为谱系 A 流向谱系 B; 还发现了谱系 D 流向谱系 B 的基因流 (图 2.9)。同时运用 D-suite 进行基因流分析结果显示, 在所有谱系组合中 D-statistic 均显著大于 0 ($D = 0.02 - 0.09$, $Z\text{-score} = 5.43 - 19.55$, $P < 0.001$), 表明不同遗传谱系之间存在不同程度的历史或持续的基因交流 (附表 2)。

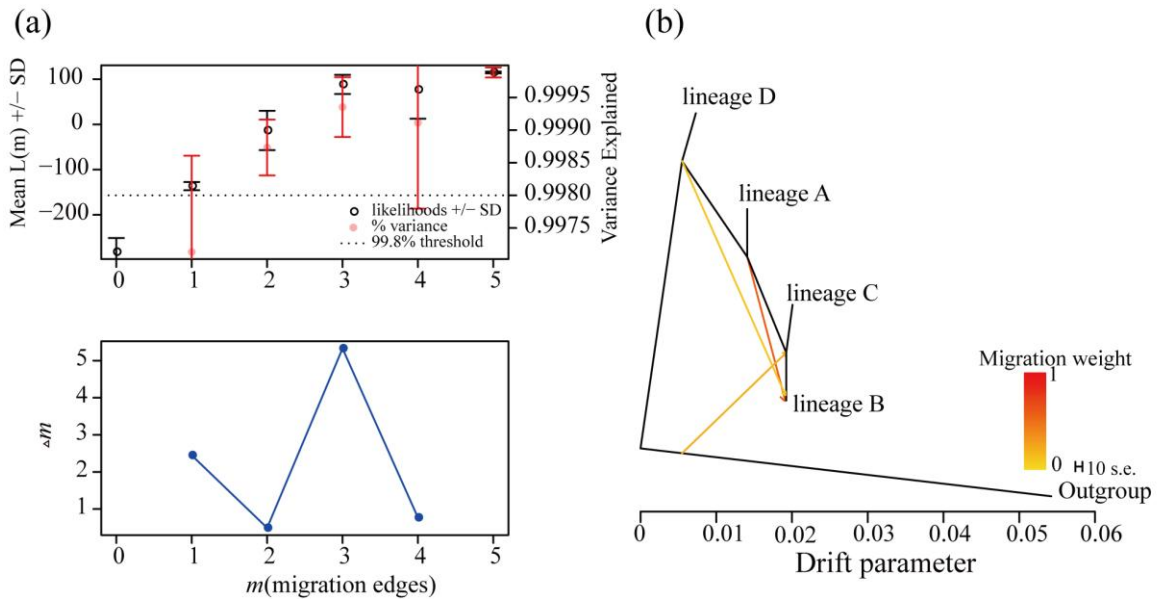


图 2.9 TreeMix 分析高山栎组植物不同谱系间基因流的结果。(a) 四个不同谱系之间的基因流动事件 ($M_{\text{best}} = 3$)。基因流动的方向由箭头指示。彩色箭头表示从 0 到 1 的迁移权重。(b) 评估 TreeMix 的不同 M 结果的最优拟合图，最佳拟合值 ($M = 3$)

Figure 2.9 Gene flow of different lineages of *Quercus* sect. *Heterobalanus* using TreeMix. (a) Gene flow events among different lineages ($M_{\text{best}} = 3$). The direction of gene flow is indicated by an arrow. The colored arrows show the migration weight from zero to one. (b) Optm fit plots, which evaluated different m results of the TreeMix, showing the best fit ($M = 3$)

利用成对序列马尔可夫共祖模型分析分别模拟高山栎组植物四个谱系的种群历史动态结果显示，四个谱系在早期历史阶段到一百五十万年前之间，经历了相似的动态历史事件，有效种群大小先不断扩张然后不断收缩的过程（图 2.10）。从一百五十万年前到一百一十万年前之间，谱系 A 和 C 的有效种群大小先趋于稳定，谱系 B 和 D 的有效种群大小出现缓慢扩张趋势（图 2.10）。从一百一十万年前到 10 万年前之间，四个谱系有效种群大小呈现由收缩到扩张的趋势（图 2.10）。

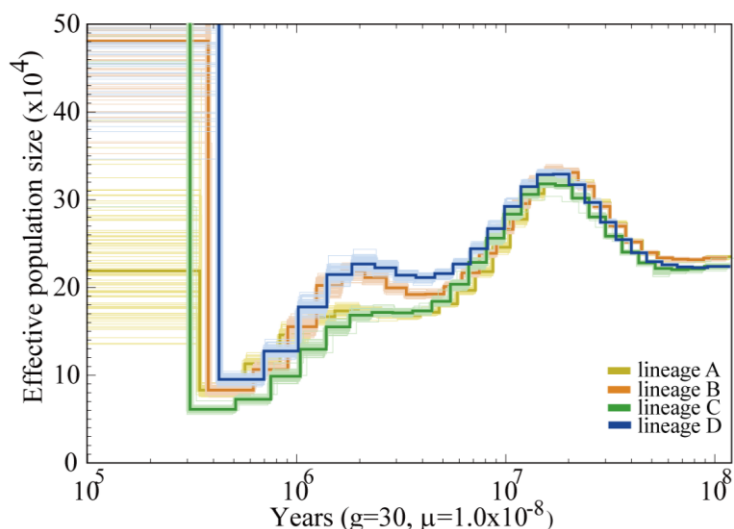


图 2.10 PSMC 模拟高山栎组植物不同谱系的有效种群大小变化

Figure 2.10 Demographic history inferred by the PSMC model among different lineages

本研究也结合序列马尔可夫共祖模型分析对四个谱系的动态历史进行模拟结果显示，一千万年前到十万年前之间，四个谱系经历反复的有效种群大小的收缩-扩张-收缩；大约在两万年前至三万年前，四个谱系均出现了有效种群大小扩张（图 2.11）。

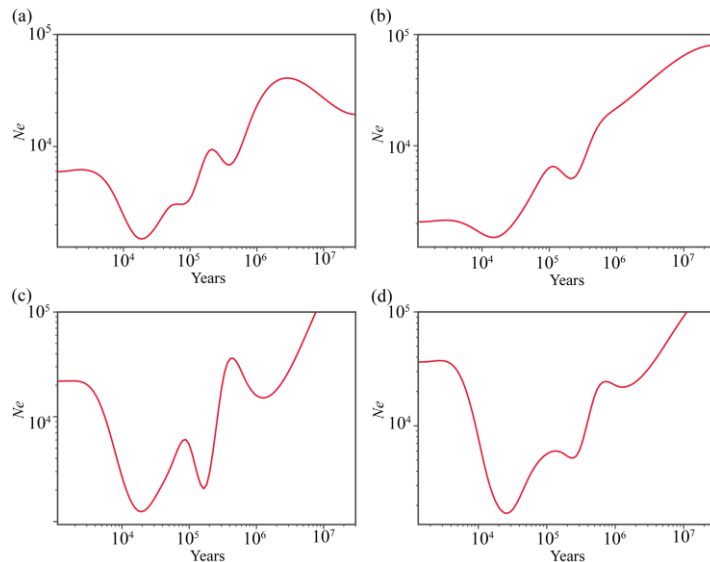


图 2.11 SMC++模拟高山栎组植物不同谱系的有效种群大小变化。(a) 谱系 A 的历史有效种群规模变化；(b) 谱系 B 的历史有效种群规模变化；(c) 谱系 C 的历史有效种群规模变化；(d) 谱系 D 的历史有效种群规模变化

Figure 2.11 The historical effective population size of four different lineages was inferred by SMC++. (a) Historical changes in effective population sizes of lineage A, (b) Historical changes in effective population sizes of lineage B, (c) Historical changes in effective population sizes of lineage C, (d) Historical changes in effective population sizes of lineage D

2.3.4 地理隔离分析

使用方差膨胀因子分析对 22 个环境因子进行去强共线性分析结果后保留了以下七个环境变量（附表 3），分别是代表温度的气候变量：平均气温日较差（Bio02），最冷月份最低温度差（Bio07），最干季度平均温度（Bio09）；代表降水的气候变量：最湿月份降水量（Bio13）和降水量季节性变化（Bio15）；及对高山植物具有较大影响的年度平均太阳辐射（Srad）和年度平均风速（Wind）（表 2.2）。对每个气候变量在谱系间的差异进行显著性检验结果显示，温度、降水和太阳辐射变量在四个谱系间均呈现显著性的差异，其中 Bio02, Bio07, Bio09, Bio15 和 Srad 呈现极显著的差异 ($P < 0.001$)，Bio13 呈现显著的差异 ($P < 0.04$)，Wind 无显著差异 ($P = 0.69$)（表 2.2，图 2.12）。

表 2.2 方差膨胀因子（VIF）筛选环境变量结果及显著性检验结果

Table 2.3 Results of variance inflation factor (VIF) analysis for screening environmental variables and their corresponding significance tests

气候变量 Climate variables	方差膨胀因子 VIF	卡方检验值 chi-squared	自由度 df	显著值 <i>P</i>
Bio02	2.53	34.73	3	1.39E-07
Bio07	2.70	37.38	3	3.82E-08
Bio09	4.62	31.92	3	5.45E-07
Bio13	2.32	8.56	3	0.03572
Bio15	2.10	21.90	3	6.85E-05
Srad	1.43	36.71	3	5.30E-08
Wind	2.67	1.48	3	0.6861

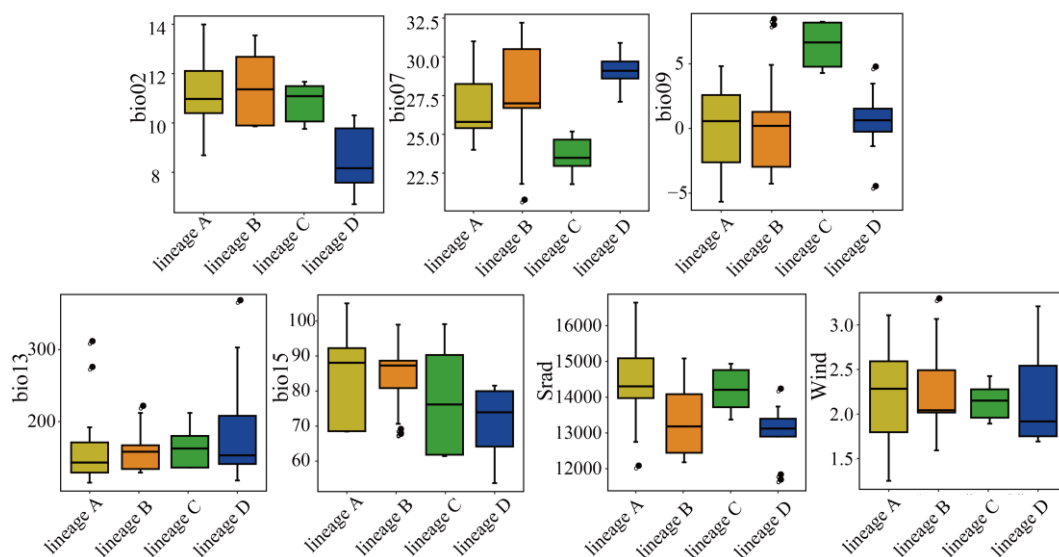


图 2.12 保留的七个气候因子进行评估不同谱系间的环境差异。y 轴代表环境因子数值

Figure 2.12 Environmental differences among lineage as measured by the seven retained climatic factors. The y-axis represents the climatic factor

通过对四个谱系气候变量的 PCA 分析结果显示，前两个主成分累计解释了谱系间环境差异的 54.1%，其中 PC1 和 PC2 分别解释了 34.6% 和 19.5%（图 2.13）。四个谱系在气候生态位中出现了一定程度的差异。

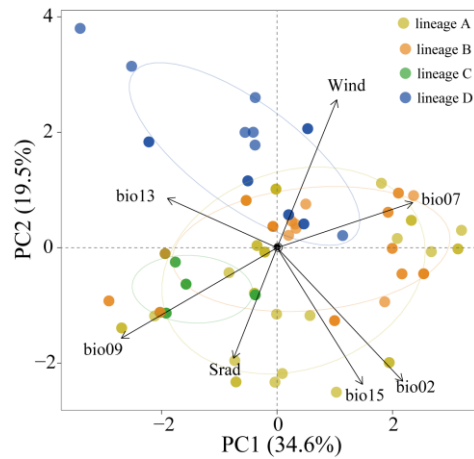


图 2.13 7 个环境因子的主成分分析图

Figure 2.13 First two axes of the principal component analysis of seven climatic variables

运用曼特尔检验（Mantel test）对高山栎组植物进行遗传分化与地理距离和环境距离的关系分析结果显示，地理距离和环境距离对高山栎组植物间的遗传距离均存在显著影响，遗传距离与地理距离存在显著正相关（Mantel $r = 0.39$, $P = 0.0001$ ），遗传距离与环境距离也存在显著正相关关系（Mantel $r = 0.18$, $P = 0.01$ ）（图 2.14）。

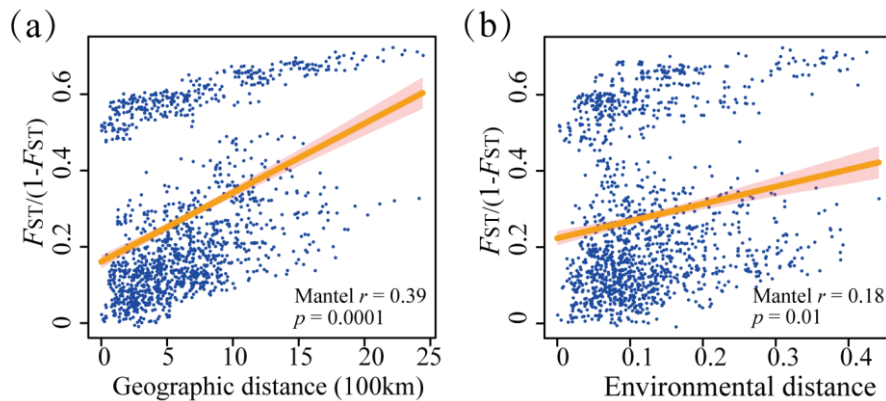


图 2.14 曼特尔检验高山栎植物的地理隔离与环境隔离效应。(a) 成对遗传距离 $F_{ST}/(1 - F_{ST})$ 与地理距离显著相关；(b) 与环境距离显著相关。采用 9999 次置换法进行显著性检验。线条展示模型预测结果，橙色阴影区域代表 95% 置信区间

Figure 2.14 Isolation by distance and environment using Mantel test. (a) Pairwise genetic distance $F_{ST}/(1 - F_{ST})$ is significantly associated with geographic distance and (b) environmental distance. The significant test using 9999 permutations. Lines show model the predictions, and the orange shading represents the 95% confidence intervals

2.3.5 生态位重叠

基于主成分分析-环境 (PCA-env) 方法, 对四个谱系进行成对生态位重叠分析, 并进一步进行了生态位等价性检验和相似性检验, 共进行了六组谱系对的比较分析 (图 2.15)。主成分分析-环境空间分布图结果显示不同谱系间的生态位重叠程度, 其重叠指数 (Schoener's D 值) 由高到低分别为: 谱系 A/B ($D = 0.149$)、谱系 B/D ($D = 0.089$)、谱系 A/D ($D = 0.008$)、谱系 B/C ($D = 0.005$)、谱系 A/C ($D = 0.003$), 以及谱系 C/D ($D = 0$)。总体上, 所有成对比较的 D 值均处于较低水平, 尤其是谱系 C 与 D 之间完全无重叠 ($D = 0$), 可以表明两个谱系在目前的环境维度上生态位高度分化, 基本不存在共享的环境空间。

生态位等价性检验结果表明, 除了谱系 B/D 之间表现出显著差异 ($P < 0.05$), 其余各谱系对之间均未达到显著水平 ($P > 0.05$), 说明多数谱系在环境空间中所占据的生态位在统计上具有等效性, 而 B 和 D 的生态位环境存在显著不同。生态位相似性检验的结果显示, 谱系 B/D、谱系 B/C、谱系 B/A 之间的生态位相似性均显著高于随机预期 ($P < 0.05$), 表明可能存在一定的生态位保守趋势, 而其余谱系对之间则未检测到显著差异 ($P > 0.05$), 可能表明它们之间的生态位差异对其谱系分化的贡献程度有限。综合重叠度指数与两类统计检验结果可以看出, 四个谱系整体呈现出生态位高度分化但多数谱系间生态位等价、仅部分谱系对存在显著生态位保守性的复杂格局。

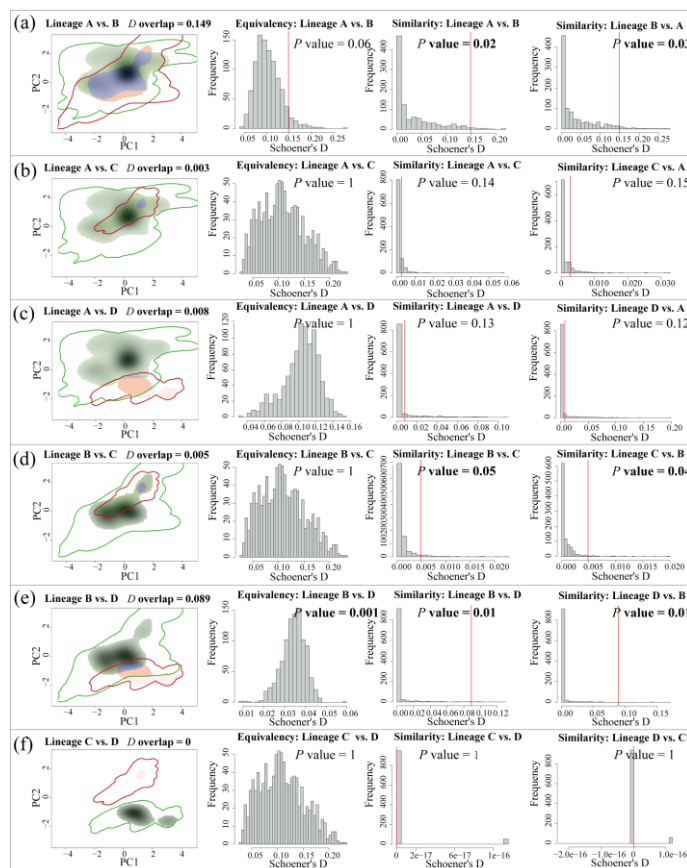


图 2.15 高山栎组植物四个谱系的相似性检验分析。谱系间的六组两两比较结果：(a) 谱系 A 与 B；(b) 谱系 A 与 C；(c) 谱系 A 与 D；(d) 谱系 B 与 C；(e) 谱系 B 与 D；(f) 谱系 C 与 D。左侧主成分分析-环境图 (PCA-env plot) 进行的相似性检验绘图。各图中阴影区域显示种群在单元格中的分布密度。实心等值线表示 100% 可利用环境 (背景)。绿色代表第一种群的生态位，红色代表第二种群的生态位。紫色阴影区域为核密度分布的生态位交集。直方图对应生态位分化等效性与相似性检验结果。直方图展示了两个分布区之间观测到的生态位重叠度 D (红线) 以及模拟的生态位重叠 (灰色条形)，生态位分化检验基于 100 次迭代计算得出。在生态位等效图中没有红线的均表示生态重叠度极低甚至为 0

Figure 2.15 Similarity test analysis for four lineages of *Quercus* sect. *Heterobalanus*. The six pairwise comparisons were tested between different lineages, (a) lineage A/B, (b) lineage A/C; (c) lineage A/D, (d) lineage B/C, (e) lineage B/D, (f) lineage C/D. The PCA-env plots (left) are shown according to the similarity test. Shaded areas in each plot show the density of the occurrences of the populations by cell. The solid contour lines illustrate 100% of the available (background) environment. The green colour represents the niche of the first population and the red colour the niche of the second population. The purple shaded areas are the niche intersections among kernel densities of occurrences. The histograms correspond to the results of equivalency and similarity tests to test for niche divergence. Histograms show the observed niche overlap D between the two ranges (red lines) and simulated niche overlaps (grey bars) on which tests of niche divergence are calculated from 100 iterations. The absence of red lines indicates extremely low ecological overlap or even zero overlap in equivalency tests

2.4 讨论

2.4.1 形态物种与遗传谱系的不一致性

尽管许多分类学家一直在努力探索物种的概念，许多物种仍处于分化的早期阶段，物种界限并不是非常清晰，许多的分类工作一直在不断地更新，特别是在物种丰富或快速辐射的谱系中^[176-178]。尤其是当同一物种在不同环境条件下可能表现出明显不同的形态特征，从而在传统形态分类中被误认为是不同物种，对于这类在形态相似、亲缘关系密切且物种界限模糊的类群即为物种复合体（Species complexes）^[176, 179]。此外，频繁的基因交流可能导致网状进化以及物种间的形态相似，从而进一步模糊物种之间的界限，增加物种或物种复合体间界定的难度，成为了运用传统形态分类进行物种界定与遗传单元界定的一个挑战^[180-182]。

基于全基因组重测序数据对高山栎组植物的七个种进行系统发育与群体遗传结构分析表明，基于传统形态分类划分的 7 个种并未形成对应的遗传聚类，而是在基因组水平上呈现出四个主要遗传谱系。这一结果与 Ju 等^[137]通过高山栎组植物的系统发育和遗传结构研究结果一致支持四个遗传谱系划分。在高通量测序技术的进步下，大量研究发现植物鉴定中观察到的形态差异并不一定对应真实的物种分化，因此基于形态特征划分的物种在遗传水平上界限不清晰的现象在植物界普遍存在^[183, 184]。尤其是在草本植物中，如毛茛科（Ranunculaceae）乌头属（*Aconitum* L.）^[185]、虎耳草属^[76]和禾本科（Poaceae）披碱草属（*Elymus* L.）^[186]等因世代周期短、辐射分化快以及易受环境驱动造成形态多样化但遗传分化较弱的近缘类群。Xu 等^[187]基于传统形态分类确定的 6 个禾本科物种使用形态学、遗传学和生态位分析的综合分析表明只能确定两个清晰的遗传谱系，并建议重新评估其物种划分。根据花序的被毛情况对青藏高原地区的乌头属进行划分的铁棒锤（*A. pendulum*）和伏毛铁棒锤（*A. flavum*）两个物种，通过广泛采样使用遗传数据和生态位数据分析发现两个物种的遗传差异不明显且生态位高度重叠，认为其并不能被视为两个独立的物种^[188]。在木本植物中，分子证据表明这种现象在松科（Pinaceae）松属（*Pinus* L.）^[189]、杨柳科（Salicaceae）杨属（*Populus* L.）^[190]和杜鹃花科（Ericaceae）杜鹃花属（*Rhododendron* L.）^[191]等类群中均有报道。

高山栎组植物的形态物种与遗传谱系之间的不一致可能由多种进化过程共同造成，如杂交、渐渗和不完全谱系筛选（Incomplete lineage sorting, ILS）等因素。基于谱系水平进行 TreeMix 分析结果检测到了谱系 A 向谱系 B 存在强烈的基因流信号，也检测到了微弱的谱系 D 流向谱系 B 的基因流。D-suite 分析结果表明所有谱系组合均检测到显著不为 0 的 D 统计量（Z-score > 3, $P < 0.001$ ），揭示了各谱系间存在普遍且显著的基因渐渗。此外，Ma 等^[136]对高山栎组植物的系统发育树拓扑结构研究发现核基因树和质体基因树存在广泛不一致性，表明在该物种分化过程中可能存在

种间基因流或不完全谱系筛选 (Incomplete lineage sorting, ILS) 现象。因此, 虽然本研究中高山栎组植物属于栎属冬青栎组的一个单支系, 但是由于具有异花授粉、花期重叠、频繁的杂交和渐渗等特点使得普遍同域分布的高山栎组植物的种群能有更多的机会进行频繁的基因交流, 共享祖先多态性, 从而造成了高山栎组植物的物种边界不清晰, 网状进化的现象^[192-194]。类似的现象在墨西哥特有类群墨西哥白栎复合体 (*Q. laeta complexes*) 中也有发现, 该物种复合体在其地理分布范围内表现出较大的形态变异, 传统分类基于叶形态差异进行分类的物种可达六个种, 但是近期对墨西哥白栎复合体进行系统发育关系研究表明该物种复合体呈现网状进化模式且可划分出 6 个地理遗传谱系^[195]。通常, 遗传结构可以按物种或地理分布的种群进行划分, 即使在生态位和地理分布重叠的种群中也是如此^[196]。本研究也发现高山栎植物的四个遗传谱系分布存在清晰的地理遗传结构。在 $K=2$ 时, 所有种群的遗传成分划分两个遗传谱系, 一个遗传谱系为分布于青藏高原地区 (QTP) 所有种群个体, 可能是由于该地区在地质历史和气候振荡时期其分布范围的变化导致了不同物种同域分布, 并由于栎属物种间存在不完全的生殖隔离, 它们之间的基因交流将会模糊物种界限^[193, 197]; 另一个遗传谱系为主要分布于秦巴山脉至华中地区 (non-QTP) 的所有个体, 可能是由于环境相似性导致了表型趋同现象, 或者在种群扩张过程中形成了固定相似的等位基因^[198]。 $K=3$ 时, 位于青藏高原地区的横断山脉地区的种群进一步产生了遗传分化, 可能是由于横断山地区在气候振荡期间形成多个冰期避难所, 不同的避难所间的长期隔离促进了进一步分化^[199]。当 $K=4$ 的最佳聚类值时, 四个遗传谱系分别主要位于西藏地区 (谱系 A)、横断山脉北部 (谱系 B)、横断山脉南部 (谱系 C) 和秦岭至华中地区 (谱系 D)。对青藏高原地区分布的物种而言, 相似的空间遗传结构形式在青藏高原地区分布的其他木本植物中均有发现, 如高山松 (*Pinus densata* Mast.)^[200]、岷江冷杉 (*Abies faxoniana* var. *faxoniana* (Rehder & E. H. Wilson) T. S. Liu)^[201]等。

除进化过程的影响因素外, 由环境差异或表型可塑性驱动的形态分化可能也是导致植物的形态物种分类与遗传谱系不一致的重要原因之一^[5, 184]。虽然本研究中缺乏对高山栎组植物的形态性状测量分析, 未能进一步结合几何形态学、表型性状等数据更深入解释造成谱系分化与形态分化不一致的机制, 但研究发现高山栎组植物会通过多种形态特征变化来适应不断变化的环境, 比如叶片绒毛、植株大小和叶缘锯齿变化等^[124]。如研究表明叶毛是喜马拉雅山脉东部的常绿栎属植物适应寒冷环境的重要性状之一, 该性状在防止叶片水分流失、防御低温胁迫以及提高抗寒、抗旱能力方面具有重要作用, 因此叶毛也是影响高山栎组植物分布格局的重要因素之一^[202]。与无毛叶片的物种相比, 如刺叶高山栎和毛脉高山栎, 具有明显叶毛的一些物种通常能够更好地耐受低温及高海拔环境条件。此外, 植株矮化也是高山植物适应高寒环境的一种重要生存策略。在高海拔地区, 较低的植株高度能够减少风害、降

低蒸腾作用等，从而提高植物在极端环境中的生存能力^[203]。因此，除了高山栎组植物中的矮高山栎，其他物种不仅能以乔木的生活型存在于高寒环境中，还能以灌木的生活型适应严酷的高山生态环境。这些均表明了高山栎组植物在剧烈变化的高山环境中，植物形态特征往往会对环境条件发生快速响应造成一定程度的形态差异而并不伴随发生明显的遗传分化，从而可能导致在形态上被划分为不同物种依旧属于同一遗传谱系。

2.4.2 谱系间的遗传分化与种群动态历史

遗传多样性作为生物多样性的重要组成部分，对物种的适应和进化潜力有着重要影响^[204, 205]。在青藏高原地区，由于经历了漫长且复杂的高原抬升和气候振荡的作用，遗传多样性呈现出多样化^[206]。尽管本研究中的部分地理种群样本量较少，然而对高山栎组植物划分的四个谱系进行遗传多样性和分化水平分析在一定程度上提高了样本代表性，并减少小样本种群对遗传多样性估计的影响，结果显示不同谱系间存在明显的遗传结构差异。其中，位于横断山脉地区的谱系 B 和谱系 C 有着较高的遗传多样性，这一结果表明横断山脉地区可能维持了高山栎组植物较高的遗传变异水平。横断山脉地区被认为是青藏高原东南缘生物多样性最为丰富的地区之一，越来越多的研究表明横断山脉地区可能是大量高山植物的冰期避难所^[199]，可能是高山栎组植物在历史演化过程中重要的遗传多样性中心或潜在的冰期避难所，从而有利于高山栎组植物遗传多样性的维持和积累。同时，高山栎组植物在横断山脉地区存在普遍的同域分布现象，促进了谱系间的基因交流，为遗传多样性的维持提供了重要的遗传变异来源。类似的在横断山脉地区表现出较高的遗传多样性的现象在许多青藏高原及其邻近地区的植物中也有发现，被认为与该地区在气候波动期间作为冰期避难所的作用有关^[200, 201, 207]。而分布于秦岭至华中地区的谱系 D 表现出最低的遗传多样性，与先前在青藏高原及其邻近地区的研究一致，表明在秦岭种群观察到的相对较低的遗传多样性可能是由于高水平的近交和有限的基因流造成的^[208]。

谱系间遗传分化指数 (F_{ST} 与 d_{xy}) 的结果表明，高山栎组植物不同谱系之间已经出现了较为明显的遗传分化，尤其是谱系 D 与位于青藏高原地区的其他谱系之间的遗传分化程度偏高。类似的遗传分化在一些高山植物如天蓝韭 (*Allium cyaneum* Regel)^[209]、华山松 (*Pinus armandii* Franch.)^[210]和掌叶大黄 (*Rheum palmatum* L.)^[211]等植物的秦岭种群与青藏高原种群之间也存在。在高山栎组植物中，谱系 C 与谱系 D 之间的遗传分化程度最为显著 ($F_{ST} = 0.20$, $d_{xy} = 0.11$)，表明分布于横断山脉南部的谱系 C 与分布于秦岭-华中地区的谱系 D 之间可能长期受到地理隔离的影响，从而限制了栎属植物的种子花粉传播，阻碍了基因交流，并促进了遗传分化。相比之下，主要分布于西藏地区的谱系 A 与位于横断山脉北部的谱系 B 相邻分布，其遗传分化程度较低。栎属植物普遍具有风媒授粉、花粉传播距离较远以及种间杂交渐渗频繁等特点，可以增强相邻分布种群间的基因交流，弱化谱系间的遗传分化程度

[192]。在其他栎属类群中也观察到类似现象，如欧洲白栎组的夏栎 (*Q. robur*) 和无梗花栎 (*Q. petraea*) 尽管在形态和生态上已经表现出明显差异，但是不同种群间依旧发生着较为频繁的基因交流^[212]。基因流分析结果显示存在明显的由谱系 A 向谱系 B 的基因流方向。此前，Jiang 等^[213]研究认为冬青栎组起源于东亚的青藏高原地区，随着地质历史的变化，向外扩散。Liu 等^[135]对川滇高山栎和刺叶高山栎进行历史基因流发现两个物种均具有从青藏高原地区由西边谱系向东边谱系的基因流方向，与本研究显著的基因流方向一致。因此，本研究进一步发现高山栎组植物的谱系分化在一定程度上符合“走出青藏高原” (Out of the Qinghai-Tibet Plateau, Out of QTP) 假说。类似的演化模式在多个高山植物类群中均有研究，例如杜鹃花属、报春花科 (Primulaceae) 报春花属 (*Primula* L.) 以及红景天属等植物均被认为在青藏高原或横断山脉地区起源，并随后向东亚其他山地地区扩散^[61, 214]。

Tajima's D 的结果表明不同谱系的种群历史可能存在差异。位于青藏高原地区的谱系 A、B 和 C 均为正值表明它们可能经历了种群收缩事件或平衡选择，这与青藏高原在更新世多次气候振荡背景下反复发生的栖息地破碎与二次接触过程高度吻合^[66, 70]。而接近零数值的谱系 D 表明其过去可能经历了轻微的扩张。已有研究表明，青藏高原在中新世至上新世期间经历了显著的地质抬升过程，并伴随着亚洲季风系统的加强，这些环境变化深刻影响了高山植物的分布和演化历史^[58, 61]。进一步的种群历史动态分析表明四个谱系的有效种群规模自中新世以来均经历了明显下降，可能是由于中新世是青藏高原的隆升最强烈的时期并伴随着气候干旱化，强烈的隆升导致青藏高原地区形成了高寒、低氧以及强紫外线的极端生境，导致高山栎组植物的适宜栖息地收缩，从而使得种群规模急剧下降。而在上新世时期，谱系 A 与谱系 C 的有效种群规模趋于稳定，而谱系 B 和谱系 D 则出现轻微扩张，这可能是由于青藏高原在此时期的隆升速率趋于平缓，气候环境和生态系统相对稳定^[58, 61]。

2.4.3 地理隔离与环境异质性驱动的遗传结构与生态位分化

地理隔离和环境异质性通常被认为是影响自然种群遗传结构的重要因素。本研究的曼特尔检验发现高山栎组植物遗传距离与地理距离存在显著正相关 (Mantel $r = 0.39$, $P = 0.0001$)，表明存在明显的地理隔离，基因交流受限。这一结果符合经典的地理隔离 (Isolation by distance, IBD) 理论，即地理隔离通过限制基因流促进种群间的遗传分化^[141]。特别是在高山植物中，由于地形复杂、栖息地破碎化以及物种扩散能力有限，地理隔离模式在许多植物类群中普遍存在^[215]。对于栎属植物而言，种群间的基因交流主要依赖于花粉和种子的传播。尽管风媒传粉可以在一定程度上实现远距离传播，而种子的扩散通常受到地形和生境条件的显著限制，从而导致种群之间的基因交流受到阻碍^[216]。高山栎组植物主要分布于青藏高原及其邻近地区，复杂的地理结构可能会显著降低花粉和种子的传播效率，从而加强地理隔离，阻碍了种群间的基因交流。青藏高原在新生代以来的持续隆升形成了复杂的山脉—峡谷地貌

格局，这种高度异质的地形结构常常将连续分布的栖息地分割为相互隔离的地理单元，从而限制种群之间的基因交流，并促进遗传分化^[131, 214]。类似的地理隔离驱动遗传结构的研究在青藏高原地区的高山植物中广泛存在。此外，本研究还检测到了显著的环境隔离（Isolation by environment, IBE）（Mantel $r = 0.18$, $P = 0.01$ ），表明环境因素可能也对高山栎组植物的遗传结构形成具有重要作用。环境隔离模式通常表明不同环境条件下的自然选择能够促进种群之间的适应性分化，即使在地理距离相对较近的情况下，不同生态环境中的种群也可能因局地适应而产生遗传差异^[217]。在青藏高原地区，由于海拔梯度显著、气候条件差异明显以及生境类型多样，环境异质性往往为植物种群提供了强烈的选择压力，从而促进适应性分化。在青藏高原分布的高山植物中，环境变化驱动的遗传分化在多种植物中已有研究，如麦吊云杉（*Picea brachytyla* (Franch.) E. Pritz.）^[218]。因此，在经历长时间的地理隔离和环境选择压力下，高山栎组植物的不同种群在适应当地生态环境变化过程中，可能发生了基因组分化，从而形成了当前的遗传结构格局。

在地理隔离和环境因素的共同作用下，种群可能进一步通过生态位分化实现对局地环境的适应。越来越多研究通过生态位重叠分析揭示了生态位分化是影响物种谱系分化的一个重要因素^[219]。不同谱系或种群能够通过生态位分化实现对局地环境的适应，进而降低种间或谱系间的竞争排斥，进一步强化对异质环境的利用与占据，促进适应性分化。本研究基于 PCA-env 方法比较了高山栎组植物不同谱系在环境空间中的生态位重叠程度，并结合生态位等价性与相似性检验，对谱系间生态位分化格局进行分析结果显示，所有谱系对之间的重叠指数（Schoener's D）值均处于较低水平，表明高山栎组各谱系在环境空间中的生态位重叠有限，具有明显的生态位分化趋势，各谱系有自己的潜在环境空间。这一结果暗示不同谱系在长期演化过程中可能经历了不同的环境选择压力，从而在气候或地理上形成了相对独立的生态位。生态位等价性检验结果显示只有谱系 B 与 D 之间拒绝了生态位等价性的原假设，表明这两个谱系在环境空间中的分布差异显著。这一差异可能主要受到地理位置和局地环境条件差异的影响：谱系 B 分布在横断山脉北部，海拔梯度大、季风影响显著、局地微气候复杂，而谱系 D 主要分布在秦岭至华中地区，环境条件相对温和稳定。长期环境选择压力可能促使两谱系形成不同的生态适应策略，从而在生态位上发生显著分化。此外，生态位差异可能有助于减少潜在的竞争排斥，维持谱系在接触区的共存，并可能促进谱系间的遗传隔离，这一机制与高山植物谱系分化的其他研究结果一致^[220, 221]。其余 5 个谱系对比较的生态位并非严格等效，表明谱系 A 和 C 很可能已经适应了当地的环境。这一模式与先前研究一致，表明亲缘关系更近的谱系在局地环境下并不必然保持严格等效的生态位，而生态位分化往往比随机期望更为显著^[172, 220]。生态位相似性检验进一步揭示了谱系 B 在多个比较中均表现出显著的生态位偏离，其生态位分布明显不同于谱系 A、C 和 D，表明谱系 B 为了适应横断

山脉北部地区的地理环境变化形成了相对独特的生态适应策略。生态位的差异能够反应不同谱系在生态适应能力以及各自的地理分布范围上的差异^[173]。因此，气候变化在促进高山栎组植物的适应性分化方面发挥了重要作用。

2.5 小结

本章节对高山栎组植物 109 个体的全基因组重测序数据进行群体基因组学分析。结合系统发育和群体遗传结构分析结果显示，高山栎组植物并未按照传统形态物种界限进行聚类，而是呈现出明显的地理结构，并形成四个遗传谱系，呈现了形态分类单元与遗传单元之间存在不完全一致的情况。位于横断山脉地区的谱系 B 和 C 有较高的遗传多样性，而分布于秦岭至华东地区的谱系 D 的遗传多样性最低。此外，谱系之间存在显著的遗传分化，其中谱系 C 与 D 之间的相对分化程度最高，而谱系 A 与 B 之间分化程度最低，且存在明显的谱系 A 向 B 的基因流。种群历史动态分析进一步表明，四个谱系在中新世早期至第四纪期间经历了反复的种群扩张与收缩事件，增加了二次接触的机会。曼特尔检验结果表明遗传距离与地理距离之间的相关性强于与环境距离之间的相关性，地理隔离对遗传结构的形成产生了更重要的作用；同时，不同遗传谱系在气候变量中存在显著差异，生态位重叠分析结果显示谱系间的重叠度偏低甚至无重叠，表明各谱系出现了生态偏好现象。因此，地理隔离可能是驱动高山栎组植物形成现今遗传结构的主要因素，而气候条件差异可能是促进谱系间产生适应性分化的重要因素。

3 高山栎组植物谱系间自然选择信号

3.1 引言

植物的进化过程通常受地质变迁和气候波动等多种因素的共同影响^[73, 138]。环境异质性通常被认为是影响栎属植物种内或种间遗传分化与局地适应的重要因素^[222, 223]。在复杂山地生态系统中, 不同谱系或种群往往会在空间隔离与环境选择的共同作用下形成显著的基因组分化格局^[224, 225]。基因组水平的分化模式不仅反映历史进化过程与基因流动态, 也反映了自然选择在局地适应与谱系形成中的作用^[32, 226]。随着高通量测序技术的发展, 越来越多的研究表明基因组分化通常呈现高度不均一的空间分布特点, 即部分区域表现出显著偏高的遗传分化水平, 而其他区域则保持较低的遗传分化程度^[224, 227]。这种空间异质性分化格局可能是由于多种进化过程的共同作用导致的, 包括选择性清除、背景选择、重组率差异以及持续或间歇性的基因流^[228, 229]。对于存在基因流的遗传分化过程中, 受到强烈选择的区域可能形成相对独立的高分化区段, 从而形成“基因组分化岛”, 这些区域往往会富集与局部适应或者生殖隔离相关的功能基因^[229]。

对分布于青藏高原及其邻近复杂山地系统的高山栎组植物而言, 长期的地质气候变化及高度的环境异质性的共同作用促使不同谱系承受了不同的生态选择压力, 并形成不同的基因流格局^[61, 70]。复杂的山地地形所形成的地理隔离可能限制了谱系间的基因交流, 而沿海拔梯度或气候梯度分布的环境差异则可能驱动谱系在不同生态环境中的适应性分化^[32, 230]。因此, 从基因组水平评估谱系间分化强度与选择信号的分布格局, 有助于进一步理解地理隔离与生态选择压力在谱系形成与维持中的相对作用。

本研究基于谱系水平运用群体遗传分化指数 (F_{ST}) 和跨群体复合似然比 (Cross-Population composite likelihood ratio, XP-CLR) 两种统计方法进行选择性清除分析, 以识别潜在受选择的基因组区域及候选基因, 进一步通过功能注释与 GO 富集分析解析受选择基因参与的主要生物学过程。 F_{ST} 通过计算群体间等位基因频率的差异检测选择信号^[138]。跨群体复合似然比通过整合连续基因组区域的多个 SNP 信息, 利用“复合似然比”模型增强对选择信号的检测能力, 有利于捕捉因正选择导致的等位基因频率偏离和连锁不平衡 (LD) 变化^[231]。此外, 栎属植物虽分布广泛, 生态适应范围广泛, 但世代周期长、遗传背景复杂、种间杂交与基因渐渗普遍, 且难以在树木中进行基因功能验证, 极大限制了对于栎属植物的受选择基因功能的研究发展。本研究将候选基因在拟南芥数据库中进行拟南芥同源基因比对, 并推测候选基因的功能类别, 从而揭示高山栎组植物在复杂的山地环境中适应性进化的遗传基础, 为理解栎属植物的生态分化机制及抗逆性状的分子调控提供理论依据。

3.2 分析方法

3.2.1 基因组水平分化

运用 VCFtools v.0.1.15^[156]评估高山栎组植物的四个遗传谱系的染色体水平的核苷酸多态性和 Tajima's D 值, 还运用 Pixy v.0.95.0^[157]评估四个遗传谱系的相对分化指数 (F_{ST})。

3.2.2 基因组分化岛检测

选择性清除 (Selective sweep) 能帮助识别与适应性变异相关的基因组信息以及强烈群体分化的异常基因组区域。本研究仅选取遗传成分纯合的个体 (Q 值 > 0.9 或 < 0.1) 进行比较, 以减少近期杂交和渐渗对选择信号的干扰, 在一定程度上降低混合个体所带来的群体结构偏差, 更能反映谱系间相对稳定的分化区域。运用群体遗传分化指数方法和群体特异性特征扫描方法进行选择扫描, 所有的统计数据均采用非重叠的 10 kb 滑动窗口。对于群体遗传分化指数的方法, 运用相对遗传分化 (F_{ST}) 进行高分化区间的鉴别, 采用前 5% 的窗口作为高分化区域 (Highly diverged regions, HDRs)。然后对高分化 (HDRs) 与非高分化区域 (non-HDRs) 进行统计分析, 采用 ggplot 进行箱线图可视化, ggsignif 用于添加显著性标记, 显著性检验采用 t 检验 (t-test) 完成。对于群体特异性特征的扫描方法, 本研究使用相同滑动窗口大小进行跨群体复合似然比 (XP-CLR) 分析。同样采用前 5% 的窗口作为选择信号区域。

随后, 将群体遗传分化指数的 F_{ST} 鉴定出的高分化区域和群体特异性特征的选择信号进行两组之间取交集, 将相同结果作为正选择高分化区域, 而正选择高分化区域内的基因定义为候选基因, 交集运用维恩图通过 jvenn 在线网站 (<http://bioinfo.genotoul.fr/jvenn>) 进行计算与可视化。

3.2.3 富集分析

将候选基因的氨基酸序列通过 BLASTP 程序比对至拟南芥 (*Arabidopsis thaliana*) 基因组注释数据库 (TAIR, <https://www.arabidopsis.org/>) 进行同源基因搜索, 设置 E-value 小于 $1e-05$ 以获取显著性匹配的同源基因标识符, 通常选取最高 BLAST 匹配结果。随后, 将这些拟南芥同源基因标识符作为查询列表, 并以拟南芥全基因组基因作为背景, 提交至 PANTHER 数据库 (<https://www.pantherdb.org/>) 进行基因本体 (GO) 富集分析^[232], 推断其生物过程与信号通路。富集显著性利用费歇尔精确检验 (Fisher's exact tests) 进行计算, 并使用错误发现率 (False discovery rate, FDR) 方法对结果进行校正。由于候选基因集规模较小, 经过严格的 FDR 校正 ($FDR < 0.05$) 后不能发现显著富集的 GO 术语。因此, 本研究参考的探索性分析方法^[233, 234], 汇报原始 P 值小于 0.05 且映射基因数量不少于 5 个的前 20 个 GO 术语, 作为潜在的功能富集趋势。

3.2.4 顺式调控元件分析

对于通过 F_{ST} 和跨群体复合似然比分析共同鉴定出的正选择高分化区域，本研究首先提取所有位于基因上游（upstream）和下游（downstream）区域的 SNP 位点。对每个位于基因上下游的 SNP，从其 5' 和 3' 两端各提取 10 bp 的短序列片段，构建包含 SNP 位点的 21 bp 基础序列，用于后续将提取的这些序列与植物顺式调控元件数据库 NewPLACE (<http://www.dna.affrc.go.jp/PLACE/>) 进行相似性搜索比对，以寻找已知的调控元件序列^[235]。然后对返回结果进行人工筛选与整理，保留包含基因 ID、元件位置及功能注释的有效信息。最后以影响植物适应性的两大因素为分类进行每个谱系的顺式调控元件统计，具体包括两类：一类是非生物环境因子，包括温度、水分和光照条件；另一类是生物类因子，包括花粉、种子及叶片。

3.3 结果

3.3.1 基因组分化水平

高山栎组植物四个谱系间的全基因组水平核苷酸多样性结果显示谱系 A 具有最高的核苷酸多样性，谱系 D 的核苷酸多样性最低，四个谱系在每条染色体上的变化趋势几乎相同，除了谱系 C 和 D 在 9 号染色体上出现连续降低的趋势，其中在 12 号染色体上出现最大的差异 ($\pi_B - \pi_D = 3.8 \times 10^{-5}$) (图 3.2-a)。谱系 A 在染色体水平上的 Tajima's D 范围为 0.36 (2 号染色体) -0.49 (12 号染色体)，谱系 B 在染色体水平上的 Tajima's D 范围为 0.38 (2 号染色体) -0.51 (12 号染色体)，谱系 C 在染色体水平上的 Tajima's D 范围为 0.17 (4 号染色体) -0.32 (10 号染色体)，谱系 D 在染色体水平上的 Tajima's D 范围为 -0.11 (10 号染色体) -0.1 (2 号染色体) (图 3.2-b)。图中显示谱系 A、谱系 B 和谱系 C 在所有染色体上的 Tajima's D 值都为正值，表明这三个谱系可能经历了平衡选择或群体收缩事件，但是谱系 D 在 1、9、10、11、12 号染色体上的 Tajima's D 值为负值，并且 Tajima's D 平均值趋近负值，表明谱系 D 经历了定向选择或者近期的群体扩张事件。从群体遗传分化指数 (F_{ST}) 程度来看，谱系 C 与 D 之间的遗传分化程度最高，谱系 A 与 B 之间的遗传分化程度最低，所有的谱系对在 1、2、3 号染色体上的遗传分化有较其他染色体高的 F_{ST} 值 (图 3.2-c)，表明在这些染色体上更可能发生选择性清除作用。

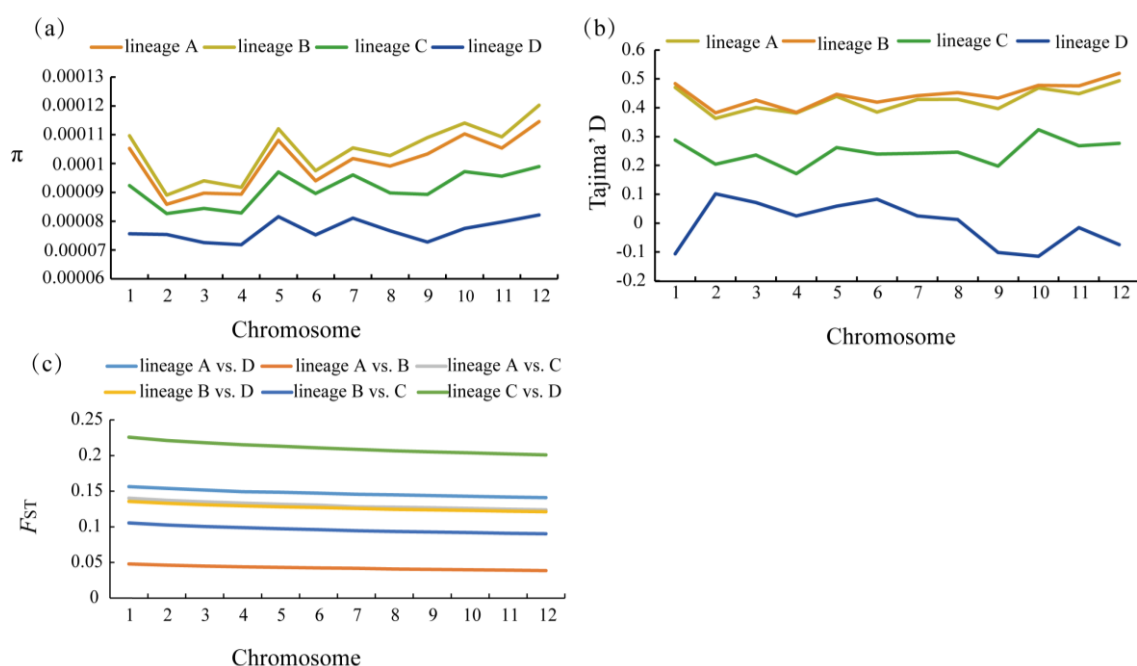


图 3.1 高山栎组植物各谱系每条染色体的核苷酸多态性和 Tajima's D 与谱系间的遗传分化指数
 Figure 3.1 Nucleotide diversity (a) and Tajima's D (b) of each chromosome in different lineages and genetic differentiation between lineages (c) of *Quercus* sect. *Heterobalanus*

3.3.2 高分化区域

为了确定遗传分化的平行性和非平行性，使用滑动窗口计算谱系对之间的群体遗传分化指数 (F_{ST})，并将 top 5% 的窗口识别为高分化区域 (HDRs) 结果表明，谱系 A 与 D 间的 F_{ST} 检测到 3660 个高分化区域，谱系 A 与 B 间的 F_{ST} 检测到 3676 个高分化区域，谱系 A 与 C 间的 F_{ST} 检测到 3670 个高分化区域，谱系 B 与 D 间的 F_{ST} 检测到 3649 个高分化区域，谱系 B 与 C 间的 F_{ST} 检测到 3654 个高分化区域，谱系 C 与 D 间的 F_{ST} 检测到 3576 个高分化区域。通过对高分化与非高分化区域进行统计分析结果显示高分化区域中 6 个谱系对之间的 F_{ST} 值明显高于非高分化区域，表明这些高分化区域可能驱动了谱系间的分化 (图 3.2)。

通过比较不同谱系对之间的高分化区域的重叠部分，发现在 470 - 1599 个共享高分化区域 (图 3.3)。这一结果表明，高分化区域只有一部分在其他遗传谱系间也存在分化，而基因组的大部分区域呈非平行进化，说明高山栎组植物在适应相似环境、形成相似表型的过程中，并非依赖相同的基因组区域/基因通路，而是通过不同的遗传变异路径实现适应性进化。

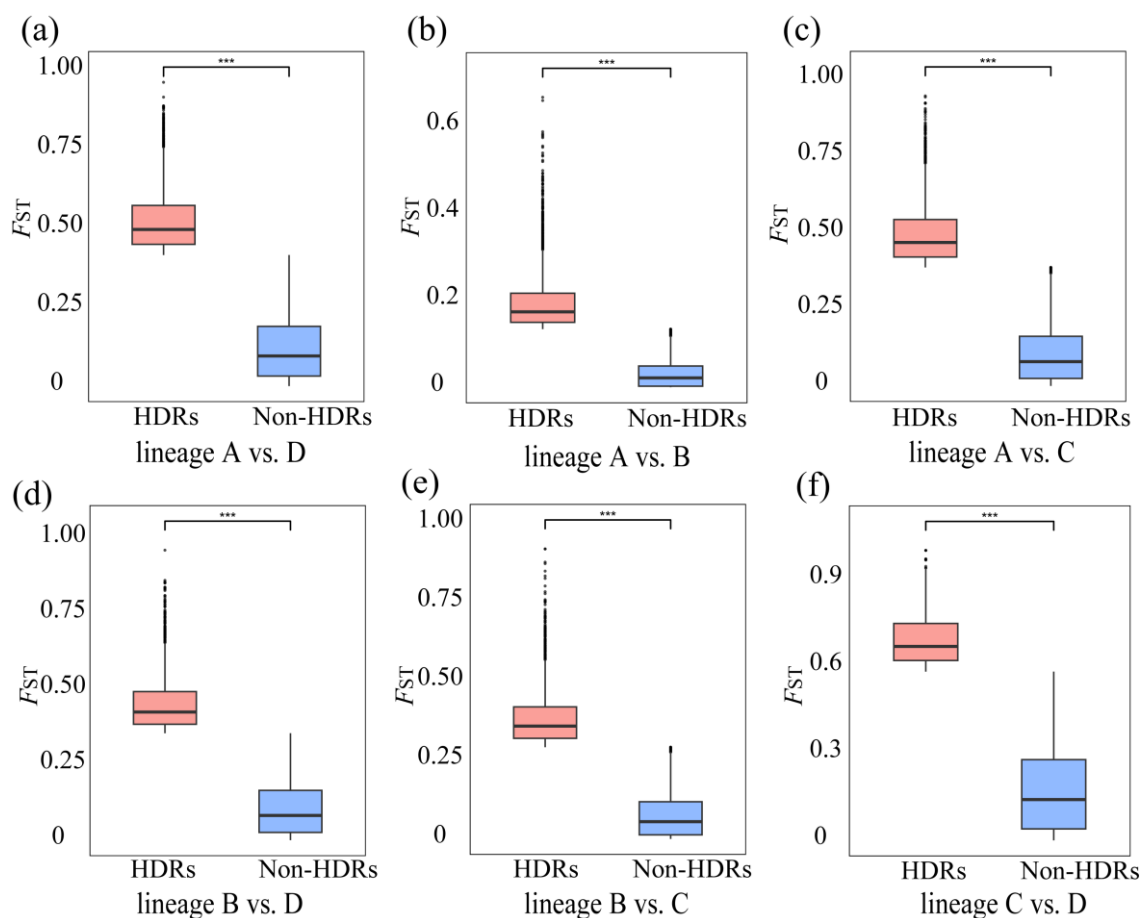


图 3.2 高山栎组植物各谱系对间的高分化区域 (HDRs) 与非高分化区域 (non-HDRs) 比较。(a) 谱系 A 与 D 之间 HDRs 和 non-HDRs 的 F_{ST} 分布情况; (b) 谱系 A 与 B 之间 HDRs 和 non-HDRs 的 F_{ST} 分布情况; (c) 谱系 A 与 C 之间 HDRs 和 non-HDRs 的 F_{ST} 分布情况; (d) 谱系 B 与 D 之间 HDRs 和 non-HDRs 的 F_{ST} 分布情况; (e) 谱系 B 与 C 之间 HDRs 和 non-HDRs 的 F_{ST} 分布情况; (f) 谱系 C 与 D 之间 HDRs 和 non-HDRs 的 F_{ST} 分布情况

Figure 3.2 Comparison of HDRs and non-HDRs between lineage pairs. (a) Boxplots showing the distributions of F_{ST} for HDRs and non-HDRs in lineage A/D, (b) Boxplots showing the distributions of F_{ST} for HDRs and non-HDRs in lineage A/B, (c) Boxplots showing the distributions of F_{ST} for HDRs and non-HDRs in lineage A/C, (d) Boxplots showing the distributions of F_{ST} for HDRs and non-HDRs in lineage B/D, (e) Boxplots showing the distributions of F_{ST} for HDRs and non-HDRs in lineage B/C, (f) Boxplots showing the distributions of F_{ST} for HDRs and non-HDRs in lineage C/D

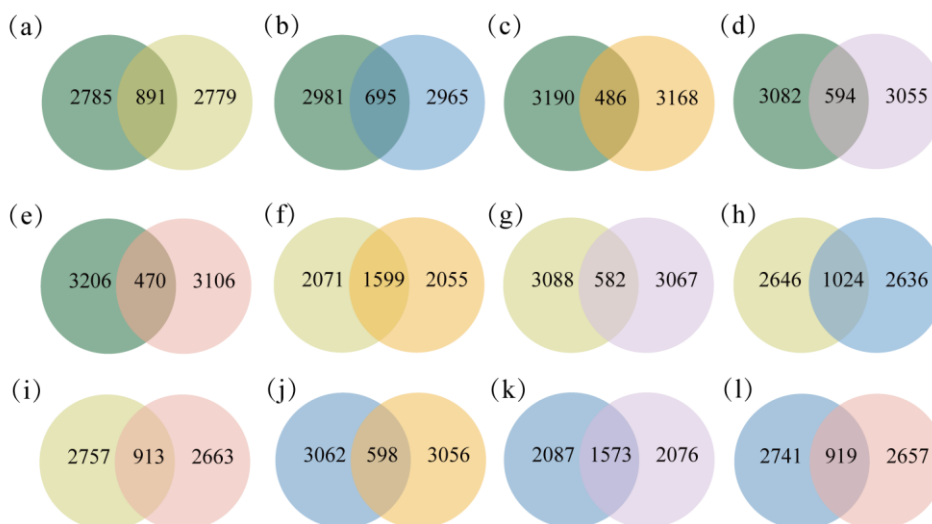


图 3.3 比较不同谱系对之间高分化区域的重叠数量。(a) 谱系 A/B 与 A/C; (b) 谱系 A/B 与 A/D; (c) 谱系 A/B 与 B/C; (d) 谱系 A/B 与 B/D; (e) 谱系 A/B 与 C/D; (f) 谱系 A/C 与 B/C; (g) 谱系 A/C 与 B/D; (h) 谱系 A/C 与 A/D; (i) 谱系 A/C 与 C/D; (j) 谱系 A/D 与 B/C; (k) 谱系 A/D 与 B/D; (l) 谱系 A/D 与 C/D

Figure 3.3 Comparison of the number of overlapping highly diverged regions between different lineage pairs. (a) lineage A/B vs. lineage A/C, (b) lineage A/B vs. lineage A/D, (c) lineage A/B vs. lineage B/C, (d) lineage A/B vs. lineage B/D, (e) lineage A/B vs. lineage C/D, (f) lineage A/C vs. lineage B/C, (g) lineage A/C vs. lineage B/D, (h) lineage A/C vs. lineage A/D, (i) lineage A/C vs. lineage C/D, (j) lineage A/D vs. lineage B/C, (k) lineage A/D vs. lineage B/D, (l) lineage A/D vs. lineage C/D

3.3.3 选择性信号检测

通过跨群体复合似然比评估选择压力，将跨群体复合似然比值前 5% 的基因组窗口视为受正选择的候选区域。谱系 A 与 D 间鉴定出 1489 个受正选择区域，谱系 A 与 B 鉴定出 2528 个受正选择区域，谱系 A 与 C 间鉴定出 2216 个受正选择区域，谱系 B 与 D 间鉴定出 2528 个受正选择区域，谱系 B 与 C 间鉴定出 2216 个受正选择区域，谱系 C 与 D 间鉴定出 2216 个受正选择区域。

根据 F_{ST} 和跨群体复合似然比进行选择清除分析筛选谱系间受正选择的高分化区域结果显示 (图 3.4): 谱系 A 与 D 间有 165 个正选择高分化区域, 占高分化区域 (HDRs) 总数的 4.5%, 涉及 765 个 SNPs。这些 SNP 中有 486 个 SNPs 位于基因间区 (Intergenic region), 280 个 SNPs 位于转录本中 (Transcript)。谱系 A 与 B 间有 749 个正选择高分化区域, 占高分化区域总数的 19.8%, 涉及 3504 个 SNPs。这些 SNPs 中有 2178 个 SNPs 位于基因间区, 1326 个 SNPs 位于转录本中。谱系 A 与 C 间有 459 个正选择高分化区域, 占高分化区域总数的 12.5%, 涉及 2020 个 SNPs。这些 SNPs 中有 1275 个 SNPs 位于基因间区, 745 个 SNPs 位于转录本中。谱系 B 与 D 间有 396 个正选择高分化区域, 占高分化区域总数的 10.1%, 涉及 2121 个 SNPs。这些 SNPs 中有 1190 个 SNPs 位于基因间区, 931 个 SNPs 位于转录本中。谱系 B 与 C 间

有 491 个正选择高分化区域，占高分化区域总数的 13.4%，涉及 2475 个 SNPs。这些 SNPs 中有 1526 个 SNPs 位于基因间区，949 个 SNPs 位于转录本中。谱系 C 与 D 间有 239 个正选择高分化区域，占高分化区域总数的 6.7%，涉及 1258 个 SNPs。这些 SNPs 中有 733 个 SNPs 位于基因间区，525 个 SNPs 位于转录本中。

此外，运用韦恩图对 6 组谱系对经历正选择的高分化区域进行可视化分析结果显示，尽管大部分共享片段具有谱系对特异性，但仍存在少量高分化区域在两个或更多不同的谱系对之间出现共享的情况（图 3.5）。

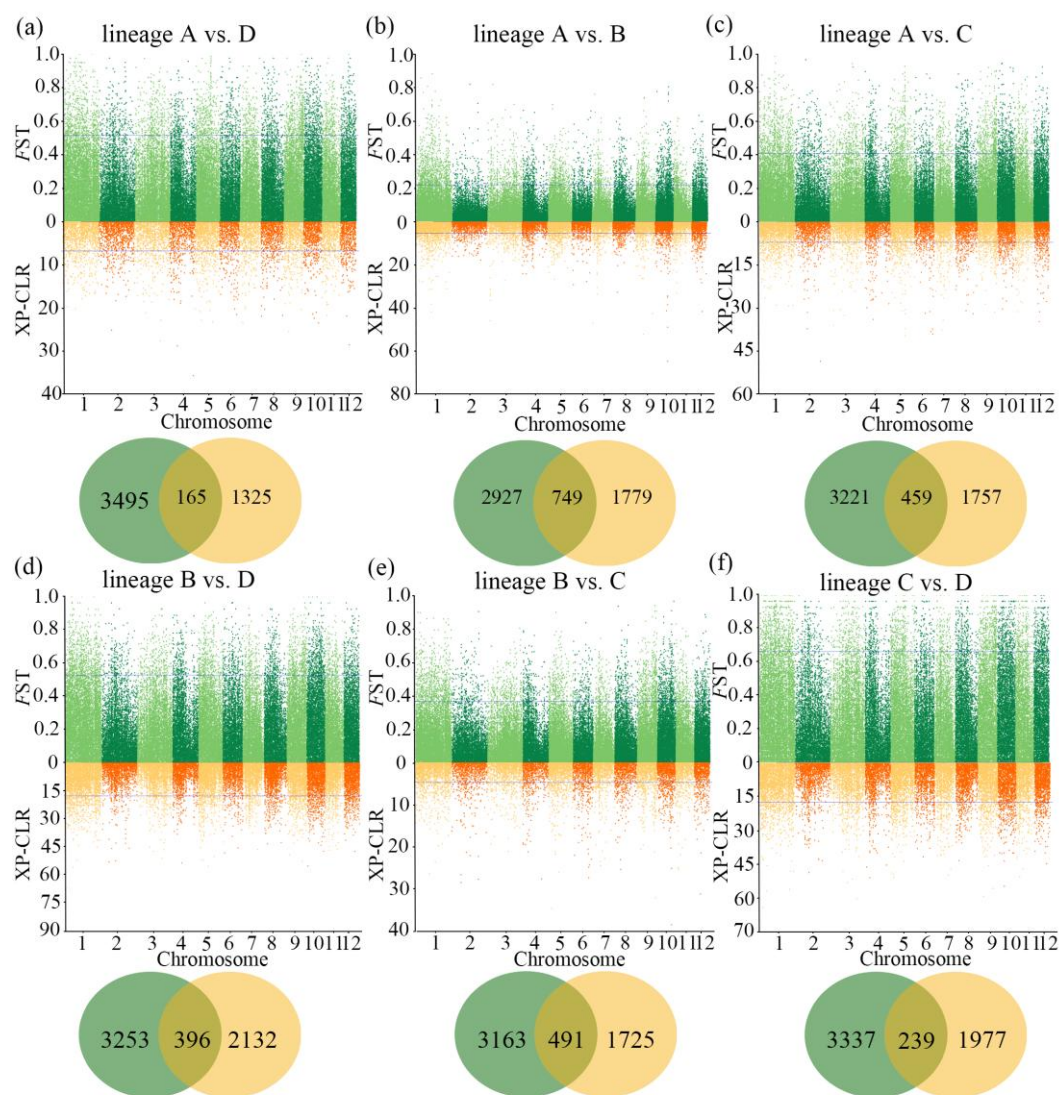


图 3.4 高山栎组植物六个谱系对的 F_{ST} 和跨群体复合似然比分析的曼哈顿图。(a) 谱系 A 与 B；(b) 谱系 A 与 C；(c) 谱系 A 与 D；(d) 谱系 B 与 C；(e) 谱系 B 与 D；(f) 谱系 C 与 D。

维恩图展示由 F_{ST} 和跨群体复合似然比方法检测到的重叠窗口区域

Figure 3.4 Manhattan plot of F_{ST} and XP-CLR analyses for six lineage pairs (a) lineage A/B, (b) lineage A/C, (c) lineage A/D, (d) lineage B/C, (e) lineage B/D, (f) lineage C/D. The Venn diagram plots with overlapping windows among HDRs detected by F_{ST} and XP-CLR

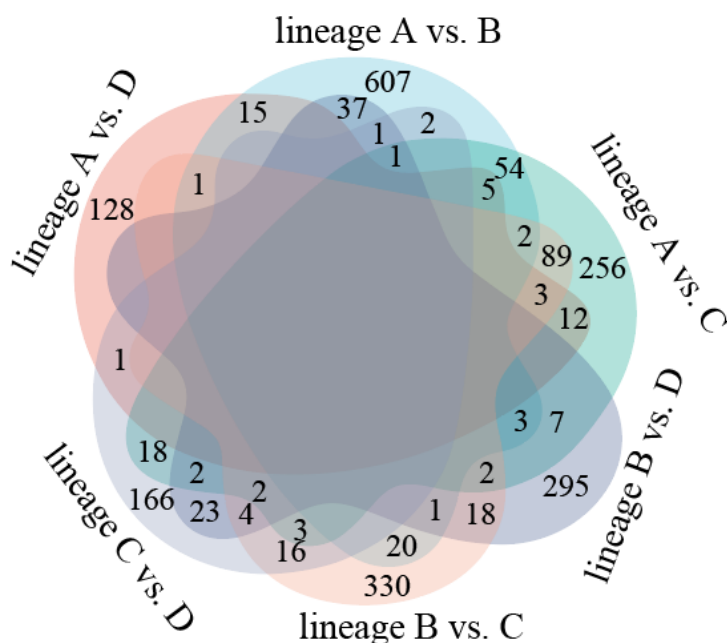


图 3.5 不同谱系对之间受正选择高分化区域的共享片段的韦恩图

Figure 3.5 Shared segments of positively selected highly diverged regions between different lineage pairs using Venn diagram

3.3.4 正选择基因

基于前面 6 个谱系对鉴定出的正选择高分化区域分别进行基因本体 (GO) 富集分析, 此处仅表述生物学过程类别中显著性最高的前 20 个条目 (图 3.6)。在谱系 A 与 B 间的正选择高分化区域共注释到 321 个受选择候选基因, 获得 302 个显著性匹配的拟南芥同源基因, GO 富集分析结果显示主要与细胞、代谢及生物合成过程, 对刺激响应、细胞胁迫及含氧化合物的响应及发育调控等过程相关。在谱系 A 与 C 间的正选择高分化区域总共注释到 204 个受选择候选基因, 获得 191 个显著性匹配的拟南芥同源基因, GO 富集分析结果表明相关基因主要富集于发育和生殖相关过程以及调控等过程。在谱系 A 与 D 间的正选择高分化区域总共注释到 67 个受选择候选基因, 获得 59 个显著性匹配的拟南芥同源基因, GO 富集分析结果表明相关基因主要富集于代谢与发育相关过程、与生殖相关过程及细胞对刺激及含氧化合物的响应等过程。在谱系 B 与 C 间的正选择高分化区域总共注释到 237 个受选择候选基因, 获得 218 个显著性匹配的拟南芥同源基因, GO 富集分析结果表明相关基因主要富集于植物器官发育及形态建成相关过程、发育、调控及离子跨膜转运等过程。在谱系 B 与 D 间的正选择高分化区域总共注释到 229 个受选择候选基因, 获得 221 个显著性匹配的拟南芥同源基因, GO 富集分析结果表明相关基因主要富集于细胞过程调控、代谢过程、基因表达及转录调控、花器官发育等过程。在谱系 C 与 D 间的正选择高分化区域总共注释到 125 个受选择候选基因, 获得 120 个显著性匹配的拟南芥同源

基因，GO富集分析结果表明相关基因主要富集于发育调控、代谢、胁迫的响应、细胞对胁迫及含氧化合物的响应等过程。

在 6 个谱系对的正选择高分化区域中总共鉴定出 910 个受选择候选基因。通过比对至拟南芥基因组注释数据库进行同源基因筛选，获得 864 个具有显著性匹配的同源基因标识符的受选择候选基因，其中只在一个谱系对中特有的同源基因有 679 (78.6%) 个，195 (22.6%) 个基因共享于 2-4 个不同谱系对。通过对 195 个共享于不同谱系对中的拟南芥同源基因的功能查找，发现 66 个拟南芥同源基因已在相应研究体现出具体功能，鉴定出一系列在高山栎组植物谱系分化过程中的关键候选基因，主要涉及生态偏好的适应（温度响应、干旱胁迫、光响应）、生殖隔离（开花时间调控、花粉识别、花粉生长及胚胎发育）、非生物胁迫耐受（重金属耐受性、盐胁迫响应）及表型发育（种子大小调控、叶片发育）等的基因（附表 4）。如 *BRS1*、*ERDJ3B*、*CNGC20*、*PRL1*、*PRR7* 和 *KOIN* 是温度响应的重要基因，*CYP707A1*、*TLP3*、*GAUT10*、*HPCA1*、*GPA1*、*GTL1* 和 *BRM* 等基因与干旱胁迫的响应有关，*PGX3*、*AtHSPR*、*ATC3H17* 和 *MYB3* 等基因与盐胁迫的响应有关，*DWF6*、*SEU*、*PSY* 和 *CK1* 等基因与光响应有关，*NIP1;2*、*ANL2* 和 *PCR2* 等基因与重金属的耐受性有关，*PWO1*、*ARI8*、*SEC3A*、*PEP*、*NDR5*、*PGM3*、*BBM* 和 *PTB1* 等基因与合子前生殖隔离有关，*TIM50*、*TTM2*、*TAF13* 和 *KUP4* 等基因可能与合子后的生殖隔离有关（附表 4）。此外，还发现了 *ANT*、*DPA4*、*PAT14*、*GVS1* 和 *DRD1* 等与叶片的发育有关的基因，*XPO1A*、*GSL04*、*IQD22* 和 *NUDX19* 等基因与氧化胁迫响应有关（附表 4）。

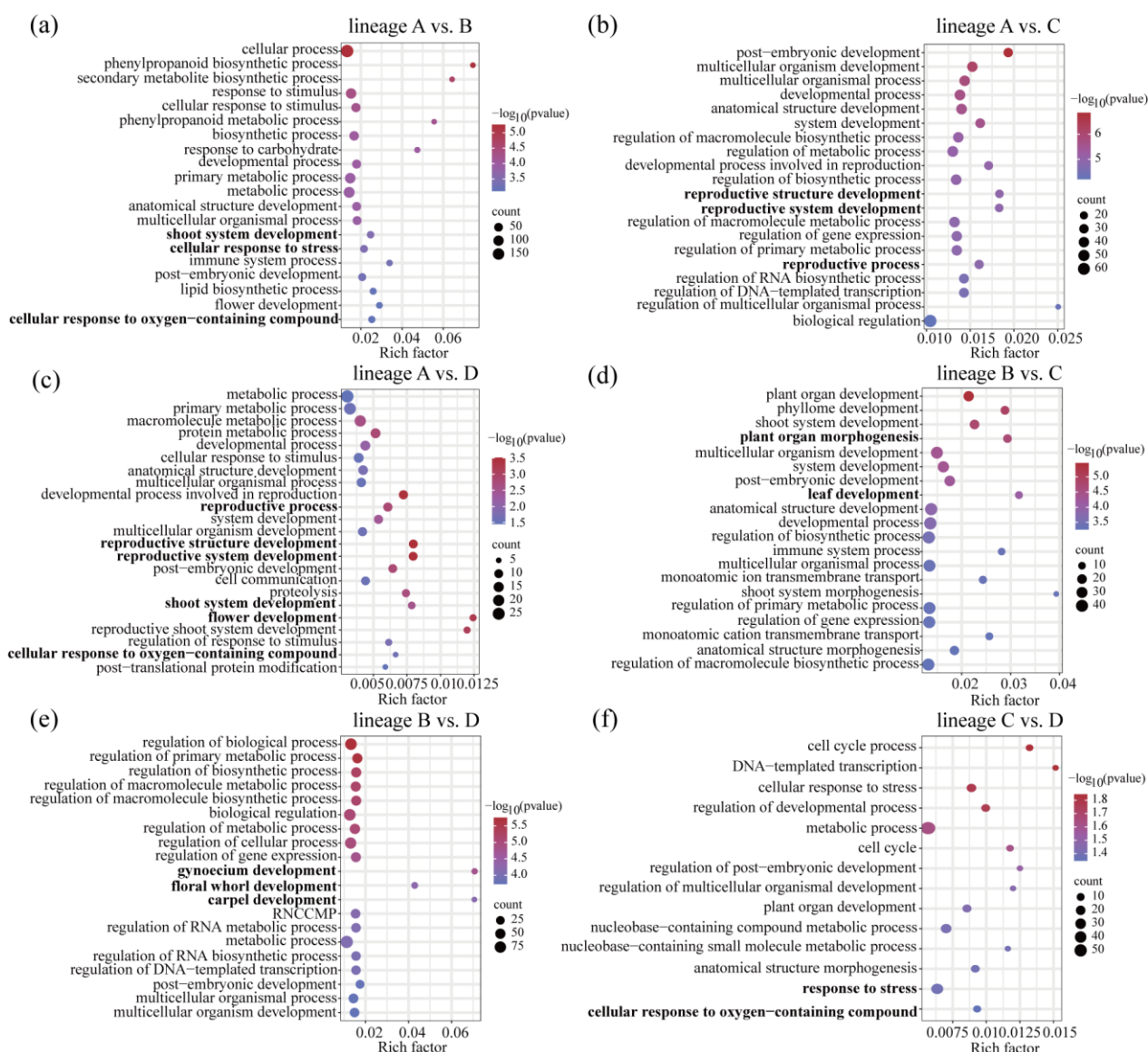


图 3.6 高山栎组植物的每个谱系间的生物学过程中显著性前 20 名的 GO 富集气泡图。气泡大小表示每个 GO 术语中富集基因数量，而气泡的颜色则表示富集的显著程度。仅显示显著富集的生物学过程

Figure 3.6 Top 20 of Gene Ontology (GO) enrichment analysis of candidate genes under selection in biological process for each lineage of *Quercus* sect. *Heterobalanus*. Bubble size represents the number of genes enriched in each GO term, while bubble color indicates the significance level of enrichment. Only significantly enriched biological processes are shown

3.3.5 顺式作用调控元件

对位于正选择高分化区域的非编码区的 SNPs 进行顺式调控元件基序结果表明 (图 3.7), 在谱系 A 与 D 间的正选择高分化区域的转录本 SNPs 中, 87 个 (11%) SNPs 位于基因上下游区域, 80% (70 个) 的 SNPs 位于已知植物顺式调控 DNA 元件的基序中, 其中 38% 的 SNPs 与光、叶片、花粉和种子发育调控相关联。在谱系 A 与

B 间的正选择高分化区域的转录本 SNPs 中，473 个（36%）SNPs 位于基因上下游区域，69%（328 个）的 SNPs 位于已知植物顺式调控 DNA 元件的基序中，其中 29% 的 SNPs 与温度、水分、光、叶片、氧气、花粉和种子发育调控相关联。在谱系 A 与 C 间的正选择高分化区域的转录本 SNPs 中，286 个（38%）SNPs 位于基因上下游区域，66%（190 个）的 SNPs 位于已知植物顺式调控 DNA 元件的基序中，其中 25% 的 SNPs 与温度、水分、光、叶片、氧气、花粉和种子发育调控相关联。在谱系 B 与 D 分化的正选择高分化区域的转录本 SNPs 中，349 个（38%）SNPs 位于基因上下游区域，73%（257 个）的 SNPs 位于已知植物顺式调控 DNA 元件的基序中，其中 30% 的 SNPs 与温度、水分、光、叶片、氧气、花粉和种子发育调控相关联。在谱系 B 与 C 间的正选择高分化区域的转录本 SNPs 中，357 个（38%）SNPs 位于基因上下游区域，74%（263 个）的 SNPs 位于已知植物顺式调控 DNA 元件的基序中，其中 29% 的 SNPs 与温度、水分、光、叶片、氧气、花粉和种子发育调控相关联。在谱系 C 与 D 间的正选择高分化区域的转录本 SNPs 中，191 个（36%）SNPs 位于基因上下游区域，68%（129 个）的 SNPs 位于已知植物顺式调控 DNA 元件的基序中，其中 34% 的 SNPs 与温度、光、叶片、氧气、花粉和种子发育调控相关联。

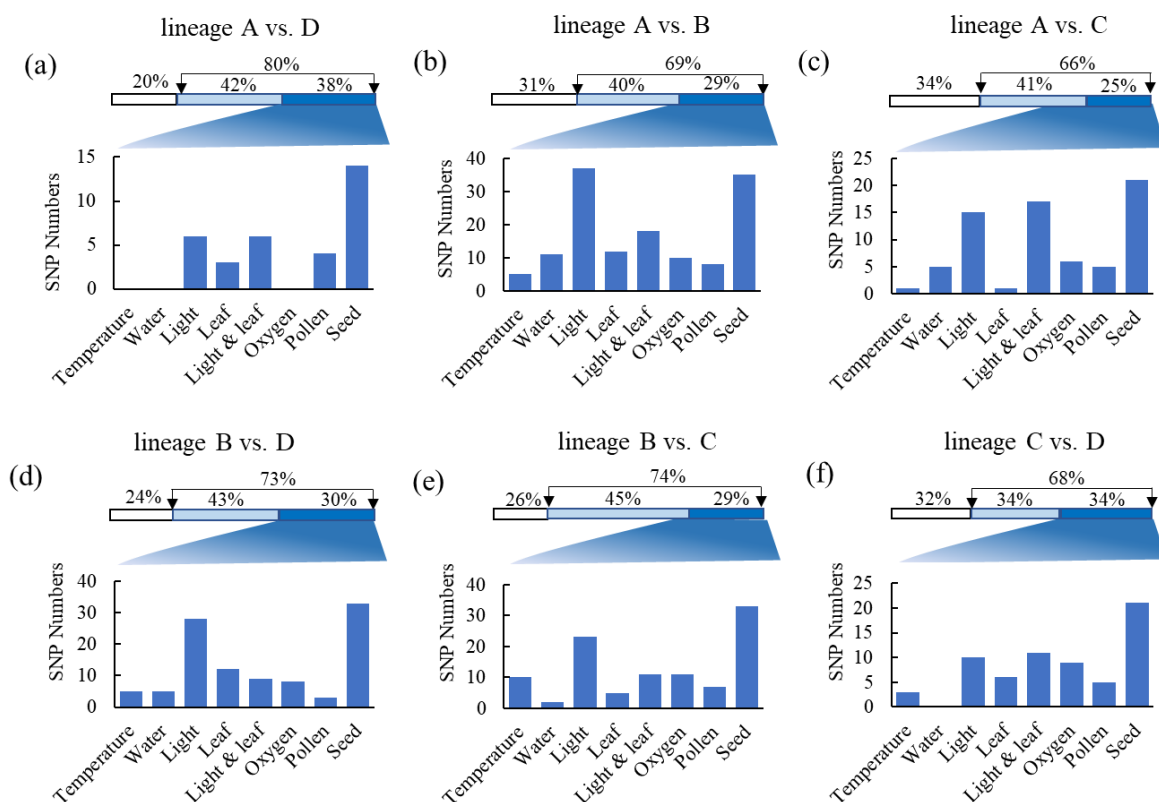


图 3.7 预测高山栎组植物各谱系间位于基因调控区域内的正选择高分化区域的顺式作用调控元件。蓝色条形图表示在已知植物顺式调控 DNA 元件的模式中发现的单核苷酸多态性的数量

Figure 3.7 Predicted functional regulatory elements of positively selected highly diverged regions located within regulatory regions for each lineage of *Quercus* sect. *Heterobalanus*. Blue bars represent the number of SNPs found in motifs of known plant cis-regulatory DNA elements

3.4 结论

3.4.1 正选择的高分化区域

物种谱系间的遗传分化是长期自然选择与环境适应协同作用的结果，其中正选择驱动的功能基因分化是解析物种适应性辐射与局地适应机制的重要内容^[236, 237]。青藏高原多变的气候条件与地形环境很可能通过改变物种间的联系与环境条件，从而形成新的栖息地与生态位，促进物种或谱系间的快速分化^[220]。 F_{ST} 与跨群体复合似然比结果分析显示，高山栎组植物6个谱系对中存在4.5% - 19.8%的正选择高分化区域，进一步发现少数的受正选择的高分化区域在多个谱系对中存在共享，且未检测到在所有谱系对中共有的正选择高分化区域。虽然基于参考基因组的重测序可能存在一定的参考基因组偏向性（Reference bias），某些区域可能被错误比对或在质量控制过程中被过滤，从而导致部分真实遗传变异信息的缺失或低估，可能会影响正选择的高分化区域的准确性，并可能在一定程度上改变对分化区域分布及跨谱系共享选择信号的识别。然而，在近期运用基因组数据开展植物正选择的高分化区域的鉴定研究中，Geng等^[238]在毛茛科（Ranunculaceae）耧斗菜属（*Aquilegia* L.）的两个近缘种无距耧斗菜（*A. ecalcarata* Maxim.）和甘肃耧斗菜（*A. oxysepala* var. *kansuensis* Brühl）所形成的东西谱系的物种对间发现正选择的高分化区域为17% - 38%。Yang等^[239]揭示了金露梅（*Dasiphora fruticosa* (L.) Rydb.）的不同遗传单元间的正选择的高分化区域为37% - 53%。尽管不同研究中筛选正选择的高分化区域所设置的阈值存在差异，但是通过比较可以发现本研究中不同谱系对之间的高分化区域与其他植物相比整体比例略微偏低。这一结果表明高山栎组植物不同谱系间的分化选择压力存在异质性，各谱系对的分化过程相对独立，呈现出谱系特异性。即使高山栎组植物间谱系分化存在特异性，但在少量的跨谱系共享选择信号能够体现出特定谱系对在相似选择压力下的进化保守性，更反映了祖先遗传特征与适应性能力的保留，可能是不同谱系对相似生境胁迫平行进化的结果。

3.4.2 选择信号揭示的谱系分化机制

不同地理种群可以通过遗传漂变和突变累积造成遗传差异，进而为局地适应与自然选择提供遗传基础^[240]。本研究基于 F_{ST} 和跨群体复合似然比特征指标对谱系水平进行的选择性清除分析结果显示，高山栎组植物各谱系受到选择的区域存在明显差异，谱系A/B间的正选择高分化区域最多，谱系A/D间最少，这在一定程度上表明主要位于西藏地区的谱系A与秦巴山-华东地区的谱系D的分化时间较早且长期隔离，适应能力出现了高度分化。这一结果与遗传结构体现出的谱系D先分化出来一致，再一次符合了高山栎组植物的“走出青藏高原”假说。基于对6个谱系对中鉴定出的正选择高分化区域进行基因本体（GO）富集分析结果显示主要参与代谢及生物合成过程，对刺激的响应、含氧化合物的响应、生殖相关过程以及调控、植物器

官发育及形态建成等过程。此外，从基因组分布情况来看，所有谱系对中正选择的高分化区域的 SNPs 位于基因间区较多，而位于转录本内的 SNPs 所占比例相对较低。这一结果表明自然选择在高山栎组植物不同谱系的分化过程中，可能更多作用于基因调控区域而非直接改变蛋白质编码序列，从而在维持核心基因功能稳定性的同时，实现在适应极端环境中对表型和适应性性状的调控^[241]。已有许多研究表明非编码的调控元件在植物的表型发育、适应和演化过程中有着重要的作用^[242, 243]。通过对谱系之间的正选择高分化区域进行顺式作用调控元件预测结果表明，在不同谱系间的选择性信号中发现与光调控（光周期、昼夜节律、UV-B 等）、花粉与花药特异、叶片发育调控以及种子发育的调控元件普遍存在。因此，光响应与植物生长发育相关的顺式调控变异是驱动高山栎组植物谱系分化的保守、共有机制；而温度、水分等逆境胁迫相关的调控变异则体现了谱系间特异性的适应性选择。

在高山栎组植物不同遗传谱系间的比较中，发现了若干功能基因在谱系间既存在共享性，也呈现谱系特异性。在 6 个谱系对的正选择的高分化区域中鉴定出共享的拟南芥同源基因进行功能查找，结果发现有 3 个基因稳定存在单一谱系与其余谱系中，表明这 3 个基因驱动该谱系与其他谱系适应性分化的关键基因，主要涉及对刺激的响应和生殖过程。第一个为在谱系 A 与其他谱系的正选择高分化区域中检测到的 *ERDJ3B* 基因。已有研究表明 *ERdj3B* 在高温胁迫下能够维持植物生殖发育和热耐受性，高温胁迫下容易导致花器官发育异常、花粉活力下降，影响繁殖成功率^[244, 245]。谱系 A 主要分布于具有高海拔、强太阳辐射以及显著的昼夜温差等环境^[246]，此地区剧烈的温度波动可能增加植物细胞内蛋白质折叠压力，从而诱导内质网应激反应，并对植物生殖发育产生不利影响。在这种环境条件下，与蛋白折叠和质量控制相关的分子伴侣系统可能受到更强的自然选择压力。相比位于地形复杂且受亚洲季风气候控制的横断山脉地区^[247]的谱系 B 和 C，以及分布于气候相对稳定^[248]的谱系 D，谱系 A 所处的极端高山环境可能对维持蛋白质稳态和生殖稳定性的相关基因产生更强的选择压力，从而导致 *ERDJ3B* 在谱系 A 与其他谱系之间表现出明显的遗传分化。第二个为在谱系 B 与其他谱系的正选择高分化区域中检测到的 *PGX3* 基因，该基因在植物的气孔动态调控和适应盐胁迫过程中具有重要作用。如 Rui 等^[249]研究发现 *PGX3* 能够通过调节果胶的降解改变细胞壁结构并影响气孔动态，从而参与植物生长与环境响应过程。此外，在盐胁迫等非生物胁迫条件下，植物常伴随活性氧（ROS）积累，通过激活抗氧化系统维持细胞膜稳定性和细胞稳态，从而提高耐盐性^[250, 251]。对于主要分布于横断山脉北部的谱系 B 而言，该区域复杂的地形结构，剧烈变化的气候条件以及水分条件差异，可能对植物产生较强的环境选择压力，促进了谱系 B 与其他谱系间的分化。第三个为在谱系 C 与其他谱系的正选择高分化区域中检测到了具有低温胁迫响应的 *KOIN* 基因^[252]。研究表明低温胁迫会破坏植物细胞膜流动性、诱导细胞内活性氧积累、抑制光合与代谢酶活性，直接威胁植物的存

活、生长与繁殖。高山栎组植物普遍生长在高海拔地区，面临严峻的低温压力^[253, 254]。已有研究表明在高山草本和木本植物中，低温胁迫响应基因的正向选择与海拔和小气候环境梯度密切相关^[255, 256]。因此，谱系 C 中 *KOIN* 基因可能反映了谱系 C 在复杂低温环境下的局地适应性进化，从而形成了与其他谱系不同的基因组特征。

此外，本研究通过跨谱系的基因组比较分析，发现一些与器官发育和形态建成密切相关的候选基因，这些基因可能是高山栎组植物在高海拔环境中表现出形态多样化的遗传基础。在谱系 A/C 和谱系 B/C 中均鉴定到了 *ANT* 基因。已有研究表明 *ANT* 是决定器官大小和形态的调控因子，具有调控叶片和花瓣表型大小或宽窄变化的功能^[257, 258]。同时，在谱系 A/B 和谱系 B/C 中也鉴定到了 *DPA4* 基因，该基因已被证实参与调控种子大小和叶片边缘发育^[259]。这些结果表明即使不同谱系在基因组水平上存在明显分化，共享的表型调控基因可能导致形态特征在谱系间的相对保守，或许可以作为形态物种与遗传谱系不完全一致现象的一个解释。谱系 A、B 和 C 分布于青藏高原，该地区高度环境异质性的地形和气候条件，为植物形态多样性的形成提供了重要的选择压力。在该地区的许多植物类群中的叶片形态特征表现出高度的形态多样化的现象^[188, 260, 261]。因此，即使在能够形成清晰的谱系 A、B 和 C 的情况下，它们仍可能在叶片、花瓣或种子等形态特征上表现出相似性，从而对传统基于形态的物种划分带来挑战。

结合在共享拟南芥基因对植物的功能作用的预测中，高山栎组植物的各谱系在对环境的适应性进化过程中，受选择的关键候选基因主要与调节开花、种子和叶形态发育以及应对高温、低温、盐和干旱等逆境防御相关。基于以上发现，未来研究可进一步结合 d_{xy} 或 π 等指标与跨群体复合似然比的交集进行严格验证，以更全面地提高选择信号的准确性以评估各谱系的遗传分化模式及其适应性进化过程。

3.5 小结

本章节运用基因组扫描鉴定不同谱系间的高分化区域与选择性清除信号结果表明, 6个谱系对间的高分化区域中只有部分区域受到正选择, 这些正选择高分化区域数量、占比及 SNPs 分布均存在显著差异, 其中谱系 A 与 B 间正选择高分化区域占比最高 (19.8%), 而谱系 A 与 D 间正选择高分化区域占比最低 (4.5%), 揭示了不同谱系对在进化过程中受到的选择压力不同, 主要位于西藏地区的谱系 A 与秦巴山-华东地区的谱系 D 的分化时间较早且长期隔离, 适应能力出现了高度分化。这一结果与遗传结构体现出的谱系 D 先分化出来一致, 符合高山栎组植物的“走出青藏高原”假说。GO 富集分析显示候选基因主要涉及发育调控、非生物胁迫响应及生殖相关过程, 与高山栎组植物在高海拔的山地环境适应能力高度相关。此外, 存在大量位于基因上下游区域的候选 SNPs 在顺式调控元件区域, 多数与光响应和叶片发育相关, 表明了调控变异在谱系分化与适应性进化中可能发挥重要作用。跨谱系的基因组比较分析仅在青藏高原地区谱系对中发现共享的表型调控基因, 可能是高山栎组植物在高海拔环境中表现出形态多样化的遗传基础, 这种共享与特异相结合的模式为理解高山植物谱系分化和形态变异提供了分子基础。

4 高山栎组植物谱系的局地适应研究

4.1 引言

从基因组水平进行遗传谱系识别进行景观基因组学研究，探讨环境因子对其遗传变异的影响，对于理解植物的适应性进化有着重要作用^[32, 225, 262, 263]。Walden等^[264]对北美地区的琴叶拟南芥 (*Arabidopsis lyrate* (L.) O'Kane & Al-Shehbaz) 的东西两个遗传谱系进行异常位点识别和 GO 富集分析研究表明两个遗传谱系之间几乎不属于平行进化，但相似的生物学过程（如应激反应、昼夜节律、花发育和繁殖）参与了适应过程。对于分布范围较广且生态适应能力较强的木本植物而言，长期的环境异质性往往会导致种群在不同区域形成相对独立的遗传谱系^[139, 265]。不同的遗传谱系在历史进化过程中可能经历了不同程度的地理隔离、基因流限制以及选择压力，从而在遗传组成和适应性特征上表现出一定差异。

景观基因组学 (Landscape genomics) 通过将基因组数据与空间环境数据相结合，可以识别与环境梯度相关的适应性基因位点，从而揭示自然选择在不同生态环境中的作用机制^[263, 266]。大量研究表明，在具有明显地理遗传结构的物种中，不同遗传谱系往往会处于不同的生态环境中，并可能受到不同的选择压力，从而导致适应性遗传变异在不同谱系之间呈现出异质性的分布格局^[267]。如果在整个物种范围内直接进行基因型-环境关联分析，群体遗传结构可能会产生较强的混杂效应，从而降低检测局地适应信号的准确性^[268]。因此，在明确遗传结构并划分遗传谱系的基础上，分别对各谱系开展局地适应分析能够在一定程度上减少种群历史与空间结构对统计结果的干扰，更加有效地识别真实的环境适应性信号。

本章节的研究将对高山栎组植物的四个遗传谱系分别开展景观基因组学分析。通过对每个谱系识别潜在的异常位点，并进一步利用冗余分析与梯度森林模型评估环境变量对基因组变异的影响，旨在揭示不同遗传谱系在复杂环境梯度下的适应性分化模式及其潜在的环境驱动机制，进而深化对物种局地适应过程以及遗传分化机制的认识，为理解该类群在复杂生态环境中的适应性进化提供理论基础。

4.2 分析方法

4.2.1 环境因子

高山栎组植物存在广泛的同域分布范围，因此光照、温度和水分等资源存在高度的共享，后续分析基于遗传谱系水平的景观基因组学分析采用第 2 章节进行共线性分析的 7 个环境变量，以探究在面临相同压力下不同谱系的环境适应性。

4.2.2 适应性位点检测

为全面鉴定与局部适应相关的遗传位点，本研究采用四种方法检测异常位点来识别可能受自然选择影响的候选 SNPs 位点并最大限度减少假阳性，分别为两种基于

遗传分化和两种基于基因型-环境关联的检测异常位点方法。

(1) 基于遗传分化 F_{ST}

首先使用基于 Logistic 回归模型的贝叶斯方法—Bayescan 分析识别异常位点，采用默认配置（5,000 iterations, thinning interval = 10, 20 pilot runs of 5,000 length, burn-in = 50,000, prior odds for the neutral model = 10）^[269]。运用此分析自动估算的 q 值以最小化误检率，并将 q 值 < 0.05 的 SNPs 识别为异常值。然后运用另一种基于主成分分析（PCA）检测适应性位点方法—PCAdapt 方法来鉴别可能受到选择的 SNPs，此分析运用 R 语言中的 PCAdapt 包进行^[270]，通过对个体基因型进行主成分分析识别异常值。通过可视化碎石图确定主成分数目，选取肘部位置或解释方差显著高于随机期望的成分，避免过度拟合或信息丢失。随后，基于选定成分运行 PCAdapt，计算每个 SNP 的 p 值。最后使用 R 包 qvalue 评估 q 值进行控制错误发现率（FDR），取 FDR 值前 10% 的 SNPs 作为 outlier SNPs。

(2) 基于基因型-环境关联分析方法

运用两种基因型-环境关联分析检测适应性位点方法—潜在因素混合模型（Latent factor mixed model, LFMM）和冗余分析（Redundancy analyses, RDA）进行 outlier SNPs 的鉴别。首先在 R 4.3.2 运用 LEA 包中进行 LFMM 分析，采用分层贝叶斯混合模型方法来进行基因组数据和环境变量的回归分析，以关联等位基因频率与环境因素，同时考虑遗传结构作为“潜在因子”^[271]。此分析环境变量被单独检验并作为固定效应引入每个模型，分别对温度、降水、辐射和风速的 7 个变量进行分析；同时通过稀疏非负矩阵分解（Sparse non-negative matrix factorization, SNMF）方法选取潜在因子数量以考虑遗传结构。首先进行遗传数据格式的转换，将 VCF 文件转化成至 0/1/2 形式并进行缺失值的填补，随后进行参数设置为：10000 迭代次数，其中 burn-in 为 5000 次，并执行十次重复运行。本研究最后采用冗余分析作为一种多元方法，以识别其遗传变异与环境因子密切相关的 SNPs。该分析使用 vegan 包中的 rda 函数进行，其中基因型矩阵（编码为 0, 1, 2）作为响应变量，标准化的环境变量作为解释变量，以捕捉不同谱系基因数据对环境变化的响应^[272]。潜在因素混合模型与冗余分析的显著性阈值设定为 $P < 0.001$ 进行异常值的筛选。

最后，本研究将通过 BayeScan、PCAdapt、潜在因素混合模型与冗余分析识别出的异常位点整合到维恩图中，以可视化呈现这些方法的重叠发现与结果。在控制假阳性的同时保留不同选择信号，本研究采用多方法交集共同鉴别出的 SNPs 作为推定的受选择 SNPs（Putative adaptive SNPs），潜在因素混合模型与冗余分析共同鉴别出的 SNPs 作为环境相关的 SNPs（Climate-associated SNPs）。

4.2.3 冗余分析

运用所有 SNPs、推定受选择的 SNPs 和与气候相关的 SNPs 位点分别作为响应变量，地理和气候因子作为解释变量进行冗余和偏冗余（ p RDA）分析，量化地理和环

境变量对高山栎组植物不同谱系的遗传变异的解释贡献度。首先将.vcf文件的SNP数据转换为0、1和2的形式，分别代表常见的纯合与杂合性的等位基因。RDA方法通过三种不同的模型对每个谱系进行对遗传变异的影响进行分析。第一种模型以地理和环境因子为解释变量；第二种模型以地理因子为解释变量；第三种以环境因子为解释变量。此外，为进一步区分地理结构与环境因子对遗传变异的独立作用，本研究采用偏冗余分析对地理结构与环境因子分别进行控制分析。在控制地理因子的条件下，将气候因子作为解释变量，以评估环境因素对遗传变异的纯效应；相应地，在控制气候因子的条件下，将地理因子作为解释变量，以量化空间结构对遗传变异的独立解释贡献。通过比较冗余及偏冗余分析结果，系统揭示地理隔离与环境异质性在塑造高山栎组植物不同谱系遗传结构中的相对作用。为避免解释变量之间的多重共线性对模型结果的影响，所有环境变量在分析前均经过共线性检验和筛选。使用vegan中的rda函数进行冗余分析和anova.cca函数进行999次排列(permutations)对环境与地理变量的显著性评估^[272]。

4.2.4 梯度森林

采用一种非参数的机器学习回归树的方法—梯度森林(Gradient Forest, GF)探究环境变量与适应性遗传变异之间的非线性关联及贡献程度。梯度森林通过整合多棵回归树模型，能够识别等位基因频率沿环境梯度变化的关键环境预测因子^[273]。本研究基于梯度森林中的周转函数(Turnover functions)，评估环境变量对等位基因频率变化的驱动作用。所有分析在R 4.3.2中使用gradientForest包完成。该分析对每个SNPs位点进行构建500棵回归树，树的最大深度参数设置 $\text{maxLevel} = \log_2(0.368n)/2$ ，并采用0.5的变量相关性阈值来计算条件变量重要性，其余参数均采用默认设置。本研究分别使用了推定的受选择SNPs和基因型-环境关联分析筛选出来的与气候相关的SNPs，以及7个气候变量进行构建梯度森林模型，探究不同谱系的适应性遗传变异的环境驱动因子。气候变量的重要性根据其加权 R^2 值进行排序并展示。同时，本研究也对气候变量的分割密度图与累积重要性曲线进行了展示。

4.2.5 适应性候选基因鉴定

为探究每个谱系的适应性遗传变异是否参与了谱系形成过程，本研究进行各谱系推定的受选择位点与3.2.2分析中正选择高分化区域的位点进行重叠分析。重叠SNPs进行基因注释后的基因为鉴定参与适应性分化的候选基因。运用韦恩图统计每个谱系中推定的适应性SNPs与谱系分化区间的重叠基因数量。

4.2.6 基因功能注释与富集分析

对于推定受选择的候选基因和适应性分化候选基因的功能注释与富集分析，同4.2.3。

4.3 结果

4.3.1 适应性位点鉴定

基于遗传分化 (F_{ST}) 鉴定异常值位点的 BayeScan 方法显示在谱系 A 中鉴定出 2308 个异常 SNPs, 谱系 B 中鉴定出 764 个异常 SNPs, 谱系 C 中鉴定出 270 个异常 SNPs, 谱系 D 中鉴定出 441 个异常 SNPs (图 4.1)。基于 PCAdapt 进行异常值位点检测结果显示, 在谱系 A 中鉴定出 4493 个异常 SNPs, 谱系 B 中鉴定出 2839 个异常 SNPs, 谱系 C 中鉴定出 5709 个异常 SNPs, 谱系 D 中鉴定出 4433 个异常 SNPs (图 4.2)。

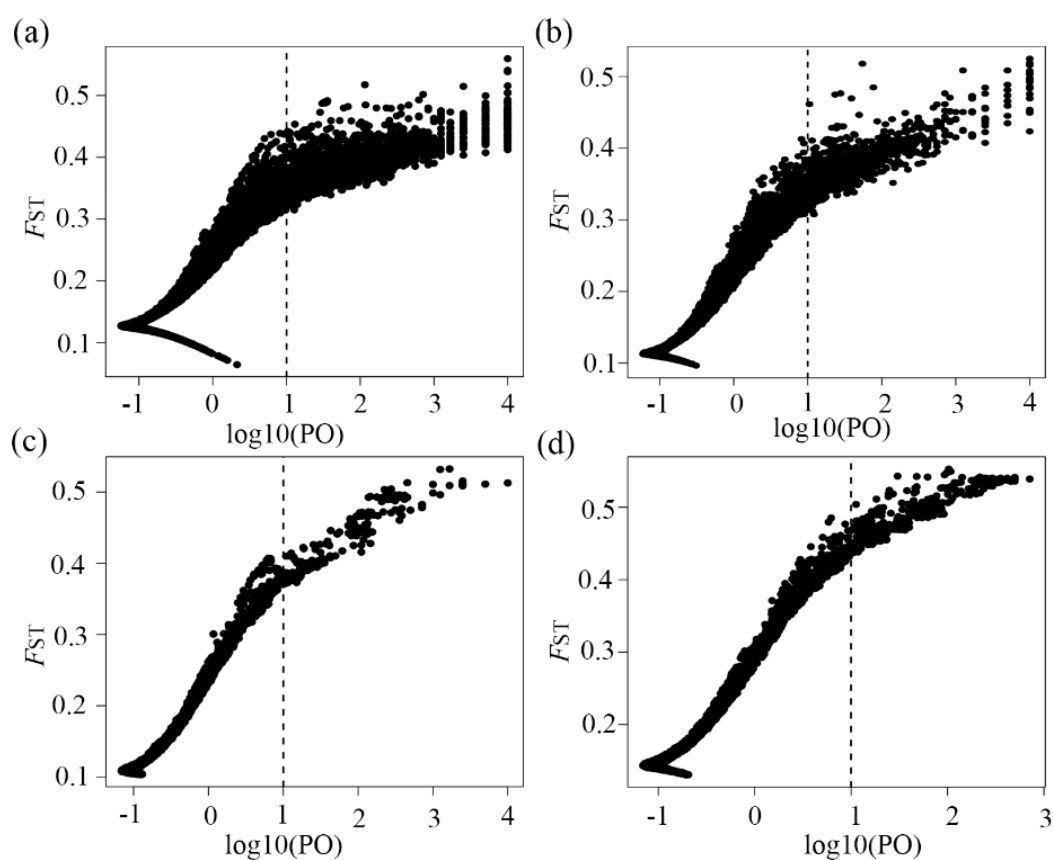


图 4.1 高山栎组植物各谱系的 Bayescan 结果, 垂直的黑色线代表 q 值的截断值 (即 q 值 = 0.05)。(a) 谱系 A 的异常位点鉴定; (b) 谱系 B 的异常位点鉴定; (c) 谱系 C 的异常位点鉴定; (d) 谱系 D 的异常位点鉴定

Figure 4.1 Plot of SNPs of different lineages of *Quercus* sect. *Heterobalanus* according to F_{ST} and $\log_{10}(PO)$, with the indication of the threshold (=1) used to identify the decisive SNPs under selection. The vertical black line represents the cutoff of q values = 0.05. (a) Outlier loci identified for lineage A, (b) Outlier loci identified for lineage B, (c) Outlier loci identified for lineage C, (d) Outlier loci identified for lineage D

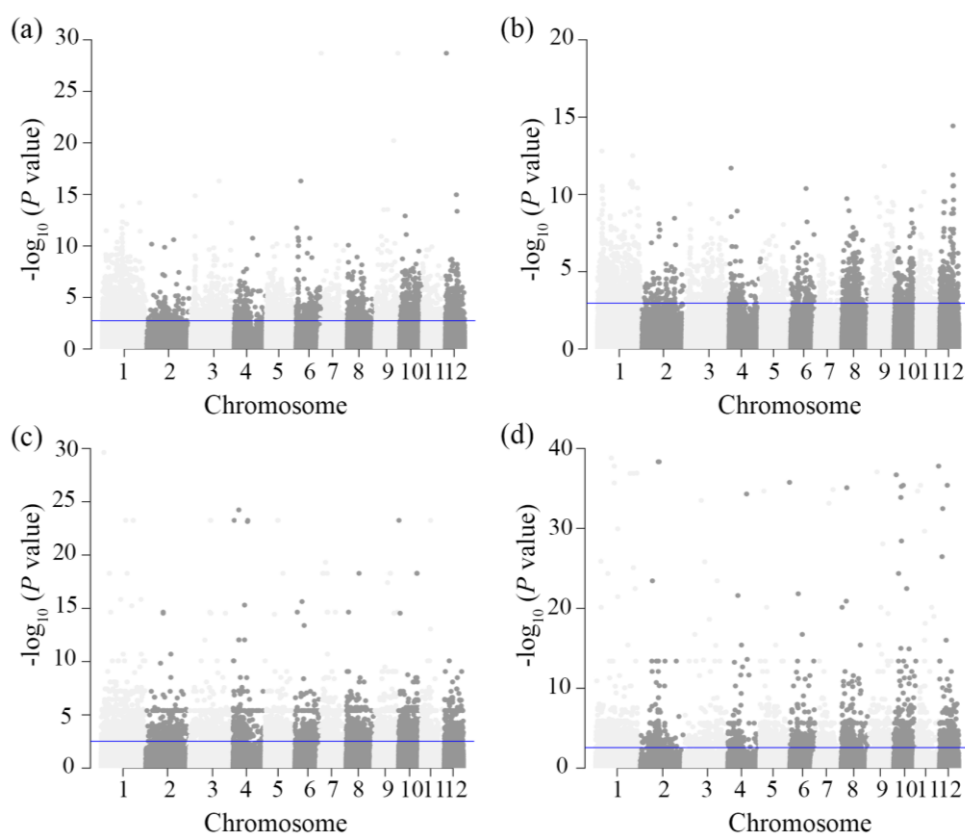


图 4.2 高山栎组植物不同谱系的 PCAdapt 结果。(a) 谱系 A 的异常位点鉴定；(b) 谱系 B 的异常位点鉴定；(c) 谱系 C 的异常位点鉴定；(d) 谱系 D 的异常位点鉴定

Figure 4.2 Manhattan plots depicting PCAdapt detection of outlier loci for four lineages in *Quercus* sect. *Heterobalanus*. (a) Outlier locus identified for lineage A, (b) Outlier locus identified for lineage B, (c) Outlier locus identified for lineage C, (d) Outlier locus identified for lineage D

表 4.1 潜在因素混合模型和冗余分析鉴定高山栎组植物各谱系的候选位点

Table 4.1 Number of candidate SNP loci under putative selection identified by LFMM and RDA of different lineages of *Quercus* sect. *Heterobalanus*

谱系 Lineage	潜在因素混合模型 (LFMM)								冗余分析 RDA
	Bio02	Bio07	Bio09	Bio13	Bio15	Srad	Wind	Sum	
A	1483	2502	2695	2861	2022	2143	2714	12204	11158
B	1279	1588	1134	1495	1902	1730	1318	8118	11824
C	455	427	460	726	408	518	494	2284	12067
D	269	2688	2113	2052	459	1	432	6802	5670

综上所述，基于任意两种方法共同鉴别的 SNPs 作为推定受选择 SNPs (Putative selected SNPs) 结果显示，谱系 A 中鉴定出 3540 个推定受选择 SNPs，谱系 B 中鉴定出 960 个推定受选择 SNPs，谱系 C 中鉴定出 689 个推定受选择 SNPs，谱系 D 中鉴定出 731 个推定受选择 SNPs (图 4.3)。此外基于潜在因素混合模型和冗余分析共同鉴别出的 SNPs 作为气候相关的 SNPs (Climate-associated SNPs) 结果显示，谱系 A

中鉴定出 568 个与气候相关的 SNPs，谱系 B 中鉴定出 352 个与气候相关的 SNPs，谱系 C 中鉴定出 123 个与气候相关的 SNPs，谱系 D 中鉴定出 217 个与环境相关的 SNPs（图 4.3）。韦恩图结果显示各谱系通过 BayeScan、PCAdapt、潜在因素混合模型和冗余分析四种分析方法共同鉴定到的 SNPs 极少，在谱系 A 中鉴定到 32 个共有的 SNPs，谱系 B 中鉴定到 2 个共有的 SNPs，谱系 C 中没有鉴定到共有的 SNP，谱系 D 中鉴定到 3 个共有的 SNPs（图 4.3）。

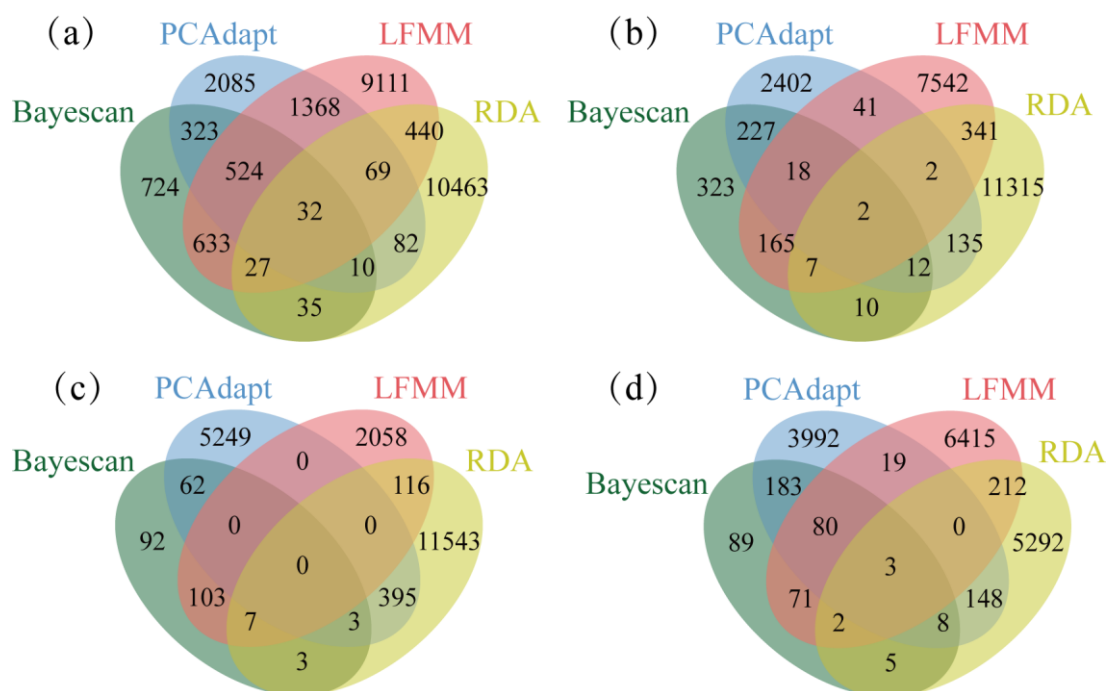


图 4.3 四种方法进行高山栎组植物各谱系的推定受选择位点与气候相关位点的韦恩图。

(a) 谱系 A; (b) 谱系 B; (c) 谱系 C; (d) 谱系 D

Figure 4.3 Venn diagram of putative selected SNP and climate-associated SNP were identified for different lineages of *Quercus* sect. *Heterobalanus* using the four methods. (a) lineage A, (b) lineage B, (c) lineage C, (d) lineage D

4.3.2 基因型-环境关联分析

对高山栎组植物各谱系所有 SNPs 位点进行冗余分析结果表明（表 4.2），四个谱系的遗传变异与地理和气候变量均呈现显著相关，气候变量在谱系 A、B、C 和 D 中分别解释了总变异的 16%，6%，17%和 18%；地理变量分别解释了总变异的 9%，3%，7%和 3%；气候和地理变量共同解释了谱系 A、B、C 和 D 中总变异的 17%，7%，18%和 19%。本研究运用偏冗余分析进一步评估地理和气候变量的相对贡献时，在控制地理变量后结果显示，谱系 A、B 和 C 的遗传变异与气候变量显著相关，谱系 D 的遗传变异与气候变量无显著相关性；在控制气候变量后结果显示，谱系 A 和 C 的遗传变异与地理变量显著相关，谱系 B 和 D 的遗传变异与地理变量无显著相关性（表 4.2）。

对各谱系推定受选择的 SNPs 位点进行冗余分析结果表明（表 4.2），四个谱系的遗传变异与地理和气候变量均呈现显著相关，气候变量在谱系 A、B、C 和 D 中分别解释了总变异的 50%，19%，56%和 63%；地理变量分别解释了总变异的 31%，11%，31%和 47%；气候和地理变量共同解释了谱系 A、B、C 和 D 中总变异的 52%，19%，61%和 65%。运用偏冗余分析进一步评估地理和气候变量的相对贡献时，在控制地理变量后结果显示，各谱系的遗传变异与气候变量显著相关；在控制气候变量后结果显示，谱系 A 和 C 的遗传变异与地理变量显著相关，谱系 B 和 D 的遗传变异与地理变量无显著相关性（表 4.2）。

对各谱系与气候相关的 SNPs 位点进行冗余分析结果表明（表 4.2），四个谱系的遗传变异与地理和气候变量均呈现显著相关，气候变量在谱系 A、B、C 和 D 中分别解释了总变异的 41%，24%，55%和 45%；地理变量分别解释了总变异的 21%，13%，26%和 29%；气候和地理变量共同解释了谱系 A、B、C 和 D 中总变异的 42%，25%，57%和 48%。运用偏冗余分析进一步评估地理和气候变量的相对贡献时，在控制地理变量后结果显示，各谱系的遗传变异与气候变量显著相关；在控制气候变量后结果显示，谱系 A、B 和 C 的遗传变异与地理变量显著相关，谱系 D 的遗传变异与地理变量无显著相关性（表 4.2）。

对于各谱系的所有 SNPs 位点、推定受选择的 SNPs 位点和气候相关的 SNPs 位点进行的冗余分析和偏冗余分析）结果综合显示，地理与气候变量的共同解释度均高于单独的气候或者地理变量。

表 4.2 高山栎组各谱系的所有 SNPs、推定受选择和与气候相关的 SNPs 的冗余分析和偏冗余分析结果

Table 4.2 Results of RDA and *p*RDA for all SNPs、putative selected SNPs and climate-associated SNPs among different lineages of *Quercus sect. Heterobalanus*

		所有 SNPs (All SNPs)			推定受选择 SNPs (Putative selected SNPs)			气候相关的 SNPs (Climate-associated SNPs)		
		R ²	adjR ²	<i>P</i>	R ²	adjR ²	<i>P</i>	R ²	adjR ²	<i>P</i>
lineage A	RDA									
	G-Env	0.31	0.16	0.001	0.59	0.50	0.001	0.52	0.41	0.001
	G-Geo	0.14	0.09	0.001	0.34	0.31	0.001	0.25	0.21	0.001
	G-Env+Geo	0.37	0.17	0.001	0.64	0.52	0.001	0.56	0.42	0.001
	<i>p</i>RDA									
	G-Env Geo	0.23	0.08	0.001	0.29	0.22	0.001	0.31	0.21	0.001
lineage B	G-Geo Env	0.06	0.02	0.001	0.04	0.02	0.001	0.04	0.02	0.002
	RDA									
	G-Env	0.24	0.06	0.001	0.35	0.19	0.001	0.39	0.24	0.001
	G-Geo	0.08	0.03	0.002	0.16	0.11	0.001	0.18	0.13	0.001
	G-Env+Geo	0.30	0.07	0.001	0.39	0.19	0.001	0.44	0.25	0.001
	<i>p</i>RDA									
G-Env Geo	0.22	0.04	0.001	0.23	0.08	0.008	0.26	0.12	0.001	
lineage C	G-Geo Env	0.06	0.01	0.078	0.04	0.00	0.421	0.05	0.01	0.033
	RDA									
	G-Env	0.45	0.17	0.001	0.70	0.56	0.002	0.70	0.55	0.001
	G-Geo	0.19	0.07	0.003	0.40	0.31	0.005	0.36	0.26	0.001
	G-Env+Geo	0.52	0.18	0.001	0.70	0.56	0.001	0.70	0.55	0.001
	<i>p</i>RDA									
G-Env Geo	0.25	0.10	0.004	0.30	0.25	0.002	0.34	0.29	0.001	
lineage D	G-Geo Env	0.11	0.07	0.01	0.42	0.46	0.001	0.22	0.20	0.001
	RDA									
	G-Env	0.54	0.18	0.001	0.79	0.63	0.001	0.69	0.45	0.001
	G-Geo	0.24	0.13	0.001	0.53	0.47	0.001	0.38	0.29	0.001
	G-Env+Geo	0.65	0.19	0.002	0.85	0.65	0.001	0.77	0.48	0.001

续表

	所有 SNPs (All SNPs)			推定受选择 SNPs (Putative selected SNPs)			气候相关的 SNPs (Climated-associated SNPs)		
	R ²	adjR ²	P	R ²	adjR ²	P	R ²	adjR ²	P
<i>p</i>RDAs									
G-Env Geo	0.41	0.06	0.069	0.32	0.19	0.001	0.39	0.19	0.004
G-Geo Env	0.10	0.01	0.45	0.05	0.02	0.186	0.08	0.03	0.18

注: G: 等位基因频率的因变量矩阵; Geo: 地理变量; Env: 环境变量; R²: 解释率; adjR²: 校正解释率; * *P*-value < 0.05; ** *P*-value < 0.01; *** *P*-value < 0.001

Note: G: Dependent matrix of population allele frequencies; Geo: geography variables; Env: environment variables; R²: coefficient of determination; adjR²: adjusted coefficient of determination; * *P*-value < 0.05; ** *P*-value < 0.01; *** *P*-value < 0.001

为进一步确定影响适应性遗传变异的两个数据集—推定的受选择 SNPs 位点和与气候相关的 SNPs 位点的主要环境因素，本研究采用梯度森林对环境变量的重要性进行排序结果如下：

在谱系 A 的推定的受选择 SNPs 位点中，梯度森林分析显示太阳辐射（Srad）是影响适应性遗传变异最重要的气候驱动因子，其次是风速（Wind），最冷月份最低温度差（Bio07）和降水量季节性变化（Bio15）（图 4.4-a）；与气候相关的 SNPs 位点的结果与推定的受选择 SNPs 位点一致（图 4.4-b）。两个数据集的等位基因转换沿环境梯度累计的重要性结果表明太阳辐射在 14000 - 15000 kWh/m²/yr 的区间发生了显著变化，风速在 2.0 - 2.5 m/s 的区间发生了显著变化，最冷月份最低温度差在 2.6 °C 左右发生了变化，降水量季节性变化在 80 - 90 mm 之间发生了明显的变化（图 4.4-c）。

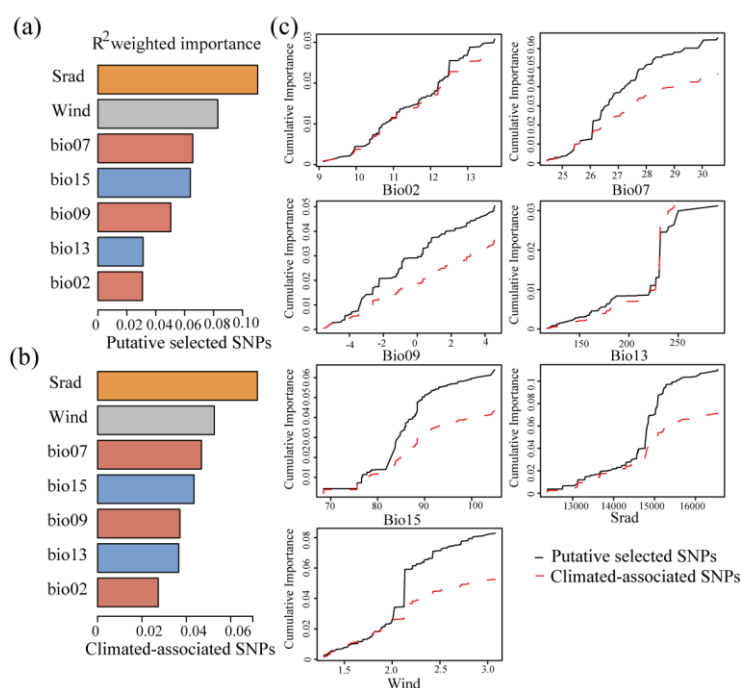


图 4.4 基于梯度森林模型对谱系 A 中推定受选择的 SNPs 和与气候相关的 SNPs 的解释变量进行比较计算环境梯度的累积重要性。(a) 推定受选择的 SNPs 和 (b) 与气候相关的 SNPs 的解释变量的 R² 加权重要性；(c) 推定受选择的 SNPs（黑色实线）和与气候相关的 SNPs（红色虚线）模型，沿着环境梯度的等位基因转换的累积重要性

Figure 4.4 Comparison of explanatory variables for putative selected SNPs and climate-associated SNPs of lineage A and cumulative importance of environmental gradients based on Gradient Forest analysis. The R² weighted importance of explanatory variables ranked for both (a) putative selected SNPs and (b) climate-associated SNPs, (c) Cumulative importance of allelic turnover along environmental gradients for both putative selected SNPs (black solid line) and climate-associated SNPs (red dashed line) models

在谱系 B 的推定的受选择 SNPs 位点中，梯度森林分析结果显示太阳辐射

(Srad) 是影响遗传变异最重要的气候驱动因子，其次是降水量季节性变化 (Bio15)，风速 (Wind) 和最干季度平均温度 (Bio09)；与气候相关的 SNP 位点的结果显示太阳辐射 (Srad) 是影响遗传变异最重要的气候驱动因子，其次是降水量季节性变化 (Bio15)，风速 (Wind) 和最冷月份最低温度差 (Bio07)。两个数据集的等位基因转换沿环境梯度累计的重要性结果共同表明太阳辐射在 13000 - 15000 kWh/m²/yr 的区间发生了显著变化，季节性降水量的变化较为平缓，风速在 2.5 - 3.0 m/s 的区间发生了显著变化，最冷月份最低温度差在 24 - 26 °C 间发生了变化。其次在推定的受选择 SNPs 位点中，最干季度平均温度在 -4 - 2 °C 间发生了明显的变化，与气候相关的 SNPs 位点中呈现稳定的趋势变化。

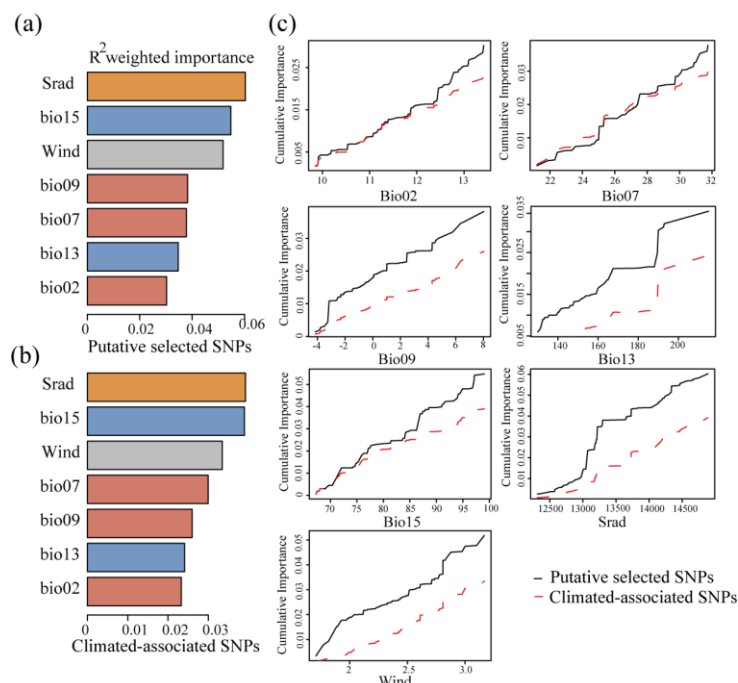


图 4.5 基于梯度森林模型对谱系 C 中推定受选择的 SNPs 和与气候相关的 SNPs 的解释变量进行比较环境梯度的累积重要性。(a) 推定受选择的 SNPs 和 (b) 与气候相关的 SNPs 的解释变量的 R² 加权重要性；(c) 推定受选择的 SNPs (黑色实线) 和与气候相关的 SNPs (红色虚线) 模型，沿着环境梯度的等位基因转换的累积重要性

Figure 4.5 Comparison of explanatory variables for putative selected SNPs and climated-associated SNPs of lineage B and cumulative importance of environmental gradients based on Gradient Forest analysis. The R² weighted importance of explanatory variables ranked for both (a) putative selected SNPs and (b) climated-associated SNPs, (c) Cumulative importance of allelic turnover along environmental gradients for both putative selected SNPs (black solid line) and climated-associated SNPs (red dashed line) models

在谱系 C 的推定的受选择 SNPs 位点中，梯度森林分析结果显示风速 (Wind) 是影响遗传变异最重要的气候驱动因子，其次是太阳辐射 (Srad)、

最湿月份降水量 (Bio13) 和最冷月份最低温度差 (Bio07) (图 4.6-a); 与气候相关的 SNPs 位点的结果显示最干季度均温 (Bio09) 是影响遗传变异最重要的气候驱动因子, 其次是最冷月份最低温度差 (Bio07), 平均气温日较差 (Bio02) 和降水量季节性变化 (Bio15) (图 4.6-b)。两个数据集的等位基因转换沿环境梯度累计的重要性结果共同表明昼夜温差月均值在 10.5 - 11 °C 间发生了显著的变化, 最冷月份最低温度差在 24 °C 附近发生了明显变化, 最干季度平均温度在 5-6 °C 间发生了剧烈变化, 降水量季节性变化在 75 - 80 mm 间发生明显变化, 风速在 2.1 - 2.2 m/s 间发生了变化。其次在推定的受选择 SNPs 位点中, 太阳辐射在 13600 - 14000 kWh/m²/yr 间发生了剧烈变化, 最湿月份降水量在 130 mm 左右发生了明显变化; 与气候相关的 SNPs 位点中, 太阳辐射在 14000 - 14200 kWh/m²/yr 间发生了剧烈变化, 最湿月份降水量在 150-160 mm 左右发生了明显变化 (图 4.6-c)。

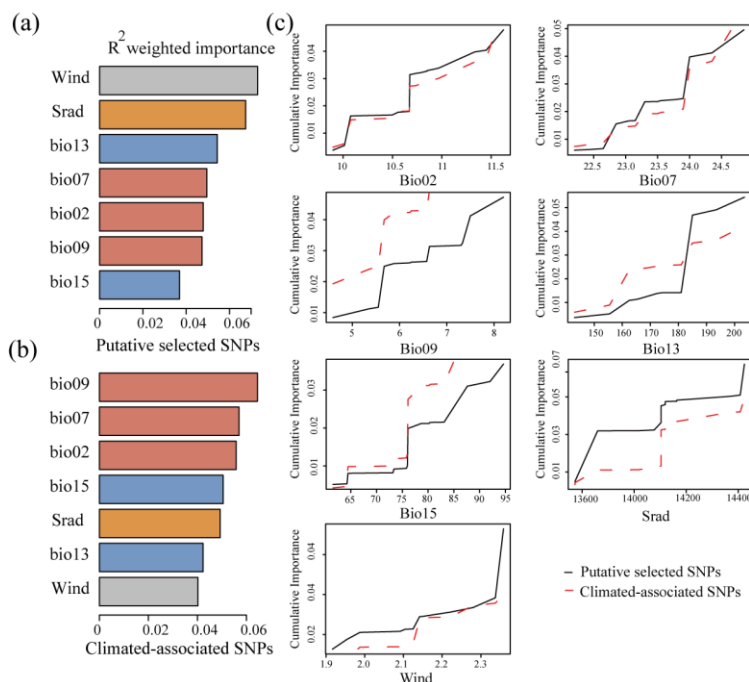


图 4.6 基于梯度森林模型对谱系 C 中推定受选择的 SNPs 和与气候相关的 SNPs 的解释变量进行比较计算环境梯度的累积重要性。(a) 推定受选择的 SNPs 和 (b) 与气候相关的 SNPs 的解释变量的 R² 加权重要性; (c) 推定受选择的 SNPs (黑色实线) 和与气候相关的 SNPs (红色虚线) 模型, 沿着环境梯度的等位基因转换的累积重要性

Figure 4.6 Comparison of explanatory variables for putative selected SNPs and climate-associated SNPs of lineage C and cumulative importance of environmental gradients based on Gradient Forest analysis. The R² weighted importance of explanatory variables ranked for both (a) putative selected SNPs and (b) climate-associated SNPs, (c) Cumulative importance of allelic turnover along environmental gradients for both putative selected SNPs (black solid line) and climate-associated SNPs (red dashed line) models

在谱系 D 的推定的受选择 SNPs 位点中，梯度森林分析结果显示太阳辐射（Srad）是影响遗传变异最重要的气候驱动因子，其次是风速（Wind）、降水量季节性变化（Bio15）和平均气温日较差（Bio02）（图 4.7-a）；与气候相关的 SNPs 位点的结果显示降水量季节性变化（Bio15）是影响遗传变异最重要的气候驱动因子，其次是最干季度平均温度（Bio09），最湿月份降水量（Bio13）和平均气温日较差（Bio02）（图 4.7-b）。在推定的受选择 SNPs 位点中，平均气温日较差（Bio02）在 9 - 10 °C 间发生了显著的变化，最湿月份降水量在 150 mm 左右发生了明显变化，降水量季节性变化在 75 - 80 mm 间发生明显变化，太阳辐射在 12000 - 12500 kWh/m²/yr 间发生了剧烈变化，风速在 1.8 - 2.0 m/s 间发生了明显变化，最冷月份最低温度差和最干季度平均温度未发现陡峭变化。与气候相关的 SNPs 位点中显示所有变量的发生平稳变化（图 4.7-c）。

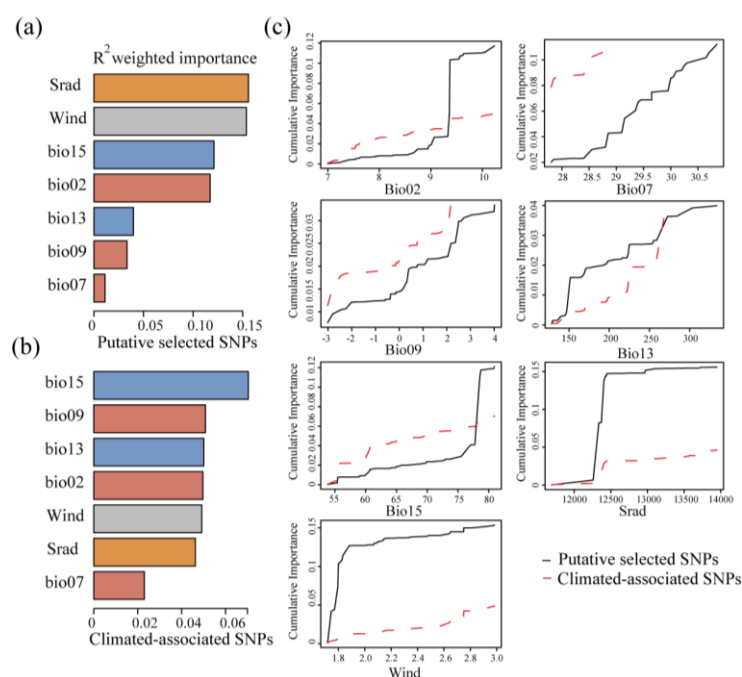


图 4.7 基于梯度森林模型对谱系 D 中推定受选择的 SNPs 和与气候相关的 SNPs 的解释变量进行比较环境梯度的累积重要性。(a) 推定受选择的 SNPs 和 (b) 与气候相关的 SNPs 的解释变量的 R² 加权重要性；(c) 推定受选择的 SNPs（黑色实线）和与气候相关的 SNPs（红色虚线）模型，沿着环境梯度的等位基因转换的累积重要性

Figure 4.7 Comparison of explanatory variables for putative selected SNPs and climate-associated SNPs of lineage D and cumulative importance of environmental gradients based on Gradient Forest analysis. (a) The R² weighted importance of explanatory variables ranked for both (a) putative selected SNPs and (b) climate-associated SNPs, (c) Cumulative importance of allelic turnover along environmental gradients for both putative selected SNPs (black solid line) and climate-associated SNPs (red dashed line) models

4.3.3 GO 富集与顺式调控元件

基于筛选出的推定受选择的 SNPs 进行基因本体 (GO) 富集分析结果表明 (图 4.7), 此处仅表述生物学过程类别中显著性最高的前 20 个条目。在谱系 A 的推定受选择的 SNPs 共注释到 1192 个基因, 获得 1127 个显著性匹配的拟南芥同源基因, GO 分析结果显示主要富集于与细胞生物发育与器官建成 (如叶性器官发育-*phyllome development*, 生殖结构发育-*reproductive structure development*)、代谢与细胞调控 (如代谢过程-*metabolic process*) 和非生物刺激响应 (*response to abiotic stimulus*) 等过程 (图 4.7-a)。在谱系 B 的推定受选择的 SNPs 共注释到 372 个基因, 获得 355 个显著性匹配的拟南芥同源基因, GO 富集分析结果显示主要富集于与光响应 (如辐射响应-*response to radiation*, 光刺激响应-*response to light stimulus*)、发育调控 (如胚胎后发育-*post-embryonic development*)、物质运输 (如铁离子转运-*iron ion transport*)、代谢过程 (如初级代谢过程-*primary metabolic process*) 及非生物刺激响应 (*response to abiotic stimulus*) 等过程 (图 4.7-b)。在谱系 C 的推定受选择的 SNPs 共注释到 257 个基因, 获得 247 个显著性匹配的拟南芥同源基因, GO 富集分析结果显示其显著富集于光合作用 (如光合作用-*photosynthesis*, 光合电子传递链-*photosynthetic electron transport chain*)、代谢与调控 (如木质素代谢过程-*lignin metabolic process*, RNA 代谢过程的积极调控-*positive regulation of RNA metabolic process*)、物质运输 (如无机阳离子跨膜转运-*inorganic cation transmembrane transport*) 以及刺激响应 (如刺激响应-*response to stimulus*、光刺激响应-*response to light stimulus*) 等过程 (图 4.7-c)。在谱系 D 的推定受选择的 SNPs 总共注释到 271 个基因, 获得 259 个显著性匹配的拟南芥同源基因, GO 富集分析结果显示其显著富集于代谢过程 (如脂质代谢过程-*lipid metabolic process*, 含氧酸代谢过程-*oxoacid metabolic process*)、刺激响应 (如刺激响应-*response to stimulus*、辐射响应-*response to radiation*) 和非生物刺激响应 (*response to abiotic stimulus*) 等过程 (图 4.7-d)。

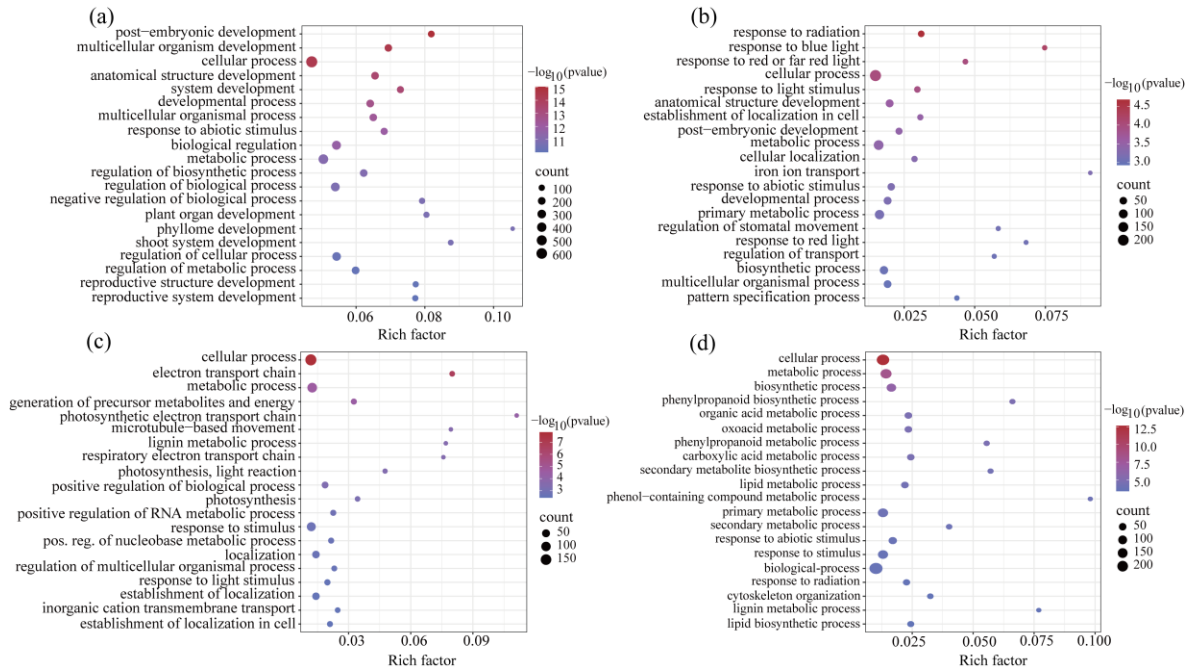


图 4.7 高山栎组植物各谱系中推定受选择 SNP 在生物过程中前 20 名的 GO 富集气泡图。(a) 谱系 A; (b) 谱系 B; (c) 谱系 C; (d) 谱系 D。气泡的大小表示每个 GO 术语中富集的基因数量，而气泡的颜色则表示富集的显著程度，仅显示了显著富集的生物过程

Figure 4.7 Top 20 of Gene Ontology (GO) enrichment analysis of putative selected SNPs in different lineages of *Quercus* sect. *Heterobalanus* in biological process. (a) lineage A, (b) lineage B, (c) lineage C, (d) lineage D. Bubble size represents the number of genes enriched in each GO term, while bubble color indicates the significance level of enrichment. Only significantly enriched biological processes are shown

基于推定受选择的非编码区上下游区域的 SNPs 的序列进行匹配的顺式调控元件基序结果表明（图 4.8），在谱系 A 的推定受选择位点的转录本 SNPs 中，554 个（40%）SNPs 位于基因上下游区域，73%（405 个）SNPs 位于已知植物顺式调控 DNA 元件的基序中，其中 34%的 SNPs 与温度、水分、光、氧气、叶片、花粉和种子发育调控相关联（图 4.8-a）。在谱系 B 的推定受选择位点的转录本 SNPs 中，140 个（35%）SNPs 位于基因上下游区域，84%（118 个）SNPs 位于已知植物顺式调控 DNA 元件的基序中，其中 38%的 SNPs 与温度、水分、光、氧气、叶片、花粉和种子发育调控相关联（图 4.8-b）。在谱系 C 的推定受选择位点的转录本 SNPs 中，99 个（36%）SNPs 位于基因上下游区域，61%（60 个）SNPs 位于已知植物顺式调控 DNA 元件的基序中，其中 14%的 SNP 与温度、水分、光、叶片、花粉和种子发育调控相关联（图 4.8-c）。在谱系 D 的推定受选择位点的转录本 SNP 中，111 个（39%）SNP 位于基因上下游区域，72%（81 个）SNP 位于已知植物顺式调控 DNA 元件的基序中，其中 42%的 SNP 与温度、水分、光、氧气、叶片、花粉和种子发育调控相关联（图 4.8-d）。

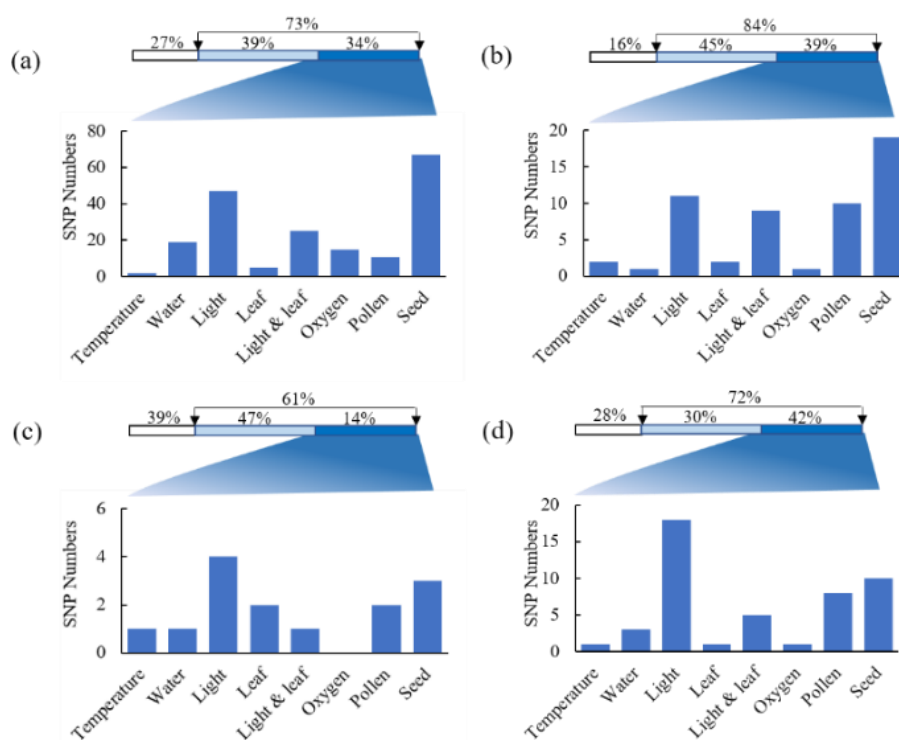


图 4.8 预测高山栎组植物各谱系位于基因调控区域内的顺式作用调控元件。(a) 谱系 A; (b) 谱系 B; (c) 谱系 C; (d) 谱系 D。蓝色条形图表示在已知植物顺式调控 DNA 元件的模式中发现的单核苷酸多态性的数量

Figure 4.8 Predicted functional regulatory elements of putative selected SNPs located within regulatory regions. (a) lineage A, (b) lineage B, (c) lineage C, (d) lineage D. Blue bars represent the number of SNPs found in motifs of known plant cis-regulatory DNA elements

4.3.4 适应性候选基因

通过对每个谱系进行推定受选择位点涉及的基因和正选择高分化域的基因重叠分析结果表明 (图 4.9), 在谱系 A 中发现了 70 个具有适应性分化的拟南芥同源基因, 这些基因总共涉及 91 个的推定受选择 SNPs 位点, 每条染色体上均有分布, 其中 2 和 9 号染色体上分布最少, 只有一个 SNP 位点 (图 4.9-a)。鉴定出一系列在谱系 A 在分化过程中的适应性关键候选基因主要涉及生长发育调控, 如调节植株大小或叶片发育以及调控花发育, *GA5*、*LAC4*、*DPA4*、*KIR1* 等基因; 环境胁迫响应, 如冷胁迫、干旱胁迫和盐胁迫, *CNGC20*、*FREE1*、*GPA1* 等基因; 防御与抗氧化调控, 如氧化胁迫, *AHL1*、*GPX1* 等基因 (附表 5)。

在谱系 B 中发现了 29 个具有适应性分化的拟南芥同源基因, 位于这些基因上的总共涉及 36 个推定受选择 SNPs 位点, 位于 1、3、5、8、9、10、11 和 12 号染色体 (图 4.9-b)。鉴定出的拟南芥同源基因主要涉及生长发育调控, 如叶缘发育和花粉萌发, *DPA4*、*ATHB1* 和 *PTB1* 等基因; 环境胁迫响应, 如冷胁迫、热胁迫、U-VB 胁迫、盐胁迫和干旱胁迫, *CNGC20*、*NBR1*、*ALDH2C4*、*HDA15*、*ACX1* 等基因;

DAN 损伤修复 (*SNII*) 与抗病性调控 (*NF-YB3*) (附表 5)。

在谱系 C 中发现了 6 个具有适应性分化的拟南芥同源基因，位于这些基因上的总共涉及 6 个推定受选择 SNPs 位点，位于 1、4、5、10 和 12 号染色体 (图 4.9-c)。鉴定出的拟南芥同源基因涉及响应干旱胁迫 (*ACXI*) 和生长发育调控的植物生殖器官 (花药) 发育过程中的关键形态建成 (*NSTI*) (附表 5)。

在谱系 D 中发现了 16 个具有适应性分化的拟南芥同源基因，位于这些基因上的总共涉及 18 个推定受选择 SNPs 位点，位于 1、3、5、6、7、9、11 和 12 号染色体 (图 4.9-d)。鉴定出的拟南芥同源基因涉及生长发育调控，如花粉/花药发育和叶片衰老调控，*NDR5*、*LHT4* 和 *GVS1*；环境胁迫响应，如低氧胁迫、水分胁迫、盐胁迫和干旱胁迫，*DGK5*、*AtHSPR*、*ALDH2C4* 等基因 (附表 5)。

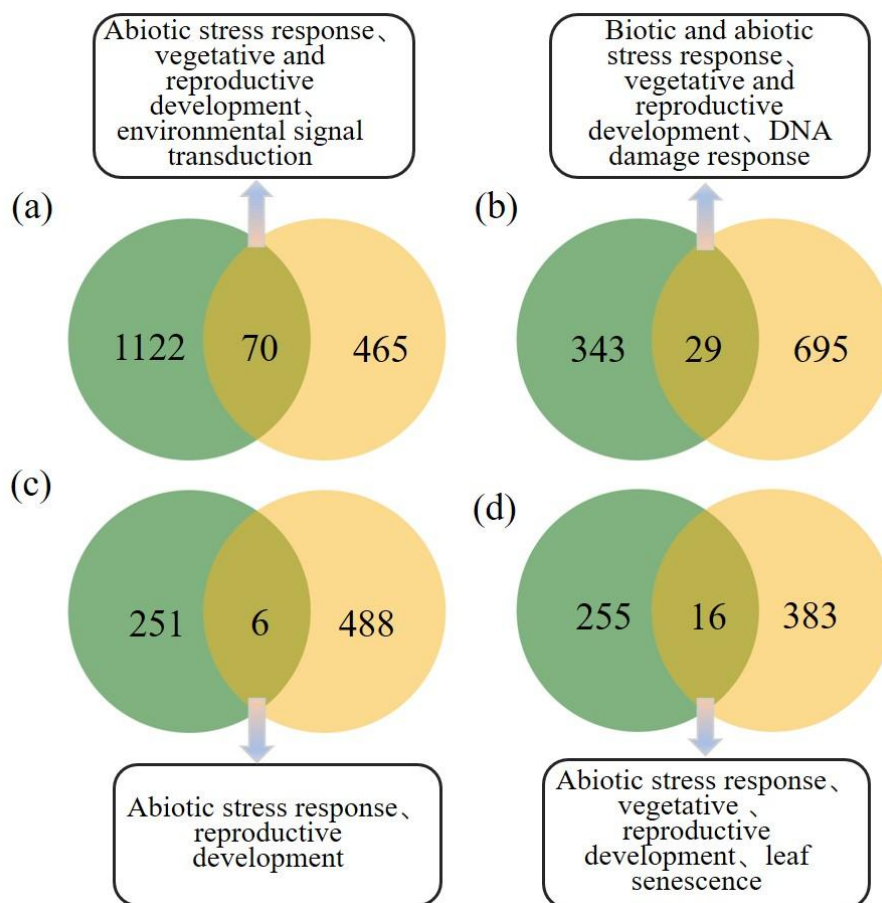


图 4.9 高山栎组植物各谱系的推定受选择位点和正选择高分化域的拟南芥同源基因数量的韦恩图。(a) 谱系 A；(b) 谱系 B；(c) 谱系 C；(d) 谱系 D。绿色和黄色分别代表位于推定受选择 SNPs 位点与正选择高分化区域内的基因数量

Figure 4.9 The number of *Arabidopsis* homologous genes between putative selected SNP and positive selected highly diverged regions for different lineages of *Quercus* sect. *Heterobalanus*. (a) lineage A, (b) lineage B, (c) lineage C, (d) lineage D. Green and yellow color represent the gene in putative selected SNPs and positive selected highly diverged regions, respectively

4.4 讨论

4.4.1 环境因子驱动适应性遗传分化

全球气候变化正在对生物多样性造成严重的威胁，面对不断变化的环境条件，适应性遗传变异可能会提高种群的适应能力和进化潜力^[274]。通过对高山栎组植物的不同谱系进行基于遗传分化和基因-环境关联分析方法识别出了大量的受选择位点，包括可能与生物气候变量相关的位点，这些位点可能在局部环境适应中发挥作用^[275]。在山地生态系统中，地理屏障和环境异质性通常被认为是影响遗传分化以及局地适应的重要因素^[276]。本研究对每个谱系的全部 SNPs、推定受选择 SNPs 以及气候相关 SNPs 三个数据集进行冗余分析结果显示，地理变量与气候变量共同驱动高山栎组植物各谱系的遗传变异，且二者的共同解释率普遍高于单一变量的解释率，表明地理隔离与气候偏好在各谱系的适应性演化过程中存在协同作用。偏冗余分析结果表明，在全部 SNPs 水平上控制地理因素后，位于青藏高原地区的谱系 A、B 和 C 的遗传变异仍与气候变量显著相关，表明高原上的极端环境对谱系的局地适应有重要影响。相反，谱系 D 的遗传变异与气候变量并无显著相关性，这一结果可能表明该谱系的遗传变异主要受到历史种群动态的影响，而非当前气候条件的直接选择作用。正如 Chen 等^[124]研究发现高山栎组植物在适应青藏高原地区的地质历史和气候振荡的作用下，随着横断山脉的抬升和全球气候的冷却，横断山脉的高海拔地区经历了越来越强烈的紫外线辐射，变得越来越凉爽和干燥，具备耐寒和耐旱特性的高山栎组植物则在多变的环境成功进行了扩张，造成了现在分布于秦巴山脉至台湾地区的分布范围。对推定受选择位点和气候相关位点进行偏冗余分析结果显示，四个谱系的遗传变异均与气候变量显著相关，说明气候环境变化可能是驱动高山栎组植物各谱系适应性遗传变异的重要因素。类似的研究在多种高山植物中也有研究发现，如高山松和青海云杉 (*Picea crassifolia* Kom.) 等物种的景观基因组学研究表明气候变量能够显著解释适应性遗传变异，反映出高山环境中强烈的局地选择压力^[200, 277]。

在青藏高原及其邻近的高海拔地区，复杂的地形和气候条件变化会对高山植物产生不同的选择压力，从而促进遗传分化或谱系分化，被认为是驱动高山植物多样性形成的重要因素^[66, 217]。梯度森林分析结果显示，谱系 A 和谱系 B 的推定受选择位点和气候相关位点均以太阳辐射为主要驱动因子，表明在高山环境中强烈的太阳辐射可能是驱动这两个谱系适应性分化的重要环境因子。已有研究证明太阳辐射是影响遗传变异和适应性的重要环境因子^[278, 279]，而在青藏高原及其邻近地区，研究发现强烈太阳辐射不仅影响植物的光合效率，还会通过紫外胁迫、蒸散作用以及温度调节等多种生理过程对植物产生选择压力，从而促进植物的局地适应和遗传分化^[280]。相比之下，谱系 C 的推定受选择位点主要受风速影响，而气候相关位点则主要与最干季度平均温度 (Bio09) 相关，表明该谱系的局地适应可能受到多种气候因子的共同作用。对于谱系 D 而言，推定受选择位点以太阳辐射为主要驱动因子，而气候相

关键位点则主要受降水量季节性变化（Bio15）的影响。该谱系所在地区作为中国南北气候和植被分界的重要区域，其明显的季风气候特征以及降水季节性变化显著，对该谱系的植物的生长周期、水分利用效率以及生理适应都具有重要影响^[66]。因此，不同遗传谱系的局地适应过程中所受的主要环境选择压力存在明显差异，与各谱系所处的地理环境和气候条件密切相关。

4.4.2 局地适应的遗传基础

通过对各谱系进行适应性位点与正选择高分化区间位点的基因重叠分析，本研究鉴定出一系列具有适应性分化信号的拟南芥同源基因，且四个谱系在适应性候选基因的数量及功能类别上表现出明显差异，表明高山栎组植物在面对青藏高原及其邻近地区不同的环境选择压力时，各遗传谱系产生了不同的适应策略。但是在各谱系中发现了具有相同功能的基因，如抗旱基因等。抗旱相关基因在当前生态背景下可能成为物种间的潜在屏障，但在特定环境条件下，这些基因也可能通过杂交和渐渗过程促进物种的适应性进化^[212, 281, 282]。本研究在高山栎组植物中的各谱系中均发现了与抗旱性相关的适应性候选基因，如谱系 A 中的 *GPA1*、*GTL1*、*ARPN*，谱系 B 中的 *ACX1*、*ALDH2C4*，谱系 C 中的 *ACX1* 和谱系中的 *ALDH2C4*，可能表明干旱是高山栎组植物适应性进化面对的共同的重要环境压力。一方面，之前研究通过对高山栎组植物进行形态特征解剖发现在叶表皮特征（如角质膜、气孔数量等）与旱生植物具有较大的相似性，并且同一种植物对不同环境的长期适应过程中，表皮的形态解剖特征表现出趋异现象，为其抗旱适应已经表现出了形态差异^[283]。另一方面，趋同进化、不同祖先的多态保留和杂交物种形成等的不同过程均可能会影响耐旱适应的进化^[284]。Leroy 等^[93]对四个欧洲白栎组的物种一无梗花栎（*Q. petraea*），夏栎（*Q. robur*），比利牛斯山栎（*Q. pyrenaica* Willd.），柔毛栎（*Q. pubescens* Willd.）研究发现其在经历了长期的地理隔离后，于冰期结束后发生了广泛的二次接触与基因流，并进一步揭示生态适应相关分化的干旱基因与内在生殖隔离机制协同驱动了栎属植物的物种形成过程。在栎属植物中，适应性相关基因不仅参与局地适应过程，还可能在物种间通过杂交与渐渗的过程促进适应性基因的扩散，并可能进一步促进了栎属植物的多样性。因此，本研究在高山栎组植物不同谱系的局地适应研究结果中的抗旱基因可能成为谱系间基因交流与适应环境变化（如干旱或高海拔环境）的重要遗传基础，能够反映出这些谱系在适应各自生境水分条件过程中经历的差异化选择压力。

通过鉴定到的适应性基因也进一步支持了花粉/花药功能分化可能作为重要的内在生殖隔离机制，促进谱系间合子前或合子后隔离，从而加快了遗传分化与物种形成进程。在各谱系中均鉴定出参与花粉/花药发育的候选基因（如谱系 A 中的 *KIR1*、谱系 B 中的 *PTB1*、谱系 C 中的 *NST1*、谱系 D 中的 *NDR5* 和 *LHT4*）。尽管各谱系受到选择的具体基因位点存在差异，这可能表明花粉发育相关基因的分化可能是高山

栎组植物不同谱系在长期地理隔离与局地适应过程的趋异进化结果，并可能作为内在生殖隔离机制，促进谱系间的合子前或合子后隔离，进一步加快遗传分化和物种形成^[284]。同时在研究中会发现某些基因可能具有多效性效应，如同时参与胁迫响应和生长发育调控的 *DPA4*，可能表明在高山栎组植物中某些基因可能在生态适应和物种形成中存在协同作用，共同促进谱系分化和物种形成。

值得注意的是，栎属高山栎组植物存在的杂交和渐渗现象^[192]。已有研究发现川滇高山栎与其近缘种在不同海拔梯度上的基因流呈现不对称性，一些与环境适应相关的基因区域在物种间更容易发生渐渗，并且在生殖隔离或核心发育调控的基因区域则表现出较强的分化维持信号^[101]。这种适应性渐渗模式可能在复杂山地环境中既可能削弱物种边界，也可能在生态选择的作用下强化各谱系的局地适应。而且对于分布在青藏高原及其邻近地区的高山栎植物而言，青藏高原的北向隆升和南向推移以及季风气候的发生，导致的亚洲内陆干旱化是高山栎组植物各谱系在演化过程中共同面临的环境选择压力，不同谱系通过选择不同的抗旱基因实现对异质生境的局地适应^[285, 286]。因此，与抗旱基因和生殖相关基因可能共同影响了高山栎组植物谱系间局地适应和潜在内在生殖隔离的遗传基础，并可能影响谱系未来的分布范围和适应性进化潜力。最后，进化单元的划分、与关键胁迫响应（如干旱、温度变化）相关的适应性基因位点及其地理分布特征，不仅可以了解不同种群或谱系的局地适应能力，还能进一步用于指导种源选择与混合配置策略。以分布于青藏高原及其邻近地区的高山栎组植物为例，在生态修复实践中可以优先选择在干旱胁迫相关基因上具有优势等位基因变异，且与目标修复区域环境匹配度更高的遗传谱系作为基础种群，同时在一定范围内引入存在适应性渐渗的近缘谱系，以增加群落的遗传多样性和提高对环境的适应能力。

4.5 小结

本章节通过对高山栎组植物的不同谱系的局地适应开展景观基因组学分析。运用遗传分化和基因型-环境关联分析识别异常位点检测的结果显示推定的受选择 SNPs 数量在不同谱系间存在差异性分布，谱系 A 的 SNP 数量最多，谱系 C 的最少。冗余分析结果表明，地理与气候变量对遗传变异的解释度相对较低；而在推定受选择 SNPs 和气候相关 SNPs 数据集中，环境变量的解释度得到了提高，且气候变量对遗传变异的影响大于地理变量。偏冗余分析显示谱系 A 和 C 同时受到地理与气候因素影响，而谱系 B 和 D 的适应性遗传变异更多受到气候因素驱动。梯度森林分析揭示太阳辐射是驱动适应性遗传变异的重要环境因子，同时不同谱系在温度、降水及风速变量上表现出不一样的重要性排序，体现了各谱系间对环境变化的响应存在特异性。功能注释表明高山栎组植物不同谱系在生态适应与生殖隔离方面存在谱系特异性策略，在干旱、高海拔及异质环境中，通过选择不同抗旱基因和发育调控基因实现对局地环境的适应，同时多效性基因可能将生态适应与内在生殖隔离相联系，从而促进谱系分化，使其能从生物多样性热点区域有效扩散至多样化环境中。此外，顺式调控元件分析显示谱系中 14% - 42% 位于基因上下游区域的推定受选择位点位于调控区域，均与光、温度和水分响应相关，揭示了基因表达调控可能在适应性分化中发挥重要作用。

5 结论与展望

5.1 主要结论

栎属高山栎组植物是青藏高原地区及其邻近地区的重要建群种，对于维持这一生物多样性热点地区的生态系统平衡和长期稳定有重要意义。结合遗传数据和环境变量对高山栎组植物进行遗传分化、适应性进化和环境适应性研究依旧有待深入。本研究综合群体基因组学和景观基因组学的方法深入解析高山栎组植物的遗传结构和局地适应机制研究，为探究高山栎组植物的物种界限和适应性进化的遗传基础提供新的理解，以及为理解栎属植物的生态分化机制及环境响应的分子调控提供理论依据。主要结论如下：

(1) 高山栎组植物的七个种的全基因组重测序数据进行遗传结构、遗传多样性以及进化历史分析结果表明，高山栎组植物存在形态分化与遗传分化不一致的现象，划分为四个谱系（A~D）且呈现按照地理分布聚类的情况，其中位于横断山脉地区的谱系 B 和 C 的遗传多样性水平最高，且位于西藏地区的谱系 A 与谱系 B 的遗传分化程度最低，并发现了谱系 A 流向谱系 B 的明显基因流。种群动态历史表明，高山栎组植物的四个谱系在进化历史上呈现反复扩张和收缩的趋势。蒙特尔检验结果显示，高山栎组植物具有显著的地理隔离和环境隔离效应，且地理隔离对遗传变异的相关性高于环境隔离的相关性。生态位重叠分析表明四个谱系间重叠度低，出现生态位分化趋势。以上研究结果表明，更新世以来的气候波动与复杂山地地理格局共同塑造了高山栎组植物的谱系分化与遗传结构格局，其中地理隔离在遗传分化过程中发挥主导作用，而环境异质性则进一步推动了谱系间的生态位分化与适应性演化。

(2) 运用基因组扫描的 F_{ST} 和跨群体复合似然比分析方法分别鉴定不同谱系间的高分化区域与选择性清除信号结果表明，主要位于西藏地区的谱系 A 与秦巴山-华东地区的谱系 D 的正选择高分化区域最少，表明分化时间较早且长期隔离。这一结果与遗传结构体现的先分化出的谱系 D 一致表明，高山栎组植物符合了“走出青藏高原”假说。在 6 个谱系对的候选基因组岛中共鉴定出 864 个具有显著性匹配的同源基因标识符的受选择候选基因，其中 195（22.6%）个基因共享于 2-4 个不同谱系对中，并发现具有胁迫响应的 *ERDJ3B*，*PGX3* 和 *KOIN* 等同源基因，主要涉及温度和盐胁迫的响应。这些与环境适应相关的基因可能会引起近缘物种形成遗传分化或者谱系分化，对近缘种之间的遗传界限的维持有重要意义。这些结果表明不同谱系间的基因组分化部分由正选择驱动，且选择信号在多个谱系对之间呈现一定程度的共享，揭示在复杂环境背景下，谱系分化过程中既存在谱系特异性的适应路径，也可能受到共同环境压力的协同作用。

(3) 基因型-环境关联分析的冗余分析结果表明，气候变量对高山栎组植物中

不同谱系的遗传变异的贡献度最大。偏冗余分析显示谱系 A 和 C 同时受到地理与气候因素影响，而谱系 B 和 D 的适应性遗传变异更多受到气候因素驱动。梯度森林分析揭示太阳辐射是驱动适应性遗传变异的重要环境因子，且各谱系对环境变化的响应存在特异性。GO 富集和功能注释结果发现适应性候选基因主要涉及光响应、非生物胁迫响应、代谢调控及生长发育过程。通过对各谱系的正选择高分化区域和局地适应共同基因鉴定出各谱系均有潜在适应干旱胁迫环境的候选同源基因，表明抗旱基因可能是谱系间基因交流与适应异质生境（如干旱或高海拔环境）的重要遗传基础之一，促进了物种在青藏高原及其邻近地区占据不同的生态位，同时也为高山栎组植物的保护和管理提供了科学依据。

5.2 创新点

本研究的创新点主要体现在以下方面：

(1) 从基因组水平的遗传聚类结果揭示了不同于传统形态分类的群体结构，为理解认识高山栎组植物的演化历史和物种界定提供思路。

(2) 从多个遗传谱系比较的角度出发，结合系统发育与群体历史分析，并整合 F_{ST} 与跨群体复合似然比检测到的分化与选择信号，综合解析了高山栎组植物可能起源于青藏高原地区的演化模式，支持“走出青藏高原”假说。

(3) 基于遗传谱系划分基础上开展景观基因组学分析，揭示各谱系对环境梯度的响应差异以及适应性分化的分子基础，为理解高山植物类群的适应性演化提供了新的见解。

5.3 展望

本研究通过对青藏高原及其邻近地区的高山栎组植物的多个近缘种的全基因组重测序数据进行群体基因组学和景观基因组学分析，阐明了高山栎组植物的遗传结构，探究了地理和气候因素等因素对其遗传变异分布格局的影响，阐述了高山栎组植物的谱系分化、基因交流和种群历史动态与历史地质气候变化的关联，鉴别了驱动不同遗传谱系的适应性遗传变异的重要气候因子，鉴定了栎属植物与拟南芥的同源基因，发现高山栎组植物各谱系进化出了不同的干旱胁迫响应策略。但是本研究中依旧存在一些不足之处，未来可以从以下几个方面继续进行深入：

(1) 本研究采集地点覆盖了高山栎组植物在中国的主要分布地点，通过重测序基因组数据分析发现高山栎组植物的七个种间存在频繁的基因交流，遗传界限不明显，但是不同地点的个体数量有限，后续可以增加个体数量对同域分布的近缘种，如在生物热点地区的横断山脉地区，进行更全面地高山栎组植物的遗传分化程度、杂交和渐渗程度研究。

(2) 栎属植物中存在频繁地杂交和渐渗现象，以及其形态特征容易随着环境变化发生可塑性改变可能会模糊基于形态特征的物种界定标准，可通过基因系统发育

建树解析该类群的亲缘关系，结合形态学特征等多种证据，帮助解析高山栎组植物界限或分类问题。

(3) 景观基因组学的研究不局限于温度、降水和辐射等气候变量，充分考虑其土壤条件、海拔和坡度等指标对该组植物的遗传变异的影响。未来可结合物种分布模型与基因组偏移分析，评估不同谱系在未来气候变化情景下的适应潜力与脆弱性，识别潜在的遗传风险区域与关键保护单元，从而为高山栎组植物在青藏高原及其邻近地区的长期保护与资源管理提供更加精准的科学依据。

(4) 尽管本研究鉴定了多个与温度、太阳辐射、干旱及盐胁迫响应以及生长发育相关的适应性候选基因，但在高山栎组植物中的功能和调控机制体现仍需结合基因功能验证、表型关联分析进行验证提高对局地适应机制的理解。

参考文献

- [1] Newbold T, Hudson L N, Hill S L, et al. Global effects of land use on local terrestrial biodiversity[J]. *Nature*, 2015, 520(7545): 45-50.
- [2] Hansen M C, Potapov P V, Moore R, et al. High-resolution global maps of 21st-century forest cover change[J]. *Science*, 2013, 342(6160): 850-853.
- [3] Aitken S N, Yeaman S, Holliday J A, et al. Adaptation, migration or extirpation: climate change outcomes for tree populations[J]. *Evolutionary Applications*, 2008, 1(1): 95-111.
- [4] Merilä J, Hendry A P. Climate change, adaptation, and phenotypic plasticity: the problem and the evidence[J]. *Evolutionary Applications*, 2014, 7(1): 1-14.
- [5] Nicotra A B, Atkin O K, Bonser S P, et al. Plant phenotypic plasticity in a changing climate[J]. *Trends in Plant Science*, 2010, 15(12): 684-692.
- [6] Wright I J, Dong N, Maire V, et al. Global climatic drivers of leaf size[J]. *Science*, 2017, 357(6354): 917-921.
- [7] Tang R, Guo Y, Tang Z. Intraspecific variation in leaf morphology of three widespread woody species along climatic gradients[J]. *Journal of Plant Ecology*, 2024, 17(6): rtae091.
- [8] Siefert A, Violle C, Chalmandrier L, et al. A global meta-analysis of the relative extent of intraspecific trait variation in plant communities[J]. *Ecology Letters*, 2015, 18(12): 1406-1419.
- [9] Kiełtyk P. Intraspecific morphological variation of *Bellidiastrum michelii* (Asteraceae) along a 1,155 m elevation gradient in the Tatra Mountains[J]. *PeerJ*, 2021, 9: e11286.
- [10] Yang Y, Hu M, Fan M, et al. Inter-and intraspecific variation patterns of leaf phenotypic traits in two sympatric pine species[J]. *Forests*, 2022, 13(12): 2014.
- [11] Westoby M, Leishman M, Lord J. Comparative ecology of seed size and dispersal[J]. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, 1996, 351(1345): 1309-1318.
- [12] Ezoe H. Optimal dispersal range and seed size in a stable environment[J]. *Journal of Theoretical Biology*, 1998, 190(3): 287-293.
- [13] Westoby M. A leaf-height-seed (LHS) plant ecology strategy scheme[J]. *Plant and Soil*, 1998, 199(2): 213-227.
- [14] Niklas K. Embryo morphology and seedling evolution. In 'Seedling ecology and evolution'. (Eds MA Leck, VT Parker, RL Simpson) pp. 103-129[Z]. Cambridge University Press: Cambridge. 2008
- [15] Lamont B B, Groom P K. Seeds as a source of carbon, nitrogen, and phosphorus for seedling establishment in temperate regions: a synthesis[J]. *American Journal of Plant Sciences*, 2013, 4(5A): 30.
- [16] Saatkamp A, Cochrane A, Commander L, et al. A research agenda for seed-trait functional ecology[J]. *New Phytologist*, 2019, 221(4): 1764-1775.
- [17] Moles A T, Westoby M. Seedling survival and seed size: a synthesis of the literature[J]. *Journal of Ecology*, 2004, 92(3): 372-383.
- [18] Donohue K, Rubio de Casas R, Burghardt L, et al. Germination, postgermination adaptation, and species ecological ranges[J]. *Annual Review of Ecology, Evolution, and Systematics*, 2010, 41(1): 293-319.
- [19] Parmesan C, Yohe G. A globally coherent fingerprint of climate change impacts across natural systems[J]. *Nature*, 2003, 421(6918): 37-42.
- [20] Cleland E E, Chuine I, Menzel A, et al. Shifting plant phenology in response to global change[J]. *Trends in Ecology & Evolution*, 2007, 22(7): 357-365.
- [21] Stinchcombe J R, Weinig C, Ungerer M, et al. A latitudinal cline in flowering time in *Arabidopsis*

- thaliana* modulated by the flowering time gene *FRIGIDA*[J]. Proceedings of the National Academy of Sciences, 2004, 101(13): 4712-4717.
- [22] Petit R J, Aguinalalde I, de Beaulieu J L, et al. Glacial refugia: hotspots but not melting pots of genetic diversity[J]. Science, 2003, 300(5625): 1563-1565.
- [23] Brewer S, Cheddadi R, De Beaulieu J, et al. The spread of deciduous *Quercus* throughout Europe since the last glacial period[J]. Forest Ecology and Management, 2002, 156(1-3): 27-48.
- [24] Keller S R, Olson M S, Silim S, et al. Genomic diversity, population structure, and migration following rapid range expansion in the Balsam Poplar, *Populus balsamifera*[J]. Molecular Ecology, 2010, 19(6): 1212-1226.
- [25] Cottrell J, Krystufek V, Tabbener H, et al. Postglacial migration of *Populus nigra* L.: lessons learnt from chloroplast DNA[J]. Forest Ecology and Management, 2005, 206(1-3): 71-90.
- [26] Perrier A, Keenan O J, Galloway L F. Revisiting evolution at the rear edge[J]. Trends in Ecology & Evolution, 2026.
- [27] Pauli H, Gottfried M, Dullinger S, et al. Recent plant diversity changes on Europe's mountain summits[J]. Science, 2012, 336(6079): 353-355.
- [28] Dullinger S, Willner W, Plutzar C, et al. Post-glacial migration lag restricts range filling of plants in the European Alps[J]. Global Ecology and Biogeography, 2012, 21(8): 829-840.
- [29] Rumpf S B, Hülber K, Wessely J, et al. Extinction debts and colonization credits of non-forest plants in the European Alps[J]. Nature Communications, 2019, 10(1): 4293.
- [30] Kawecki T J, Ebert D. Conceptual issues in local adaptation[J]. Ecology Letters, 2004, 7(12): 1225-1241.
- [31] Hereford J. A quantitative survey of local adaptation and fitness trade-offs[J]. The American Naturalist, 2009, 173(5): 579-588.
- [32] Savolainen O, Lascoux M, Merilä J. Ecological genomics of local adaptation[J]. Nature Reviews Genetics, 2013, 14(11): 807-820.
- [33] Linhart Y B, Grant M C. Evolutionary significance of local genetic differentiation in plants[J]. Annual Review of Ecology and Systematics, 1996, 27(1): 237-277.
- [34] Martínez-Sancho E, Rellstab C, Fonti P, et al. Genetic and plastic effects on trait variability in two major tree species: insights from common garden experiments across Europe[J]. Forest Ecology and Management, 2025, 597: 123126.
- [35] Antonovics J. The nature of limits to natural selection[J]. Annals of the Missouri Botanical Garden, 1976: 224-247.
- [36] Lenormand T. Gene flow and the limits to natural selection[J]. Trends in Ecology & Evolution, 2002, 17(4): 183-189.
- [37] Lewis J A, Kandala P, Penley M J, et al. Gene flow accelerates adaptation to a parasite[J]. Evolution, 2023, 77(6): 1468-1478.
- [38] Galloway L F, Fenster C B. Population differentiation in an annual legume: local adaptation[J]. Evolution, 2000, 54(4): 1173-1181.
- [39] Sork V L. Genomic studies of local adaptation in natural plant populations[J]. Journal of Heredity, 2018, 109(1): 3-15.
- [40] Leimu R, Fischer M. A meta-analysis of local adaptation in plants[J]. PloS One, 2008, 3(12): e4010.
- [41] Kreyling J, Buhk C, Backhaus S, et al. Local adaptations to frost in marginal and central populations of the dominant forest tree *Fagus sylvatica* L. as affected by temperature and extreme drought in common garden experiments[J]. Ecology and Evolution, 2014, 4(5): 594-605.
- [42] Capblancq T, Lachmuth S, Fitzpatrick M C, et al. From common gardens to candidate genes:

- Exploring local adaptation to climate in red spruce[J]. *New Phytologist*, 2023, 237(5): 1590-1605.
- [43] Mátyás C. Climatic adaptation of trees: rediscovering provenance tests[J]. *Euphytica*, 1996, 92(1): 45-54.
- [44] De Kort H, Vandepitte K, Bruun H H, et al. Landscape genomics and a common garden trial reveal adaptive differentiation to temperature across Europe in the tree species *Alnus glutinosa*[J]. *Molecular Ecology*, 2014, 23(19): 4709-4721.
- [45] Dauphin B, Rellstab C, Wüest R O, et al. Re-thinking the environment in landscape genomics[J]. *Trends in Ecology & Evolution*, 2023, 38(3): 261-274.
- [46] Fitzpatrick M C, Keller S R. Ecological genomics meets community-level modelling of biodiversity: Mapping the genomic landscape of current and future environmental adaptation[J]. *Ecology Letters*, 2015, 18(1): 1-16.
- [47] Sork V, Aitken S, Dyer R, et al. Putting the landscape into the genomics of trees: Approaches for understanding local adaptation and population responses to changing climate[J]. *Tree Genetics & Genomes*, 2013, 9(4): 901-911.
- [48] Steane D A, Potts B M, McLean E, et al. Genome-wide scans detect adaptation to aridity in a widespread forest tree species[J]. *Molecular Ecology*, 2014, 23(10): 2500-2513.
- [49] Ahrens C W, Rymer P D, Stow A, et al. The search for loci under selection: trends, biases and progress[J]. *Molecular Ecology*, 2018, 27(6): 1342-1356.
- [50] Feng L, Du F K. Landscape genomics in tree conservation under a changing environment[J]. *Frontiers in Plant Science*, 2022, 13: 822217.
- [51] 王天瑞, 冯力, 杜芳. 生态适应研究新方法:从种群遗传学到景观基因组学[J]. *中国科学:生命科学*, 2021, 51(02): 167-178.
- [52] Lu Y, Dong H, Fan S, et al. Local adaptation and climate change vulnerability of the relict tree species *Taiwania cryptomerioides* provide insights into its conservation and restoration[J]. *Evolutionary Applications*, 2025, 18(5): e70113.
- [53] Uchiyama K, Ujino-Ihara T, Nakao K, et al. Climate-associated genetic variation and projected genetic offsets for *Cryptomeria japonica* D. Don under future climate scenarios[J]. *Evolutionary Applications*, 2025, 18(2): e70077.
- [54] Zhu X, Zhang H, Lu Z, et al. Common gardens reveal genomic susceptibility and vulnerability to climate change in *Eucalyptus*[J]. *The Plant Journal*, 2025, 123(1): e70336.
- [55] Yao T, Thompson L G, Mosbrugger V, et al. Third pole environment (TPE)[J]. *Environmental Development*, 2012, 3: 52-64.
- [56] Kuang X, Jiao J J. Review on climate change on the Tibetan Plateau during the last half century[J]. *Journal of Geophysical Research: Atmospheres*, 2016, 121(8): 3979-4007.
- [57] Zhisheng A, Kutzbach J E, Prell W L, et al. Evolution of Asian monsoons and phased uplift of the Himalaya-Tibetan plateau since Late Miocene times[J]. *Nature*, 2001, 411(6833): 62-66.
- [58] Molnar P, Boos W R, Battisti D S. Orographic controls on climate and paleoclimate of Asia: thermal and mechanical roles for the Tibetan Plateau[J]. *Annual Review of Earth and Planetary Sciences*, 2010, 38(1): 77-102.
- [59] 张镜铨, 李炳元, 郑度. 论青藏高原范围与面积[J]. *地理研究*, 2002, (01): 1-8.
- [60] Liu J Q, Li J L, Lai Y J. Plant diversity and ecology on the Qinghai-Tibet Plateau[J]. *Journal of Systematics and Evolution*, 2021, 59(6): 1139.
- [61] Favre A, Päckert M, Pauls S U, et al. The role of the uplift of the Qinghai-Tibetan Plateau for the evolution of Tibetan biotas[J]. *Biological Reviews*, 2015, 90(1): 236-253.
- [62] Sun H, Zhang J, Deng T, et al. Origins and evolution of plant diversity in the Hengduan Mountains,

- China[Z]. Elsevier. 2017: 161-166
- [63] Singh D, Hajra P. Floristic diversity[J]. Biodiversity Status in the Himalaya New Delhi: British Council, 1996: 23-38.
- [64] Herrando-Moraira S, Roquet C, Calleja J A, et al. Testing evolutionary hypotheses for the world's main hotspot of temperate alpine flora using *Saussurea* (Compositae) as a model[J]. Molecular Phylogenetics and Evolution, 2025: 108420.
- [65] Wen J, Zhang J Q, Nie Z L, et al. Evolutionary diversifications of plants on the Qinghai-Tibetan Plateau[J]. Frontiers in Genetics, 2014, 5: 4.
- [66] Qiu Y X, Fu C X, Comes H P. Plant molecular phylogeography in China and adjacent regions: tracing the genetic imprints of Quaternary climate and environmental change in the world's most diverse temperate flora[J]. Molecular Phylogenetics and Evolution, 2011, 59(1): 225-244.
- [67] Liu J Q, Wang Y J, Wang A L, et al. Radiation and diversification within the *Ligularia-Cremanthodium-Parasenecio* complex (Asteraceae) triggered by uplift of the Qinghai-Tibetan Plateau[J]. Molecular Phylogenetics and Evolution, 2006, 38(1): 31-49.
- [68] Mao K S, Wang Y, Liu J Q. Evolutionary origin of species diversity on the Qinghai-Tibet Plateau[J]. Journal of Systematics and Evolution, 2021, 59(6): 1142-1158.
- [69] Wang A, Yang M, Liu J. Molecular phylogeny, recent radiation and evolution of gross morphology of the rhubarb genus *Rheum* (Polygonaceae) inferred from chloroplast DNA trn LF sequences[J]. Annals of Botany, 2005, 96(3): 489-498.
- [70] Hewitt G. The genetic legacy of the Quaternary ice ages[J]. Nature, 2000, 405(6789): 907-913.
- [71] Chen S, Wu G, Zhang D, et al. Potential refugium on the Qinghai-Tibet Plateau revealed by the chloroplast DNA phylogeography of the alpine species *Metagentiana striata* (Gentianaceae)[J]. Botanical Journal of the Linnean Society, 2008, 157(1): 125-140.
- [72] López-Pujol J, Ren M X. China: a hot spot of relict plant taxa[J]. Biodiversity Hotspots, 2010: 123-137.
- [73] Hewitt G M. Genetic consequences of climatic oscillations in the Quaternary[J]. Philosophical Transactions of the Royal Society of London Series B: Biological Sciences, 2004, 359(1442): 183-195.
- [74] Yang L, Zhou G. Phylogeography and ecological niche modeling implicate multiple microrefugia of *Swertia tetraptera* during quaternary glaciations[J]. BMC Plant Biology, 2023, 23(1): 450.
- [75] Shimono A, Ueno S, Gu S, et al. Range shifts of *Potentilla fruticosa* on the Qinghai-Tibetan Plateau during glacial and interglacial periods revealed by chloroplast DNA sequence variation[J]. Heredity, 2010, 104(6): 534-542.
- [76] Ebersbach J, Muellner-Riehl A, Michalak I, et al. In and out of the Qinghai-Tibet Plateau: divergence time estimation and historical biogeography of the large arctic-alpine genus *Saxifraga* L[J]. Journal of Biogeography, 2017, 44(4): 900-910.
- [77] Lai Y-J, Han Y, Schuiteman A, et al. Diversification in Qinghai-Tibet Plateau: *Orchidinae* (Orchidaceae) clades exhibiting pre-adaptations play critical role[J]. Molecular Phylogenetics and Evolution, 2021, 157: 107062.
- [78] Zhang J Q, Meng S Y, Allen G A, et al. Rapid radiation and dispersal out of the Qinghai-Tibetan Plateau of an alpine plant lineage *Rhodiola* (Crassulaceae)[J]. Molecular Phylogenetics and Evolution, 2014, 77: 147-158.
- [79] Jia D R, Abbott R J, Liu T L, et al. Out of the Qinghai-Tibet Plateau: evidence for the origin and dispersal of Eurasian temperate plants from a phylogeographic study of *Hippophaë rhamnoides* (Elaeagnaceae)[J]. New Phytologist, 2012, 194(4): 1123-1133.
- [80] Yang W, Feng L, Jiao P, et al. Out of the Qinghai-Tibet plateau: genomic biogeography of the alpine

- monospecific genus *Megadenia* (Biscutelleae, Brassicaceae)[J]. *Molecular Ecology*, 2023, 32(2): 492-503.
- [81] Favre A, Michalak I, Chen C H, et al. Out-of-Tibet: the spatio-temporal evolution of *Gentiana* (Gentianaceae)[J]. *Journal of Biogeography*, 2016, 43(10): 1967-1978.
- [82] Liu P F, He Y X, Di Y X, et al. The uplift of the Qinghai-Xizang Plateau and global climate cooling triggered spatiotemporal evolution of *Lonicera* (Caprifoliaceae)[J]. *Journal of Systematics and Evolution*.
- [83] Weigold M H. Südost-Tibet als Lebensraum, &c[M]. 1935.
- [84] Weigold H. "Die" Biographie Tibets und seiner Vorländer[M]. Verein Sächsischer Ornithologen, 2005.
- [85] Kremer A, Abbott A G, Carlson J E, et al. Genomics of fagaceae[J]. *Tree Genetics & Genomes*, 2012, 8(3): 583-610.
- [86] Denk T, Grimm G W, Manos P S, et al. An updated infrageneric classification of the oaks: review of previous taxonomic schemes and synthesis of evolutionary patterns[J]. *Oaks Physiological Ecology Exploring the Functional Diversity of Genus Quercus L*, 2017: 13-38.
- [87] Hipp A L, Manos P S, Hahn M, et al. Genomic landscape of the global oak phylogeny[J]. *New Phytologist*, 2020, 226(4): 1198-1212.
- [88] Cavender-Bares J. Diversification, adaptation, and community assembly of the American oaks (*Quercus*), a model clade for integrating ecology and evolution[J]. *New Phytologist*, 2019, 221(2): 669-692.
- [89] Ducouso A. Endangered oak resources in Europe and around the Mediterranean basin: marginal populations and minor species[J]. *International Oaks*, 2012, 24: 108-115.
- [90] Plomion C, Aury J M, Amselem J, et al. Oak genome reveals facets of long lifespan[J]. *Nature Plants*, 2018, 4(7): 440-452.
- [91] Kremer A, Hipp A L. Oaks: An evolutionary success story[J]. *New Phytologist*, 2020, 226(4): 987-1011.
- [92] Gugger P F, Cokus S J, Sork V L. Association of transcriptome-wide sequence variation with climate gradients in valley oak (*Quercus lobata*)[J]. *Tree Genetics & Genomes*, 2016, 12(2): 15.
- [93] Leroy T, Louvet J M, Lalanne C, et al. Adaptive introgression as a driver of local adaptation to climate in European white oaks[J]. *New Phytologist*, 2020, 226(4): 1171-1182.
- [94] Sork V L, Fitz-Gibbon S T, Puiu D, et al. First draft assembly and annotation of the genome of a California endemic oak *Quercus lobata* Née (Fagaceae)[J]. *G3: Genes, Genomes, Genetics*, 2016, 6(11): 3485-3495.
- [95] Ai W, Liu Y, Mei M, et al. A chromosome-scale genome assembly of the Mongolian oak (*Quercus mongolica*)[J]. *Molecular Ecology Resources*, 2022, 22(6): 2396-2410.
- [96] Sork V L, Cokus S J, Fitz-Gibbon S T, et al. High-quality genome and methylomes illustrate features underlying evolutionary success of oaks[J]. *Nature Communications*, 2022, 13(1): 2047.
- [97] Kapoor B, Jenkins J, Schmutz J, et al. A haplotype-resolved chromosome-scale genome for *Quercus rubra* L. provides insights into the genetics of adaptive traits for red oak species[J]. *G3: Genes, Genomes, Genetics*, 2023, 13(11): jkad209.
- [98] Wang W B, He X F, Yan X M, et al. Chromosome-scale genome assembly and insights into the metabolome and gene regulation of leaf color transition in an important oak species, *Quercus dentata*[J]. *New Phytologist*, 2023, 238(5): 2016-2032.
- [99] Ramos A M, Usié A, Barbosa P, et al. The draft genome sequence of cork oak[J]. *Scientific Data*, 2018, 5(1): 180069.
- [100] Rey M D, Labella-Ortega M, Guerrero-Sánchez V M, et al. A first draft genome of holm oak

- (*Quercus ilex* subsp. *ballota*), the most representative species of the Mediterranean forest and the Spanish agrosylvopastoral ecosystem “dehesa”[J]. *Frontiers in Molecular Biosciences*, 2023, 10: 1242943.
- [101] Ren Y B, Yu S L, Sun H, et al. Genomic insights into introgression between *Quercus aquifolioides* and its sympatric relatives across elevational gradients[J]. *Molecular Ecology*, 2025, 34(21): e17747.
- [102] Zhou X, Liu N, Jiang X, et al. A chromosome-scale genome assembly of *Quercus gilva*: insights into the evolution of *Quercus* section *Cyclobalanopsis* (Fagaceae)[J]. *Frontiers in Plant Science*, 2022, 13: 1012277.
- [103] Liu X, Zhang W, Zhang Y, et al. Chromosome-scale genomes of *Quercus sichourensis* and *Quercus rex* provide insights into the evolution and adaptation of Fagaceae[J]. *Journal of Genetics and Genomics*, 2024, 51(5): 554-565.
- [104] Fu R, Zhu Y, Liu Y, et al. Genome-wide analyses of introgression between two sympatric Asian oak species[J]. *Nature Ecology & Evolution*, 2022, 6(7): 924-935.
- [105] Chang E, Guo W, Chen J, et al. Chromosome-level genome assembly of *Quercus variabilis* provides insights into the molecular mechanism of cork thickness[J]. *Plant Science*, 2023, 337: 111874.
- [106] Mead A, Fitz-Gibbon S T, Escalona M, et al. The genome assembly of Island Oak (*Quercus tomentella*), a relictual island tree species[J]. *Journal of Heredity*, 2024, 115(2): 221-229.
- [107] Larson D A, Staton M E, Kapoor B, et al. A haplotype-resolved reference genome of *Quercus alba* sheds light on the evolutionary history of oaks[J]. *New Phytologist*, 2025, 246(1): 331-348.
- [108] Martins K, Gugger P F, Llanderal-Mendoza J, et al. Landscape genomics provides evidence of climate-associated genetic variation in Mexican populations of *Quercus rugosa*[J]. *Evolutionary Applications*, 2018, 11(10): 1842-1858.
- [109] Gugger P F, Fitz-Gibbon S T, Albarrán-Lara A, et al. Landscape genomics of *Quercus lobata* reveals genes involved in local climate adaptation at multiple spatial scales[J]. *Molecular Ecology*, 2021, 30(2): 406-423.
- [110] Sun P W, Chang J T, Luo M X, et al. Genomic insights into local adaptation and vulnerability of *Quercus longinux* to climate change[J]. *BMC Plant Biology*, 2024, 24(1): 279.
- [111] Gómez Quijano M J, Gross B L, Etterson J R. Genetic differentiation across a steep and narrow environmental gradient: quantitative genetic and genomic insights into Lake Superior populations of *Quercus rubra*[J]. *Molecular Ecology*, 2024, 33(17): e17483.
- [112] Mead A, Fitz-Gibbon S, Knapp J, et al. Comparison of conservation strategies for California Channel Island Oak (*Quercus tomentella*) using climate suitability predicted from genomic data[J]. *Evolutionary Applications*, 2024, 17(12): e70057.
- [113] Meger J, Ulaszewski B, Chmura D J, et al. Signatures of local adaptation to current and future climate in phenology-related genes in natural populations of *Quercus robur*[J]. *BMC Genomics*, 2024, 25(1): 78.
- [114] Vanhove M, Pina-Martins F, Coelho A C, et al. Using gradient Forest to predict climate response and adaptation in Cork oak[J]. *Journal of Evolutionary Biology*, 2021, 34(6): 910-923.
- [115] Denk T, Grimm G W, Hipp A L, et al. Niche evolution in a northern temperate tree lineage: biogeographical legacies in cork oaks (*Quercus* section *Cerris*)[J]. *Annals of Botany*, 2023, 131(5): 769-787.
- [116] Zhou B F, Yuan S, Crowl A A, et al. Phylogenomic analyses highlight innovation and introgression in the continental radiations of Fagaceae across the Northern Hemisphere[J]. *Nature Communications*, 2022, 13(1): 1320.
- [117] Denk T, Grimm G W. The oaks of western Eurasia: traditional classifications and evidence from two nuclear markers[J]. *Taxon*, 2010, 59(2): 351-366.

- [118] 徐仁, 陶君容, 孙湘君. 希夏邦马峰高山栎化石层的发现及其在植物学和地质学上的意义[J]. *Journal of Integrative Plant Biology*, 1973, (01): 103-119.
- [119] 周浙昆, 普春霞, 陈文允. 青藏高原隆起和高山栎组(壳斗科)分布的关系[J]. *地球科学进展*, 2003, (06): 884-890.
- [120] 黄成就, 张永田, 徐永椿. 中国植物志[M]. 北京: 科学出版社, 1999.
- [121] Singh A, Samant S, Naithani S. Population ecology and habitat suitability modelling of *Quercus semecarpifolia* Sm. in the sub-alpine ecosystem of Great Himalayan National Park, north-western Himalaya, India[J]. *South African Journal of Botany*, 2021, 141: 158-170.
- [122] Shrestha B B. *Quercus semecarpifolia* Sm. in the Himalayan region: ecology, exploitation and threats[J]. *Himalayan Journal of Sciences*, 2003, 1(2): 126-128.
- [123] Rawat B, Rawat J M, Purohit S, et al. A comprehensive review of *Quercus semecarpifolia* Sm.: an ecologically and commercially important Himalayan tree[J]. *Frontiers in Ecology and Evolution*, 2022, 10: 961345.
- [124] Chen L, Deng W, Su T, et al. Late Eocene sclerophyllous oak from Markam Basin, Tibet, and its biogeographic implications[J]. *Science China Earth Sciences*, 2021, 64(11): 1969-1981.
- [125] Sun M, Su T, Shi B Z, et al. Variations in leaf morphological traits of *Quercus guyavifolia* (Fagaceae) were mainly influenced by water and ultraviolet irradiation at high elevations on the Qinghai-Tibet Plateau, China[J]. *International Journal of Agriculture and Biology*, 2016, 18(2).
- [126] Bai X L, Feng T, Zou S, et al. Differences in leaf functional traits of *Quercus rehderiana* Hand.-Mazz. in forests with rocky and non-rocky desertification in Southwest China[J]. *Forests*, 2024, 15(8): 1439.
- [127] Li Y, Zhang Y, Liao P C, et al. Genetic, geographic, and climatic factors jointly shape leaf morphology of an alpine oak, *Quercus aquifolioides* Rehder & EH Wilson[J]. *Annals of Forest Science*, 2021, 78(3): 64.
- [128] 张小莹, 何梅香, 迟荣波, et al. 滇中川滇高山栎叶片气孔性状特征及其生态适应性[J]. *云南民族大学学报(自然科学版)*: 1-9.
- [129] Zhang Y, Luo Y, Qi M, et al. Leaf morphological trait integration and modularity provide insights into ecological adaptation in evergreen oaks[J]. *Forest Ecosystems*, 2025, 14: 100350.
- [130] 王钰雯, 何梅香, 江国东, et al. 乌蒙乡不同海拔梯度刺叶高山栎叶功能性状特征及其对环境因子的响应[J]. *生态学报*, 2024, 44(16): 7238-7248.
- [131] Meng H H, Su T, Gao X Y, et al. Warm-cold colonization: Response of oaks to uplift of the Himalaya-Hengduan Mountains[J]. *Molecular Ecology*, 2017, 26(12): 3276-3294.
- [132] Ju M-M, Feng L, Yang J, et al. Evaluating population genetic structure and demographic history of *Quercus spinosa* (Fagaceae) based on specific length amplified fragment sequencing[J]. *Frontiers in Genetics*, 2019, 10: 965.
- [133] Du F K, Wang T, Wang Y, et al. Contrasted patterns of local adaptation to climate change across the range of an evergreen oak, *Quercus aquifolioides*[J]. *Evolutionary Applications*, 2020, 13(9): 2377-2391.
- [134] Du F K, Hou M, Wang W, et al. Phylogeography of *Quercus aquifolioides* provides novel insights into the Neogene history of a major global hotspot of plant diversity in south-west China[J]. *Journal of Biogeography*, 2017, 44(2): 294-307.
- [135] Liu K, Qi M, Du F K. Ppopulation and landscape genetics provide insights into species conservation of two evergreen oaks in Qinghai-Tbet plateau and adjacent regions[J]. *Frontiers in Plant Science*, 2022, 13: 858526.
- [136] Ma X G, Ren Y B, Sun H. Introgression and incomplete lineage sorting blurred phylogenetic

- relationships across the genomes of sclerophyllous oaks from southwest China[J]. *Cladistics*, 2024, 40(4): 357-373.
- [137] Ju M M, Yue M, Zhao G F. Integrative population genomics reveals niche differentiation and gene flow in chinese sclerophyllous oaks (*Quercus* Sect. *Ilex*)[J]. *Plants*, 2025, 14(15): 2403.
- [138] Wright S. Evolution in Mendelian populations[J]. *Genetics*, 1931, 16(2): 97.
- [139] Avise J C. *Phylogeography: the history and formation of species*[M]. Harvard university press, 2000.
- [140] Slatkin M. Gene flow and the geographic structure of natural populations[J]. *Science*, 1987, 236(4803): 787-792.
- [141] Wright S. Isolation by distance[J]. *Genetics*, 1943, 28(2): 114.
- [142] Manel S, Schwartz M K, Luikart G, et al. Landscape genetics: combining landscape ecology and population genetics[J]. *Trends in Ecology & Evolution*, 2003, 18(4): 189-197.
- [143] Sork V L. Gene flow and natural selection shape spatial patterns of genes in tree populations: Implications for evolutionary processes and applications[J]. *Evolutionary Applications*, 2016, 9(1): 291-310.
- [144] Liu W, Zhao Y, Qi D, et al. The Tanggula Mountains enhance population divergence in *Carex moorcroftii*: a dominant sedge on the Qinghai-Tibetan Plateau[J]. *Scientific Reports*, 2018, 8(1): 2741.
- [145] Nazareno A G, Bemmels J B, Dick C W, et al. Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species[J]. *Molecular Ecology Resources*, 2017, 17(6): 1136-1147.
- [146] Aguirre-Liguori J A, Luna-Sánchez J A, Gasca-Pineda J, et al. Evaluation of the minimum sampling design for population genomic and microsatellite studies: an analysis based on wild maize[J]. *Frontiers in Genetics*, 2020, 11: 870.
- [147] Porebski S, Bailey L G, Baum B R. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components[J]. *Plant Molecular Biology Reporter*, 1997, 15(1): 8-15.
- [148] Chen Y, Chen Y, Shi C, et al. SOAPnuke: A MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data[J]. *Gigascience*, 2018, 7(1): gix120.
- [149] Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform[J]. *Bioinformatics*, 2010, 26(5): 589-595.
- [150] Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform[J]. *Bioinformatics*, 2009, 25(14): 1754-1760.
- [151] McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data[J]. *Genome Research*, 2010, 20(9): 1297-1303.
- [152] Cingolani P, Platts A, Wang L L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3[J]. *Fly*, 2012, 6(2): 80-92.
- [153] Alexander D H, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals[J]. *Genome Research*, 2009, 19(9): 1655-1664.
- [154] Team R C. R A language and environment for statistical computing, R Foundation for Statistical Computing, 2020.
- [155] Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses[J]. *The American Journal of Human Genetics*, 2007, 81(3): 559-575.
- [156] Danecek P, Auton A, Abecasis G, et al. The variant call format and VCFtools[J]. *Bioinformatics*, 2011, 27(15): 2156-2158.

- [157] Korunes K L, Samuk K. pixy: unbiased estimation of nucleotide diversity and divergence in the presence of missing data[J]. *Molecular Ecology Resources*, 2021, 21(4): 1359-1368.
- [158] Xu L, He W, Tai S, et al. VCF2Dis: an ultra-fast and efficient tool to calculate pairwise genetic distance and construct population phylogeny from VCF files[J]. *Gigascience*, 2025, 14: giaf032.
- [159] 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.
- [160] Pickrell J, Pritchard J. Inference of population splits and mixtures from genome-wide allele frequency data[J]. *Nature Precedings*, 2012: 1-1.
- [161] Fitak R R. OptM: Estimating the optimal number of migration edges on population trees using Treemix[J]. *Biology Methods and Protocols*, 2021, 6(1): bpab017.
- [162] Malinsky M, Matschiner M, Svardal H. Dsuite-Fast D-statistics and related admixture evidence from VCF files[J]. *Molecular Ecology Resources*, 2021, 21(2): 584-595.
- [163] 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.
- [164] Li H, Durbin R. Inference of human population history from individual whole-genome sequences[J]. *Nature*, 2011, 475(7357): 493-496.
- [165] Terhorst J, Kamm J A, Song Y S. Robust and scalable inference of population history from hundreds of unphased whole genomes[J]. *Nature Genetics*, 2017, 49(2): 303-309.
- [166] Fick S E, Hijmans R J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas[J]. *International Journal of Climatology*, 2017, 37(12): 4302-4315.
- [167] Naimi B, Hamm N A, Groen T A, et al. Where is positional uncertainty a problem for species distribution modelling?[J]. *Ecography*, 2014, 37(2): 191-203.
- [168] Lischer H E, Excoffier L. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs[J]. *Bioinformatics*, 2012, 28(2): 298-299.
- [169] Hijmans R J, Williams E, Vennes C, et al. Package ‘geosphere’[J]. *Spherical Trigonometry*, 2017, 1(7): 1-45.
- [170] Goslee S, Urban D, Goslee M S. Package ‘ecodist’[J]. Package ‘ecodist’, 2020.
- [171] Oksanen J. *Vegan: an introduction to ordination*[J]. URL <http://cran.r-project.org/web/packages/vegan/vignettes/introvegan.pdf>, 2015, 8: 19.
- [172] Broennimann O, Fitzpatrick M C, Pearman P B, et al. Measuring ecological niche overlap from occurrence and spatial environmental data[J]. *Global Ecology and Biogeography*, 2012, 21(4): 481-497.
- [173] Warren D L, Glor R E, Turelli M. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution[J]. *Evolution*, 2008, 62(11): 2868-2883.
- [174] Dray S, Dufour A-B. The ade4 package: Implementing the duality diagram for ecologists[J]. *Journal of Statistical Software*, 2007, 22: 1-20.
- [175] Di Cola V, Broennimann O, Petitpierre B, et al. ecospat: An R package to support spatial analyses and modeling of species niches and distributions[J]. *Ecography*, 2017, 40(6): 774-787.
- [176] De Queiroz K. Species concepts and species delimitation[J]. *Systematic Biology*, 2007, 56(6): 879-886.
- [177] Stankowski S, Ravinet M. Defining the speciation continuum[J]. *Evolution*, 2021, 75(6): 1256-1273.
- [178] Aldhebiani A Y. Species concept and speciation[J]. *Saudi journal of Biological Sciences*, 2018, 25(3): 437-440.
- [179] Bickford D, Lohman D J, Sodhi N S, et al. Cryptic species as a window on diversity and conservation[J]. *Trends in Ecology & Evolution*, 2007, 22(3): 148-155.

- [180] Xu S. Phylogenetic analysis under reticulate evolution[J]. *Molecular Biology and Evolution*, 2000, 17(6): 897-907.
- [181] Seehausen O. Hybridization and adaptive radiation[J]. *Trends in Ecology & Evolution*, 2004, 19(4): 198-207.
- [182] Liu J, Yu L, Arnold M L, et al. Reticulate evolution: frequent introgressive hybridization among Chinese hares (genus *Lepus*) revealed by analyses of multiple mitochondrial and nuclear DNA loci[J]. *BMC Evolutionary Biology*, 2011, 11(1): 223.
- [183] Mallet J. Hybridization as an invasion of the genome[J]. *Trends in Ecology & Evolution*, 2005, 20(5): 229-237.
- [184] Abbott R, Albach D, Ansell S, et al. Hybridization and speciation[J]. *Journal of Evolutionary Biology*, 2013, 26(2): 229-246.
- [185] Wang L, Abbott R J, Zheng W, et al. History and evolution of alpine plants endemic to the Qinghai-Tibetan Plateau: *Aconitum gymnantrum* (Ranunculaceae)[J]. *Molecular Ecology*, 2009, 18(4): 709-721.
- [186] Han M, Zhang J, Li D, et al. Phylogeographical pattern and population evolution history of Indigenous *Elymus sibiricus* L. on Qinghai-Tibetan plateau[J]. *Frontiers in Plant Science*, 2022, 13: 882601.
- [187] Su X, Wu G, Li L, et al. Species delimitation in plants using the Qinghai-Tibet Plateau endemic *Orinus* (Poaceae: Tridentinae) as an example[J]. *Annals of Botany*, 2015, 116(1): 35-48.
- [188] Li Q, Yu J, Ding X J, et al. Population genetics and ecological niche modelling shed light on species boundaries and evolutionary history of *Aconitum pendulum* and *A. flavum*[J]. *Flora*, 2024, 314: 152507.
- [189] Guo J F, Zhao W, Andersson B, et al. Genomic clines across the species boundary between a hybrid pine and its progenitor in the eastern Tibetan Plateau[J]. *Plant Communications*, 2023, 4(4).
- [190] Shi Y J, Huang J L, Mi J X, et al. A model of hybrid speciation process drawn from three new poplar species originating from distant hybridization between sections[J]. *Molecular Phylogenetics and Evolution*, 2024, 190: 107966.
- [191] Zheng W, Yan L J, Burgess K S, et al. Natural hybridization among three *Rhododendron* species (Ericaceae) revealed by morphological and genomic evidence[J]. *BMC Plant Biology*, 2021, 21(1): 529.
- [192] Petit R J, Bodénès C, Ducouso A, et al. Hybridization as a mechanism of invasion in oaks[J]. *New Phytologist*, 2004, 161(1): 151-164.
- [193] Curtu A L, Gailing O, Finkeldey R. Evidence for hybridization and introgression within a species-rich oak (*Quercus* spp.) community[J]. *BMC Evolutionary Biology*, 2007, 7(1): 218.
- [194] Harrison R G, Larson E L. Hybridization, introgression, and the nature of species boundaries[J]. *Journal of Heredity*, 2014, 105(S1): 795-809.
- [195] Morales-Saldaña S, Hipp A L, Valencia-Ávalos S, et al. Divergence and reticulation in the Mexican white oaks: Ecological and phylogenomic evidence on species limits and phylogenetic networks in the *Quercus laeta* complex (Fagaceae)[J]. *Annals of Botany*, 2024, 133(7): 1007-1024.
- [196] Holsinger K E, Weir B S. Genetics in geographically structured populations: defining, estimating and interpreting F_{ST}[J]. *Nature Reviews Genetics*, 2009, 10(9): 639-650.
- [197] Peñaloza-Ramírez J M, González-Rodríguez A, Mendoza-Cuenca L, et al. Interspecific gene flow in a multispecies oak hybrid zone in the Sierra Tarahumara of Mexico[J]. *Annals of Botany*, 2010, 105(3): 389-399.
- [198] Excoffier L, Ray N. Surfing during population expansions promotes genetic revolutions and structuration[J]. *Trends in Ecology & Evolution*, 2008, 23(7): 347-351.
- [199] Meng L, Chen G, Li Z, et al. Refugial isolation and range expansions drive the genetic structure of *Oxyria sinensis* (Polygonaceae) in the Himalaya-Hengduan Mountains[J]. *Scientific Reports*, 2015, 5(1):

10396.

- [200] Wang B, Mao J F, Gao J, et al. Colonization of the Tibetan Plateau by the homoploid hybrid pine *Pinus densata*[J]. *Molecular Ecology*, 2011, 20(18): 3796-3811.
- [201] Xu T, Abbott R J, Milne R I, et al. Phylogeography and allopatric divergence of cypress species (*Cupressus* L.) in the Qinghai-Tibetan Plateau and adjacent regions[J]. *BMC Evolutionary Biology*, 2010, 10(1): 194.
- [202] Hernandez J O, Park B B. The leaf trichome, venation, and mesophyll structural traits play important roles in the physiological responses of oak seedlings to water-deficit stress[J]. *International Journal of Molecular Sciences*, 2022, 23(15): 8640.
- [203] Zhang K L, Leng Y N, Hao R R, et al. Adaptation of high-altitude plants to harsh environments: Application of phenotypic-variation-related methods and multi-omics techniques[J]. *International Journal of Molecular Sciences*, 2024, 25(23): 12666.
- [204] Reed D H, Frankham R. Correlation between fitness and genetic diversity[J]. *Conservation Biology*, 2003, 17(1): 230-237.
- [205] Hughes A R, Inouye B D, Johnson M T, et al. Ecological consequences of genetic diversity[J]. *Ecology Letters*, 2008, 11(6): 609-623.
- [206] 于海彬, 张镫铤. 青藏高原及其周边地区高山植物谱系地理学研究进展[J]. *西北植物学报*, 2013, 33(06): 1268-1278.
- [207] Zhang J Q, Meng S Y, Rao G Y. Phylogeography of *Rhodiola kirilowii* (Crassulaceae): A story of Miocene divergence and Quaternary expansion[J]. *PloS One*, 2014, 9(11): e112923.
- [208] Fu P C, Mo B J, Wan H X, et al. Divergence of alpine plant populations of three *Gentianaceae* species in the Qinling sky Island[J]. *BMC Plant Biology*, 2025, 25(1): 144.
- [209] Xie C, Xie D F, Zhong Y, et al. The effect of Hengduan Mountains Region (HMR) uplift to environmental changes in the HMR and its eastern adjacent area: tracing the evolutionary history of *Allium* section *Sikkimensia* (Amaryllidaceae)[J]. *Molecular Phylogenetics and Evolution*, 2019, 130: 380-396.
- [210] Liu Y Y, Jin W T, Wei X X, et al. Cryptic speciation in the Chinese white pine (*Pinus armandii*): Implications for the high species diversity of conifers in the Hengduan Mountains, a global biodiversity hotspot[J]. *Molecular Phylogenetics and Evolution*, 2019, 138: 114-125.
- [211] Feng L, Ruhsam M, Wang Y H, et al. Using demographic model selection to untangle allopatric divergence and diversification mechanisms in the *Rheum palmatum* complex in the Eastern Asiatic Region[J]. *Molecular Ecology*, 2020, 29(10): 1791-1805.
- [212] Leroy T, Roux C, Villate L, et al. Extensive recent secondary contacts between four European white oak species[J]. *New Phytologist*, 2017, 214(2): 865-878.
- [213] Jiang X L, Hipp A L, Deng M, et al. East Asian origins of European holly oaks (*Quercus* section *Ilex* Loudon) via the Tibet-Himalaya[J]. *Journal of Biogeography*, 2019, 46(10): 2203-2214.
- [214] Xing Y, Ree R H. Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot[J]. *Proceedings of the National Academy of Sciences*, 2017, 114(17): E3444-E3451.
- [215] Stöcklin J, Kuss P, Pluess A R. Genetic diversity, phenotypic variation and local adaptation in the alpine landscape: case studies with alpine plant species[J]. *Botanica Helvetica*, 2009, 119(2): 125-133.
- [216] Hamrick J L. Consequences of dispersal in plants[J]. *Population Dynamics in Ecological Space and Time*, 1996.
- [217] Wang I J, Bradburd G S. Isolation by environment[J]. *Molecular Ecology*, 2014, 23(23): 5649-5662.
- [218] Lu S, Liu L, Lei W, et al. Cryptic divergence in and evolutionary dynamics of endangered hybrid *Picea brachytyla* sensu stricto in the Qinghai-Tibet Plateau[J]. *BMC Plant Biology*, 2024, 24(1): 1202.

- [219] Terrones-Ramírez A K, Robles-Bello S M, Vázquez-López M, et al. Recent genetic, phenetic and ecological divergence across the Mesoamerican highlands: A study case with *Diglossa baritula* (Aves: Thraupidae)[J]. PeerJ, 2024, 12: e16797.
- [220] Feng L, Zhou L, Zhang T, et al. Niche dynamics below the species level: evidence from evaluating niche shifts within *Quercus aquifolioides*[J]. Forests, 2023, 14(4): 690.
- [221] Rödder D, Engler J O. Quantitative metrics of overlaps in *Grinnellian* niches: advances and possible drawbacks[J]. Global Ecology and Biogeography, 2011, 20(6): 915-927.
- [222] Ortego J, Riordan E C, Gugger P F, et al. Influence of environmental heterogeneity on genetic diversity and structure in an endemic southern Californian oak[J]. Molecular Ecology, 2012, 21(13): 3210-3223.
- [223] Gugger P F, Ikegami M, Sork V L. Influence of late Quaternary climate change on present patterns of genetic variation in valley oak, *Quercus lobata* Née[J]. Molecular Ecology, 2013, 22(13): 3598-3612.
- [224] Nosil P, Funk D J, Ortiz-Barrientos D. Divergent selection and heterogeneous genomic divergence[J]. Molecular Ecology, 2009, 18(3): 375-402.
- [225] Seehausen O, Butlin R K, Keller I, et al. Genomics and the origin of species[J]. Nature Reviews Genetics, 2014, 15(3): 176-192.
- [226] Feder J L, Egan S P, Nosil P. The genomics of speciation-with-gene-flow[J]. Trends in Genetics, 2012, 28(7): 342-350.
- [227] Turner T L, Hahn M W, Nuzhdin S V. Genomic islands of speciation in *Anopheles gambiae*[J]. PLoS Biology, 2005, 3(9): e285.
- [228] Cruickshank T E, Hahn M W. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow[J]. Molecular Ecology, 2014, 23(13): 3133-3157.
- [229] Wolf J B, Ellegren H. Making sense of genomic islands of differentiation in light of speciation[J]. Nature Reviews Genetics, 2017, 18(2): 87-100.
- [230] Körner C. The use of 'altitude' in ecological research[J]. Trends in Ecology & Evolution, 2007, 22(11): 569-574.
- [231] Chen H, Patterson N, Reich D. Population differentiation as a test for selective sweeps[J]. Genome Research, 2010, 20(3): 393-402.
- [232] Mi H, Muruganujan A, Casagrande J T, et al. Large-scale gene function analysis with the PANTHER classification system[J]. Nature Protocols, 2013, 8(8): 1551-1566.
- [233] Rendón-Anaya M, Wilson J, Sveinsson S, et al. Adaptive introgression facilitates adaptation to high latitudes in European aspen (*Populus tremula* L.)[J]. Molecular Biology and Evolution, 2021, 38(11): 5034-5050.
- [234] Guillaume A S, Leempoel K, Rogivue A, et al. Integrating very high resolution environmental proxies in genotype-environment association studies[J]. Evolutionary Applications, 2024, 17(7): e13737.
- [235] Higo K, Ugawa Y, Iwamoto M, et al. Plant cis-acting regulatory DNA elements (PLACE) database: 1999[J]. Nucleic Acids Research, 1999, 27(1): 297-300.
- [236] Wray G A. The evolutionary significance of cis-regulatory mutations[J]. Nature Reviews Genetics, 2007, 8(3): 206-216.
- [237] Nielsen R. Molecular signatures of natural selection[J]. Annu Rev Genet, 2005, 39(1): 197-218.
- [238] Geng F D, Liu M Q, Zhang X D, et al. Genomics of hybrid parallel origin in *Aquilegia ecalcarata*[J]. BMC Ecology and Evolution, 2024, 24(1): 75.
- [239] Yang F S, Liu M, Guo X, et al. Signatures of adaptation and purifying selection in highland populations of *Dasiphora fruticosa*[J]. Molecular Biology and Evolution, 2024, 41(6): msae099.
- [240] Losos J B, Glor R E. Phylogenetic comparative methods and the geography of speciation[J]. Trends

- in *Ecology & Evolution*, 2003, 18(5): 220-227.
- [241] Moses A M. Statistical tests for natural selection on regulatory regions based on the strength of transcription factor binding sites[J]. *BMC Evolutionary Biology*, 2009, 9(1): 286.
- [242] Schmitz R J, Grotewold E, Stam M. Cis-regulatory sequences in plants: their importance, discovery, and future challenges[J]. *The Plant Cell*, 2022, 34(2): 718-741.
- [243] Marand A P, Eveland A L, Kaufmann K, et al. Cis-Regulatory elements in plant development, adaptation, and evolution[J]. *Annual Review of Plant Biology*, 2023, 74(1): 111-137.
- [244] Leng Y J, Yao Y S, Yang K Z, et al. Arabidopsis ERdj3B coordinates with ERECTA-family receptor kinases to regulate ovule development and the heat stress response[J]. *The Plant Cell*, 2022, 34(10): 3665-3684.
- [245] Yamamoto M, Uji S, Sugiyama T, et al. ERdj3B-mediated quality control maintains anther development at high temperatures[J]. *Plant Physiology*, 2020, 182(4): 1979-1990.
- [246] Yu H, Wang G, Hu Z, et al. Long-term land-atmosphere energy and water exchange observational dataset over central Tibetan Plateau[J]. *Earth System Science Data*, 2025, 17(12): 6871-6888.
- [247] Lu H L, Zhou H, Zuo H M, et al. Analysis of the evolutionary patterns and drivers of Capturability of Atmospheric Water (CAW) on the Qinghai-Tibet Plateau[J]. *Atmospheric Research*, 2025: 108484.
- [248] Zhang S H, Bai H Y, Qi G Z, et al. Changes of climate zone boundary of the Qinling Mountains from 1960 to 2019[J]. *Journal of Natural Resources*, 2021, 36(10): 2491-2506.
- [249] Rui Y, Xiao C, Yi H, et al. POLYGALACTURONASE INVOLVED IN EXPANSION3 functions in seedling development, rosette growth, and stomatal dynamics in *Arabidopsis thaliana*[J]. *The Plant Cell*, 2017, 29(10): 2413-2432.
- [250] Miao Y, Lv D, Wang P, et al. An Arabidopsis glutathione peroxidase functions as both a redox transducer and a scavenger in abscisic acid and drought stress responses[J]. *The Plant Cell*, 2006, 18(10): 2749-2766.
- [251] Zheng M, Liu X, Lin J, et al. Histone acetyltransferase GCN 5 contributes to cell wall integrity and salt stress tolerance by altering the expression of cellulose synthesis genes[J]. *The Plant Journal*, 2019, 97(3): 587-602.
- [252] Zhang X, Li M, Zhang X, et al. A receptor-kinase cascade confers cold-induced root growth inhibition in Arabidopsis[J]. *Nature Plants*, 2025, 11(7): 1441-1454.
- [253] Ruelland E, Cantrel C, Gawer M, et al. Activation of phospholipases C and D is an early response to a cold exposure in Arabidopsis suspension cells[J]. *Plant Physiology*, 2002, 130(2): 999-1007.
- [254] Feng Y, Li Z, Kong X, et al. Plant coping with cold stress: molecular and physiological adaptive mechanisms with future perspectives[J]. *Cells*, 2025, 14(2): 110.
- [255] Manel S, Andrello M, Henry K, et al. Predicting genotype environmental range from genome-environment associations[J]. *Molecular Ecology*, 2018, 27(13): 2823-2833.
- [256] Li S L, Li Z G, Yang L T, et al. Differential effects of cold stress on chloroplasts structures and photosynthetic characteristics in cold-sensitive and cold-tolerant cultivars of sugarcane[J]. *Sugar Tech*, 2018, 20(1): 11-20.
- [257] Jung J M, Jung S, Ahmed M R, et al. Invasion risk of the yellow crazy ant (*Anoplolepis gracilipes*) under the Representative Concentration Pathways 8.5 climate change scenario in South Korea[J]. *Journal of Asia-Pacific Biodiversity*, 2017, 10(4): 548-554.
- [258] Mizukami Y, Fischer R L. Plant organ size control: AINTEGUMENTA regulates growth and cell numbers during organogenesis[J]. *Proceedings of the National Academy of Sciences*, 2000, 97(2): 942-947.
- [259] Schroeder R Y, Zhu A, Eubel H, et al. The ribokinases of *Arabidopsis thaliana* and *Saccharomyces*

- cerevisiae* are required for ribose recycling from nucleotide catabolism, which in plants is not essential to survive prolonged dark stress[J]. *New Phytologist*, 2018, 217(1): 233-244.
- [260] Jin Y, Liao M, Hou Y, et al. Spatial patterns and variations in leaf traits of alpine plants on the interior Tibetan Plateau[J]. *Global Ecology and Conservation*, 2024, 53: e03037.
- [261] Fan B, Ma Z, Gao P, et al. Functional traits of male and female leaves of *Hippophae tibetana* on the eastern edge of the Tibetan Plateau and their altitudinal variability[J]. *Plants*, 2022, 11(19): 2484.
- [262] Nosil P, Feder J L, Gompert Z. How many genetic changes create new species?[J]. *Science*, 2021, 371(6531): 777-779.
- [263] Rellstab C, Gugerli F, Eckert A J, et al. A practical guide to environmental association analysis in landscape genomics[J]. *Molecular Ecology*, 2015, 24(17): 4348-4370.
- [264] Walden N, Lucek K, Willi Y. Lineage-specific adaptation to climate involves flowering time in North American *Arabidopsis lyrata*[J]. *Molecular Ecology*, 2020, 29(8): 1436-1451.
- [265] Moritz C. Defining 'evolutionarily significant units' for conservation[J]. *Trends in Ecology & Evolution*, 1994, 9(10): 373-375.
- [266] Manel S, Joost S, Epperson B K, et al. Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field[J]. *Molecular Ecology*, 2010, 19(17): 3760-3772.
- [267] Forester B R, Jones M R, Joost S, et al. Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes[J]. *Molecular Ecology*, 2016, 25(1): 104-120.
- [268] Lotterhos K E, Whitlock M C. Evaluation of demographic history and neutral parameterization on the performance of F_{ST} outlier tests[J]. *Molecular Ecology*, 2014, 23(9): 2178-2192.
- [269] Foll M. BayeScan v2. 1 user manual[J]. *Ecology*, 2012, 20(10): 1450-1462.
- [270] Luu K, Bazin E, Blum M G. pcadapt: an R package to perform genome scans for selection based on principal component analysis[J]. *Molecular Ecology Resources*, 2017, 17(1): 67-77.
- [271] Frichot E, François O. LEA: an R package for landscape and ecological association studies[J]. *Methods in Ecology and Evolution*, 2015, 6(8): 925-929.
- [272] Oksanen J, Blanchet F, Friendly M, et al. vegan: Community Ecology Package. Ordination methods, diversity analysis and other functions for community and vegetation ecologists. Version 2.3-5[J]. 2016.
- [273] Ellis N, Smith S J, Pitcher C R. Gradient forests: calculating importance gradients on physical predictors[J]. *Ecology*, 2012, 93(1): 156-168.
- [274] Chung M Y, Merilä J, Li J, et al. Neutral and adaptive genetic diversity in plants: an overview[J]. *Frontiers in Ecology and Evolution*, 2023, 11: 1116814.
- [275] Nocchi G, Wang J, Yang L, et al. Genomic signals of local adaptation and hybridization in Asian white birch[J]. *Molecular Ecology*, 2023, 32(3): 595-612.
- [276] Hughes C E, Atchison G W. The ubiquity of alpine plant radiations: from the Andes to the Hengduan Mountains[J]. *New Phytologist*, 2015, 207(2): 275-282.
- [277] Feng S, Xi E, Wan W, et al. Genomic signals of local adaptation in *Picea crassifolia*[J]. *BMC Plant Biology*, 2023, 23(1): 534.
- [278] Bosco L, Bergamaschi H, Marodin G. Solar radiation effects on growth, anatomy, and physiology of apple trees in a temperate climate of Brazil[J]. *International Journal of Biometeorology*, 2020, 64(11): 1969-1980.
- [279] Wang Y, Qiu N, Wang X, et al. Effects of enhanced UV-B radiation on fitness of an alpine species *Cerastium glomeratum* Thuill[J]. *Journal of Plant Ecology*, 2008, 1(3): 197-202.
- [280] Sun P, Hao R, Fan F, et al. Adaptation of high-altitude plants to plateau abiotic stresses: a case study of the Qinghai-Tibet plateau[J]. *International Journal of Molecular Sciences*, 2025, 26(5): 2292.
- [281] Bontrager M, Angert A L. Gene flow improves fitness at a range edge under climate change[J].

Evolution Letters, 2019, 3(1): 55-68.

[282] Suarez-Gonzalez A, Lexer C, Cronk Q C. Adaptive introgression: a plant perspective[J]. Biology Letters, 2018, 14(3).

[283] 贺金生, 陈伟烈, 王勋陵. 高山栎叶的形态结构及其与生态环境的关系[J]. 植物生态学报, 1994, (03): 219-227.

[284] Lee K M, Coop G. Distinguishing among modes of convergent adaptation using population genomic data[J]. Genetics, 2017, 207(4): 1591-1619.

[285] Qiang X, An Z, Song Y, et al. New eolian red clay sequence on the western Chinese Loess Plateau linked to onset of Asian desertification about 25 Ma ago[J]. Science China Earth Sciences, 2011, 54(1): 136-144.

[286] Sun J, Windley B F. Onset of aridification by 34 Ma across the Eocene-Oligocene transition in Central Asia[J]. Geology, 2015, 43(11): 1015-1018.

附录

附表 1 高山栎组植物样本的测序质量及群体遗传参数统计

Appendix Table 1: Summary statistics of whole genome resequencing data and population genetic for each individual of *Quercus* sect. *Heterobalanus*

物种 Species	样本名称 Sample	比对上的 读段数 Mapping read	比对率 Mapping ratio	干净读段 数 Clean read	干净碱基数 Clean base	测序 深度 Depth	正确配对 的读段数 Properly paired read	正确配 对率 Properly paired ratio	观测杂 合度 H_o	期望 杂合 度 H_e	近交 系数 F_{IS}
<i>Q. guyavifolia</i>	QG_BMXS_3	221263159	98.86%	216823750	32523562500	28.342	181831064	83.86%	0.23	0.25	0.10
<i>Q. guyavifolia</i>	QG_BMXS_4	226524468	98.89%	221688750	33253312500	28.978	183042272	82.57%	0.23	0.25	0.07
<i>Q. guyavifolia</i>	QG_ELS_2	265155237	98.80%	259447370	38917105500	33.914	215408858	83.03%	0.25	0.26	0.03
<i>Q. guyavifolia</i>	QG_ELS_3	292730471	98.73%	286578738	42986810700	37.46	234399194	81.79%	0.25	0.26	0.02
<i>Q. guyavifolia</i>	QG_MLXS_5	222592533	98.92%	218024510	32703676500	28.499	183257198	84.05%	0.23	0.25	0.10
<i>Q. guyavifolia</i>	QG_MLXS_6	244939040	98.91%	240108704	36016305600	31.386	201913566	84.09%	0.24	0.25	0.05
<i>Q. guyavifolia</i>	QG_MNP_10	207938538	98.44%	204060370	30609055500	26.674	170110196	83.36%	0.21	0.25	0.17
<i>Q. guyavifolia</i>	QG_MNP_5	218544000	98.59%	214633012	32194951800	28.056	178956438	83.38%	0.34	0.25	-0.36
<i>Q. guyavifolia</i>	QG_MPS_1	242251986	98.50%	237899474	35684921100	31.097	198338286	83.37%	0.18	0.26	0.30
<i>Q. guyavifolia</i>	QG_MPS_8	218851043	98.56%	214893590	32234038500	28.09	178006742	82.83%	0.18	0.26	0.29
<i>Q. guyavifolia</i>	QG_MYL_3	241317979	98.79%	236405338	35460800700	30.902	198873182	84.12%	0.24	0.26	0.07
<i>Q. guyavifolia</i>	QG_MYL_5	243022297	98.46%	239039458	35855918700	31.246	200309326	83.80%	0.23	0.26	0.08
<i>Q. monimotricha</i>	QO_ELS_5	275370409	98.20%	270453496	40568024400	35.352	221881832	82.04%	0.25	0.26	0.04
<i>Q. monimotricha</i>	QO_ELS_7	243857271	98.39%	238955656	35843348400	31.235	197605674	82.70%	0.22	0.26	0.15
<i>Q. monimotricha</i>	QO_GZX_4	214694418	98.54%	211188228	31678234200	27.605	173709420	82.25%	0.22	0.26	0.12
<i>Q. monimotricha</i>	QO_LS_1	212663306	95.68%	215248054	32287208100	28.136	170676252	79.29%	0.22	0.26	0.12
<i>Q. monimotricha</i>	QO_LS_2	270378259	96.58%	270188968	40528345200	35.318	218740866	80.96%	0.23	0.26	0.10
<i>Q. monimotricha</i>	QO_MNP_1	217434374	98.24%	214427414	32164112100	28.029	173243300	80.79%	0.23	0.26	0.09
<i>Q. monimotricha</i>	QO_MNP_7	201378880	98.62%	197785776	29667866400	25.853	163864272	82.85%	0.23	0.26	0.09
<i>Q. monimotricha</i>	QO_PMS_1	222542066	98.13%	219185882	32877882300	28.651	177166546	80.83%	0.23	0.26	0.08

物种 Species	样本名称 Sample	比对上的 读段数 Mapping read	比对率 Mapping ratio	干净读段 数 Clean read	干净碱基数 Clean base	测序 深度 Depth	正确配对 的读段数 Properly paired read	正确配 对率 Properl y paired ratio	观测杂 合度 H_0	期望 杂合 度 H_E	近交 系数 F_{IS}
<i>Q. monimotricha</i>	QO_PMS_4	278841548	97.96%	274816664	41222499600	35.922	224695088	81.76%	0.24	0.26	0.08
<i>Q. monimotricha</i>	QO_XZ_2	246949666	98.97%	241106298	36165944700	31.516	198642960	82.39%	0.23	0.26	0.08
<i>Q. monimotricha</i>	QO_ZR_11	256325835	98.53%	251258096	37688714400	32.843	206014872	81.99%	0.34	0.26	-0.34
<i>Q. monimotricha</i>	QO_ZR_5	239948991	83.45%	278779128	41816869200	36.44	196712328	70.56%	0.33	0.26	-0.28
<i>Q. aquifolioides</i>	Qa_GBX_2	293953794	98.34%	289635454	43445318100	37.86	247092146	85.31%	0.21	0.25	0.17
<i>Q. aquifolioides</i>	Qa_GBX_21	288732832	93.21%	300426938	45064040700	39.27	237390922	79.02%	0.20	0.25	0.22
<i>Q. aquifolioides</i>	Qa_LZA_26	205774543	98.52%	202433570	30365035500	26.461	171432134	84.69%	0.20	0.25	0.20
<i>Q. aquifolioides</i>	Qa_MLJ_5	255304548	98.39%	251491496	37723724400	32.874	213893978	85.05%	0.20	0.25	0.19
<i>Q. aquifolioides</i>	Qa_LZ_12	279542837	98.72%	274349092	41152363800	35.861	242081624	88.24%	0.21	0.25	0.16
<i>Q. aquifolioides</i>	Qa_LZD_17	214637501	97.71%	212825472	31923820800	27.819	179637680	84.41%	0.20	0.25	0.20
<i>Q. aquifolioides</i>	Qa_BMR_9	208253034	98.84%	204214316	30632147400	26.694	173382152	84.90%	0.20	0.25	0.22
<i>Q. aquifolioides</i>	Qa_BMS_3	260739646	97.98%	257300196	38595029400	33.633	212027224	82.40%	0.19	0.25	0.24
<i>Q. aquifolioides</i>	Qa_CYX_14	286702832	0.9798	283183644	42477546600	37.016	242199348	0.8553	0.21	0.25	0.16
<i>Q. aquifolioides</i>	Qa_CYX_15	268900769	0.9839	264487044	39673056600	34.572	228039462	0.8622	0.21	0.25	0.15
<i>Q. aquifolioides</i>	Qa_MKD_14	238883057	0.9805	235942986	35391447900	30.841	202820592	0.8596	0.26	0.25	-0.03
<i>Q. aquifolioides</i>	Qa_MKR_15	278985092	0.9903	273197366	40979604900	35.711	240355052	0.8798	0.22	0.25	0.11
<i>Q. aquifolioides</i>	Qa_BZL_19	213556436	0.9818	209978532	31496779800	27.447	175910050	0.8378	0.27	0.25	-0.08
<i>Q. aquifolioides</i>	Qa_ZD_3	339307657	0.9927	330187906	49528185900	43.16	292771056	0.8867	0.24	0.24	0.26
<i>Q. aquifolioides</i>	Qa_XCR_6	189526594	0.9909	184831208	27724681200	24.16	160012222	0.8657	0.21	0.25	0.17
<i>Q. aquifolioides</i>	Qa_DCE_8	221203731	0.9912	215914040	32387106000	28.223	185902214	0.861	0.22	0.25	0.12
<i>Q. aquifolioides</i>	Qa_YJH_2	303603641	0.9908	296933602	44540040300	38.813	260595544	0.8776	0.23	0.25	0.10
<i>Q. aquifolioides</i>	Qa_YJH_5	254219146	0.9803	251259264	37688889600	32.843	212268758	0.8448	0.23	0.25	0.10
<i>Q. aquifolioides</i>	Qa_DMX_7	246823137	0.9911	240920886	36138132900	31.492	203603042	0.8451	0.24	0.26	0.06
<i>Q. aquifolioides</i>	Qa_DMX_8	249906247	0.9885	244506206	36675930900	31.96	205854600	0.8419	0.24	0.26	0.07
<i>Q. aquifolioides</i>	Qa_JCK_7	287274376	0.9899	280434602	42065190300	36.657	238945994	0.8521	0.22	0.22	0.26
<i>Q. aquifolioides</i>	Qa_JCK_12	296268091	0.9881	290033146	43504971900	37.912	247272228	0.8526	0.23	0.26	0.10

物种 Species	样本名称 Sample	比对上的 读段数 Mapping read	比对率 Mapping ratio	干净读段 数 Clean read	干净碱基数 Clean base	测序 深度 Depth	正确配对 的读段数 Properly paired read	正确配 对率 Properl y paired ratio	观测杂 合度 H_0	期望 杂合 度 H_E	近交 系数 F_{IS}
<i>Q. aquifolioides</i>	Qa_JCK_14	304303485	0.9866	298097130	44714569500	38.966	252757514	0.8479	0.22	0.25	0.10
<i>Q. aquifolioides</i>	Qa_PW_6	236103940	0.8983	255220584	38283087600	33.361	197339182	0.7732	0.23	0.25	0.08
<i>Q. spinosa</i>	Qs_TM_1	230946538	84.99%	264265746	39639861900	34.543	188343576	71.27%	0.21	0.25	0.16
<i>Q. spinosa</i>	Qs_TM_3	243482023	98.21%	239637564	35945634600	31.324	200601516	83.71%	0.20	0.25	0.19
<i>Q. spinosa</i>	Qs_CYXC_2	252158298	98.50%	247746774	37162016100	32.384	209865832	84.71%	0.21	0.25	0.16
<i>Q. spinosa</i>	Qs_BLLK_9	290648860	98.77%	284711856	42706778400	37.216	243649370	85.58%	0.22	0.22	0.26
<i>Q. spinosa</i>	Qs_XGJQ_1	261197630	98.56%	256733246	38509986900	33.559	219415256	85.46%	0.25	0.28	0.11
<i>Q. spinosa</i>	Qs_XGJQ_3	289949375	98.21%	285670630	42850594500	37.341	241929122	84.69%	0.25	0.28	0.10
<i>Q. spinosa</i>	Qs_XGJQ_2	268489177	98.82%	262654124	39398118600	34.333	217171454	82.68%	0.26	0.28	0.07
<i>Q. spinosa</i>	Qs_GZ_4	256704476	98.88%	251272332	37690849800	32.845	213213282	84.85%	0.22	0.22	0.26
<i>Q. spinosa</i>	Qs_GDS_8	370248570	98.48%	362399486	54359922900	47.371	306803070	84.66%	0.24	0.24	0.26
<i>Q. spinosa</i>	Qs_SWJ_2	314182966	98.68%	307630914	46144637100	40.212	261313206	84.94%	0.23	0.23	0.26
<i>Q. spinosa</i>	Qs_LBZ_6	291006406	98.50%	285544338	42831650700	37.325	236074446	82.68%	0.28	0.25	-0.16
<i>Q. spinosa</i>	Qs_LBZ_8	220469770	98.51%	216571112	32485666800	28.309	180806940	83.49%	0.30	0.25	-0.21
<i>Q. spinosa</i>	Qs_JZG_2	276390854	98.53%	270799202	40619880300	35.397	226423010	83.61%	0.23	0.25	0.07
<i>Q. spinosa</i>	Qs_ZJH_9	295977651	96.71%	295903288	44385493200	38.679	243173652	82.18%	0.19	0.25	0.21
<i>Q. spinosa</i>	Qs_ZJH_1	218117300	89.33%	236960106	35544015900	30.974	178231826	75.22%	0.19	0.25	0.22
<i>Q. spinosa</i>	Qs_TJS_3	282522152	96.65%	282344292	42351643800	36.906	231936076	82.15%	0.19	0.25	0.21
<i>Q. spinosa</i>	Qs_TJS_5	280101984	97.62%	277038418	41555762700	36.213	229200030	82.73%	0.20	0.25	0.20
<i>Q. spinosa</i>	Qs_BDX_9	235280539	97.96%	232013088	34801963200	30.327	192494590	82.97%	0.19	0.25	0.22
<i>Q. spinosa</i>	Qs_TDY_17	264378723	98.23%	259634468	38945170200	33.938	218603870	84.20%	0.19	0.25	0.22
<i>Q. spinosa</i>	Qs_DJP_22	259259378	98.56%	253609760	38041464000	33.15	211927866	83.56%	0.19	0.25	0.22
<i>Q. spinosa</i>	Qs_WDS_1	297886509	98.98%	290150614	43522592100	37.927	248437828	85.62%	0.19	0.25	0.22
<i>Q. spinosa</i>	Qs_TMS_3	237607961	98.62%	232455222	34868283300	30.385	193138422	83.09%	0.17	0.24	0.31
<i>Q. spinosa</i>	Qs_TMS_8	288664087	98.60%	281800222	42270033300	36.835	237604304	84.32%	0.30	0.25	-0.21
<i>Q. spinosa</i>	Qs_WKS_15	287862925	98.49%	282084352	42312652800	36.872	240303386	85.19%	0.16	0.25	0.34

物种 Species	样本名称 Sample	比对上的 读段数 Mapping read	比对率 Mapping ratio	干净读段 数 Clean read	干净碱基数 Clean base	测序 深度 Depth	正确配对 的读段数 Properly paired read	正确配 对率 Properl y paired ratio	观测杂 合度 H_0	期望 杂合 度 H_E	近交 系数 F_{IS}
<i>Q. spinosa</i>	Qs_SQS_3	266046498	98.08%	261569978	39235496700	34.191	220933326	84.46%	0.14	0.25	0.42
<i>Q. rehderiana</i>	QR_ELS_3	221642850	98.17%	218151744	32722761600	28.516	179656092	82.35%	0.23	0.23	0.26
<i>Q. rehderiana</i>	QR_ELS_5	222105181	97.04%	221452610	33217891500	28.947	181414474	81.92%	0.25	0.25	0.26
<i>Q. rehderiana</i>	QR_FMC_2	226328430	98.93%	221484978	33222746700	28.951	186297812	84.11%	0.29	0.28	-0.02
<i>Q. rehderiana</i>	QR_FMC_4	194829705	99.00%	190487790	28573168500	24.899	161230860	84.64%	0.26	0.28	0.08
<i>Q. rehderiana</i>	QR_JZS_3	242205445	98.98%	236510534	35476580100	30.915	195535056	82.67%	0.24	0.28	0.14
<i>Q. rehderiana</i>	QR_JZS_6	271013887	98.35%	266459386	39968907900	34.83	217987608	81.81%	0.25	0.28	0.11
<i>Q. rehderiana</i>	QR_KKZ_1	220320748	98.49%	215847314	32377097100	28.214	177689442	82.32%	0.29	0.25	-0.17
<i>Q. rehderiana</i>	QR_KKZ_2	221153348	98.45%	216887528	32533129200	28.35	179723832	82.86%	0.31	0.25	-0.25
<i>Q. rehderiana</i>	QR_LS_1	201100892	97.99%	198437394	29765609100	25.939	160857964	81.06%	0.24	0.28	0.16
<i>Q. rehderiana</i>	QR_LS_4	220579690	98.51%	216423288	32463493200	28.29	176911958	81.74%	0.24	0.28	0.16
<i>Q. rehderiana</i>	QR_XYZ_2	215784647	98.42%	211995516	31799327400	27.711	174716342	82.42%	0.28	0.28	-0.01
<i>Q. rehderiana</i>	QR_XYZ_4	223005145	98.45%	218892648	32833897200	28.612	178068940	81.35%	0.28	0.28	0.01
<i>Q. rehderiana</i>	QR_ZR_4	267407317	98.91%	261336050	39200407500	34.16	225552188	86.31%	0.21	0.21	0.26
<i>Q. rehderiana</i>	QR_ZR_7	257252364	97.67%	254630598	38194589700	33.284	212989072	83.65%	0.21	0.21	0.26
<i>Q. semecarpifolia</i>	QM_BMXS_3	234945640	98.97%	230136972	34520545800	30.082	194367604	84.46%	0.23	0.25	0.10
<i>Q. semecarpifolia</i>	QM_BMXS_4	254567163	98.88%	249484824	37422723600	32.611	209152154	83.83%	0.24	0.25	0.04
<i>Q. semecarpifolia</i>	QM_FMC_5	231041671	99.08%	227223308	34083496200	29.701	194134646	85.44%	0.16	0.25	0.35
<i>Q. semecarpifolia</i>	QM_FMC_7	235959194	98.95%	232165640	34824846000	30.347	197453324	85.05%	0.17	0.25	0.33
<i>Q. semecarpifolia</i>	QM_GSL_3	234182027	98.71%	229779116	34466867400	30.035	193704284	84.30%	0.21	0.25	0.16
<i>Q. semecarpifolia</i>	QM_GSL_4	241020575	98.87%	235991350	35398702500	30.847	200137476	84.81%	0.21	0.25	0.15
<i>Q. semecarpifolia</i>	QM_MEX_1	230295208	98.32%	226841394	34026209100	29.651	186140634	82.06%	0.31	0.25	-0.22
<i>Q. semecarpifolia</i>	QM_MEX_2	262833341	98.44%	258298024	38744703600	33.763	215375708	83.38%	0.23	0.25	0.08
<i>Q. semecarpifolia</i>	QM_MLXS_1	211399593	98.98%	206917694	31037654100	27.047	171701436	82.98%	0.23	0.25	0.09
<i>Q. semecarpifolia</i>	QM_MLXS_6	222925366	98.93%	218361040	32754156000	28.543	182741076	83.69%	0.23	0.25	0.10
<i>Q. senescens</i>	QN_ELS_4	216901407	98.48%	213050364	31957554600	27.849	173101872	81.25%	0.25	0.25	0.26

物种 Species	样本名称 Sample	比对上的 读段数 Mapping read	比对率 Mapping ratio	干净读段 数 Clean read	干净碱基数 Clean base	测序 深度 Depth	正确配对 的读段数 Properly paired read	正确配 对率 Properly paired ratio	观测杂 合度 H_0	期望 杂合 度 H_E	近交 系数 F_{IS}
<i>Q. senescens</i>	QN_ELS_6	234459272	98.60%	230008800	34501320000	30.065	192466788	83.68%	0.25	0.25	0.26
<i>Q. senescens</i>	QN_FMC_3	212394545	98.54%	208539058	31280858700	27.259	171379704	82.18%	0.27	0.28	0.05
<i>Q. senescens</i>	QN_FMC_4	226455225	98.90%	221865396	33279809400	29.001	186725044	84.16%	0.25	0.28	0.09
<i>Q. senescens</i>	QN_JZS_3	260078526	96.31%	261123016	39168452400	34.133	202044166	77.38%	0.22	0.22	0.26
<i>Q. senescens</i>	QN_MNP_4	234640910	98.82%	230328434	34549265100	30.107	192131148	83.42%	0.21	0.25	0.15
<i>Q. senescens</i>	QN_MNP_5	237637701	98.58%	233870470	35080570500	30.57	193628760	82.79%	0.24	0.25	0.06
<i>Q. senescens</i>	QN_PMS_2	242147442	98.42%	238303492	35745523800	31.15	196796912	82.58%	0.23	0.23	0.26
<i>Q. senescens</i>	QN_PMS_3	221195522	97.83%	218907298	32836094700	28.614	176801378	80.77%	0.23	0.23	0.26
<i>Q. senescens</i>	QN_XYZ_1	225045491	98.60%	220598790	33089818500	28.835	186865822	84.71%	0.24	0.28	0.14
<i>Q. senescens</i>	QN_ZJY_1	282121239	96.96%	280508618	42076292700	36.667	228407498	81.43%	0.24	0.28	0.15
<i>Q. senescens</i>	QN_ZJY_2	349272991	98.23%	343255934	51488390100	44.868	287736820	83.83%	0.24	0.28	0.15

附表 2 高山栎组植物四个谱系 (P1、P2、P3、外类群 O) 的 D 统计量与 f4 比率, 及其 Z 值、P 值, 以及对应的 BBAA、ABBA、BABA 位点计数; 四个谱系的系统发育关系为有根树: [(P1, P2), P3], O

Appendix Table 2: The D statistic and f4 ratio, with z-score and P-value, as well as BBAA, ABBA, BABA-values that apply to biallelic SNPs across four lineages of *Quercus* sect. *Heterobalanus* in China: P1, P2, P3, and O, related by the rooted tree [(P1, P2), P3], O

P1	P2	P3	Dstatistic	Z-score	p-value	f4-ratio	BBAA	ABBA	BABA
B	A	D	0.04	16.31	2.30E-16	0.21	6027.3	4224.7	3898.39
C	A	D	0.09	19.55	2.30E-16	0.39	5854.32	4656.94	3899.25
B	A	C	0.02	5.43	5.79E-08	0.11	5239.74	4634.52	4454.04
C	B	D	0.05	11.58	2.30E-16	0.22	5733.63	4390.41	3959.03

附表 3 WorldClim version 2.1 中下载的环境变量

Appendix Table 3: Environmental variables were downloaded from WorldClim version 2.1

名称	变量	单位
Name	Variables	Unit
BIO01	Annual Mean Temperature 年平均气温	(°C)
BIO02	Mean Diurnal Range 平均气温日较差	(°C)
BIO03	Isothermality 等温性	-
BIO04	Temperature Seasonality 气温季节性变动系数	-
BIO05	Max Temperature of Warmest Month 最热月份最高温度	(°C)
BIO06	Min Temperature of Coldest Month 最冷月份最低温度	(°C)
BIO07	Temperature Annual Range 最冷月份最低温度差	(°C)
BIO08	Mean Temperature of Wettest Quarter 最湿季度平均温度	(°C)
BIO09	Mean Temperature of Driest Quarter 最干季度平均温度	(°C)
BIO10	Mean Temperature of Warmest Quarter 最暖季度平均温度	(°C)
BIO11	Mean Temperature of Coldest Quarter 最冷季度平均温度	(°C)
BIO12	Annual Precipitation 年降水量	(mm)
BIO13	Precipitation of Wettest Month 最湿月份降水量	(mm)
BIO14	Precipitation of Driest Month 最干月份降水量	(mm)
BIO15	Precipitation Seasonality 降水量季节性变化	(mm)
BIO16	Precipitation of Wettest Quarter 最干季度降水量	(mm)
BIO17	Precipitation of Driest Quarter 最湿季度降水量	(mm)
BIO18	Precipitation of Warmest Quarter 最暖季度降水量	(mm)
BIO19	Precipitation of Coldest Quarter 最冷季度降水量	(mm)
Srad	Solar radiation 太阳辐射	(kJ m ⁻² day ⁻¹)
Wind	Wind speed 风速	(m s ⁻¹)
Vapor	Water vapor pressure 水汽压	(kPa)

附表 4 在两个或多个谱系对中出现重叠候选基因的功能注释信息

Appendix Table 4: Functional annotation of overlap candidate genes was identified in two or more lineage comparison

基因 ID Gene ID	拟南芥 ID Arabidopsis ID	基因 Gene name	功能 Function	参考 Reference	谱系对 Lineage pairs
evm8.model.Scaffold48.128	AT5G03730	<i>CTR1</i>	regulator of ethylene signaling	Kieber et al., 1993	A vs. B; A vs. C; C vs. D
evm8.model.Scaffold437.14	AT4G37750	<i>ANT</i>	Regulation of the size and shape of leaves	Jung et al., 2025	A vs. C; B vs. C;
evm8.model.Scaffold945.20	AT4G30610	<i>BRS1</i>	response to plant redox regulation and cold stress	Zhang et al., 2021	A vs. B; A vs. C;
evm8.model.Scaffold534.57	AT4G19230	<i>CYP707A1</i>	regulating ABA degradation during drought stress	Eysholdt-Derzso et al., 2024; Cai et al., 2024	B vs. D; C vs. D;
evm8.model.Scaffold85.34	AT4G18910	<i>NIP1;2</i>	involved in As tolerance	Kamiya et al., 2009	A vs. B; A vs. C;
evm8.model.Scaffold270.8	AT4G16760	<i>ACX1</i>	response to drought tolerance	Ebeed et al., 2018	A vs. C; B vs. C;
evm8.model.Scaffold267.11	AT4G00730	<i>ANL2</i>	regulator of arsenic tolerance and anthocyanin biosynthesis	Koor et al., 1999; Navarro et al., 2024	B vs. D; B vs. C;
evm8.model.Scaffold474.12	AT3G62600	<i>ERDJ3B</i>	involved in plant reproduction and response to heat stress	Yamamoto et al., 2020	A vs. D; A vs. B; A vs. C
evm8.model.Scaffold240.27	AT3G03140	<i>PWO1</i>	regulate flower timing	Hohenstatt et al., 2018	A vs. C; B vs. C;
evm8.model.Scaffold199.49	AT2G47900	<i>TLP3</i>	response to drought stress	Bano et al., 2022	A vs. B; A vs. C; B vs. C
evm8.model.Scaffold4.72	AT2G46870	<i>NGA1</i>	regulate organ formation and cell identity	Alvarez et al., 2006	A vs. D; B vs. D;
evm8.model.Scaffold7.78	AT2G46020	<i>BRM</i>	regulation of appropriate water stress responses	Han et al., 2012	A vs. C; C vs. D;
evm8.model.Scaffold500.60	AT2G38050	<i>DWF6</i>	light-regulated development	Li et al., 1996	A vs. C; B vs. C;
evm8.model.Scaffold312.74	AT2G20810	<i>GAUT10</i>	response to drought tolerance	Guo et al., 2023	A vs. C; B vs. C;
evm8.model.Scaffold829.17	AT2G02850	<i>ARPN</i>	Involved in another development and pollination	Dong et al., 2005	A vs. C; B vs. C;
evm8.model.Scaffold448.15	AT2G01660	<i>PDLP6</i>	affect plant growth and starch accumulation	Li et al., 2024	A vs. B; A vs. C;
evm8.model.Scaffold382.46	AT1G75540	<i>BBX21</i>	regulate seedling photomorphogenesis	Xu et al., 2014	A vs. B; A vs. C;
evm8.model.Scaffold540.5	AT1G55900	<i>TIM50</i>	regulate seedling development	Kumar et al., 2012	A vs. B; A vs. C;
evm8.model.Scaffold805.23	AT1G43850	<i>SEU</i>	response to light and temperature	Huai et al., 2018	A vs. C; B vs. C;

基因 ID Gene ID	拟南芥 ID Arabidopsis ID	基因 Gene name	功能 Function	参考 Reference	谱系对 Lineage pairs
evm8.model.Scaffold678.21	AT1G26190	<i>TTM2</i>	regulation of seed germination and pathogen defense responses	Ung et al., 2017; Feng et al., 2023	B vs. D; B vs. C;
evm8.model.Scaffold474.11	AT1G02680	<i>TAF13</i>	contribute to seed development	Lindner et al., 2013	A vs. D; A vs. B; A vs. C
evm8.model.Scaffold3.115	AT1G65430	<i>ARI8</i>	highly expressed in mature pollen stage	Marín, 2005	A vs. D; A vs. C; B vs. D
evm8.model.Scaffold22.99	AT1G17440	<i>NOBIRO6</i>	contribute to environment-responsive root growth	Kimi et al., 2022	A vs. D; A vs. B;
evm8.model.Scaffold48.127	AT2G36490	<i>ROS1</i>	regulation of stress tolerance and pathogen defense	Jin et al., 2025	A vs. D; A vs. C; C vs. D
evm8.model.Scaffold557.36	AT1G47550	<i>SEC3A</i>	regulation of pollen germination	Li et al., 2017	A vs. D; A vs. B;
evm8.model.Scaffold822.19	AT4G18470	<i>SNII</i>	response to DNA damage	Yin et al., 2023	A vs. D; A vs. B;
evm8.model.Scaffold1.215	AT3G17700	<i>CNGC20</i>	response to cold stress	Peng et al., 2024	A vs. B; B vs. D;
evm8.model.Scaffold24.151	AT4G23640	<i>KUP4</i>	contribute to embryonic growth during seed maturation	Rubén Tenorio-Berrío et al., 2018	A vs. B; A vs. C;
evm8.model.Scaffold50.96	AT5G17020	<i>XPO1A</i>	Response to heat and oxidative stress	Wu et al., 2010	A vs. B; A vs. C;
evm8.model.Scaffold97.62	AT1G02400	<i>GA2OX6</i>	regulation of root growth	Kubalová et al., 2025	A vs. B; B vs. D;
evm8.model.Scaffold97.80	AT1G48100	<i>PGX3</i>	response to salt stress	Zheng et al., 2018	A vs. B; B vs. D; B vs. C
evm8.model.Scaffold108.101	AT5G49760	<i>HPCA1</i>	response to drought stress	Chen et al., 2025	A vs. B; A vs. C; B vs. D
evm8.model.Scaffold120.27	AT4G26000	<i>PEP</i>	related to flowering time	Xie et al., 2024	A vs. B; B vs. D;
evm8.model.Scaffold120.34	AT4G29920	<i>AtHSPR</i>	response to salt stress and drought	Yang et al., 2015	A vs. B; B vs. D;
evm8.model.Scaffold131.116	AT5G06250	<i>DPA4</i>	regulator of seed size and leaf margin development	Y. Schroeder et al., 2017	A vs. B; B vs. C
evm8.model.Scaffold188.92	AT4G15900	<i>PRL1</i>	responses to plant hormones, sugars and cold stress	Németh et al., 1998;Zheng et al., 2023	A vs. B; B vs. C
evm8.model.Scaffold199.48	AT5G18670	<i>BAM9</i>	responses to cold stress	Kaplan and Guy, 2005	A vs. B; B vs. C
evm.model.Scaffold208.32	AT3G57230	<i>AGL16</i>	regulate to flowering time, responses to salt stress and drought resistance	Dong et al., 2023	A vs. B; A vs. C
evm8.model.Scaffold246.56	AT2G20470	<i>NDR5</i>	regulate pollen development	Yoon et al., 2021	A vs. B; B vs. D
evm8.model.Scaffold257.5	AT1G23190	<i>PGM3</i>	regulate pollen development	Egli et al., 2010	A vs. B; B vs. D;

基因 ID Gene ID	拟南芥 ID Arabidopsis ID	基因 Gene name	功能 Function	参考 Reference	谱系对 Lineage pairs
evm8.model.Scaffold486.2	AT1G14870	<i>PCR2</i>	increased resistance to heavy metals	Song et al., 2010	A vs. B; C vs. D;
evm8.model.Scaffold522.46	AT2G02160	<i>ATC3H17</i>	regulates salt stress tolerance	Seok et al., 2018	A vs. B; A vs. C;
evm8.model.Scaffold582.23	AT5G17230	<i>PSY</i>	response to light and temperature change	Sun et al., 2025	A vs. B; B vs. C;
evm8.model.Scaffold584.30	AT2G26300	<i>GPA1</i>	responses to drought stress	Chen et al., 2025	A vs. B; B vs. D;
evm8.model.Scaffold994.1	AT1G19330	<i>AFR2</i>	regulation of photoperiodic flowering	Gu et al., 2013	A vs. B; A vs. C;
evm8.model.Scaffold6.75	AT3G14570	<i>GSL04</i>	response to reactive oxygen species stress	Vu et al., 2022	A vs. C; B vs. D; B vs. C
evm8.model.Scaffold12.82	AT4G26100	<i>CK1</i>	regulate blue light signaling	Tan et al., 2013	A vs. C; C vs. D;
evm8.model.Scaffold22.97	AT1G22640	<i>MYB3</i>	response to salt stress	Kim et al., 2022	A vs. C; B vs. D;
evm8.model.Scaffold236.70	AT4G10090	<i>ELP6</i>	regulation of defense response signaling	An et al., 2106	A vs. C; B vs. C;
evm8.model.Scaffold236.71	AT5G02810	<i>PRR7</i>	response to low temperature stress	Kim et al., 2024	A vs. C; B vs. D; B vs. C
evm8.model.Scaffold256.58	AT2G25490	<i>EBF1</i>	response to many stimuli	Hao et al., 2020	A vs. C; B vs. D;
evm8.model.Scaffold350.13	AT5G58300	<i>KOIN</i>	response to cold stress	Zhang et al., 2025	A vs. C; B vs. C; C vs. D
evm8.model.Scaffold360.36	AT1G33240	<i>GTL1</i>	regulate drought resistance	Chen et al., 2025	A vs. C; C vs. D;
evm8.model.Scaffold3.102	AT5G17430	<i>BBM</i>	regulate early embryo development	Boutillier et al., 2002	B vs. D; B vs. C;
evm8.model.Scaffold15.68	AT4G23060	<i>IQD22</i>	response to plant hypoxia	Zhao et al., 2025	B vs. D; C vs. D;
evm8.model.Scaffold77.115	AT3G01470	<i>ATHB1</i>	regulate leaf margin development	Miguel et al., 2019	B vs. D; B vs. C;
evm8.model.Scaffold108.65	AT5G24870	<i>MREL57</i>	response to drought stress and ABA	Dou et al., 2021	B vs. D; B vs. C;
evm8.model.Scaffold181.32	AT5G06600	<i>UBP12</i>	regulate blue light signaling	N. Lindbäck et al., 2022	B vs. D; B vs. C;
evm8.model.Scaffold484.30	AT3G01150	<i>PTB1</i>	involved in pollen germination	Wang & Okamoto, 2009	B vs. D; B vs. C; C vs. D
evm8.model.Scaffold493.15	AT1G78590	<i>NADK3</i>	response to abiotic stress and ABA	Chai et al., 2006	B vs. D; C vs. D;
evm8.model.Scaffold522.49	AT5G20070	<i>NUDX19</i>	response to photooxidative stress	Maruta et al., 2016	B vs. D; C vs. D;
evm8.model.Scaffold616.1	AT3G60800	<i>PAT14</i>	regulation of leaf senescence	Zhao et al., 2016	B vs. D; C vs. D;
evm8.model.Scaffold673.13	AT4G34540	<i>GVS1</i>	regulation of leaf senescence	Lyu et al., 2018	B vs. D; C vs. D;
evm8.model.Scaffold751.16	AT2G16390	<i>DRD1</i>	regulation of leaf senescence	Cho et al., 2016	B vs. D; B vs. C;
evm8.model.Scaffold11.119	AT1G47670	<i>LHT4</i>	involved in reproduction	Foster et al., 2008	B vs. C; C vs. D;
evm8.model.Scaffold15.95	AT4G11650	<i>OSM34</i>	related to the ABA signaling	Park et al., 2021	B vs. C; C vs. D;

附表 5 四个谱系中候选适应性分化基因的功能注释信息

Appendix Table 5: Functional annotation of candidate adaptive differentiative genes was identified in each lineage

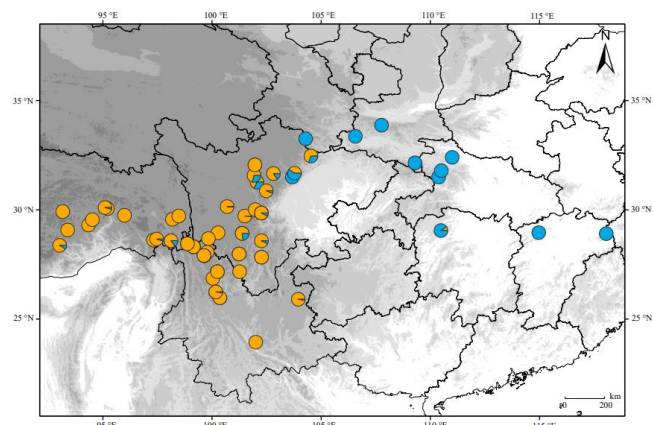
基因 ID Gene ID	拟南芥 ID Arabidopsis ID	基因 Gene name	功能 Function	功能 Reference	谱系 Lineage
evm8.model.Scaffold10.82	AT5G44180	<i>RLT2</i>	regulation of development and flowering time	Li et al., 2012	A
evm8.model.Scaffold11.186	AT1G20670	<i>BRD1</i>	repressors of flowering	Bardani et al., 2023	A
evm8.model.Scaffold34.136	AT5G03620	<i>TSBT4.15</i>	NA	NA	A
evm8.model.Scaffold45.37	AT1G19690	NA	NA	NA	A
evm8.model.Scaffold437.5	AT4G25420	<i>GA5</i>	regulate the balance between plant height, and response to far red light and temperature	Zhang et al., 2025	A
evm8.model.Scaffold619.45	AT2G38080	<i>LAC4</i>	cause plants to be dwarf with little lignin	Zhao et al., 2012	A
evm8.model.Scaffold1.215	AT3G17700	<i>CNGC20</i>	response to cold stress	Peng et al., 2024	A
evm8.model.Scaffold4.149	AT2G46700	<i>CRK3</i>	defense responses	Miyamoto et al., 2019	A
evm8.model.Scaffold11.18	AT1G20180	NA	NA	NA	A
evm8.model.Scaffold11.28	AT1G60420	<i>NRX1</i>	protect plants from oxidative stress	Ho Kang et al., 2020	A
evm8.model.Scaffold12.186	AT5G16560	<i>KAN</i>	regulate the development of reproductive tissues	Hagelthorn et al., 2023	A
evm8.model.Scaffold12.46	AT3G60310	<i>FANCC</i>	NA	NA	A
evm8.model.Scaffold18.31	AT3G13530	<i>MAP3KE1</i>	embryo development	Chaiwongsar et al., 2012	A
evm8.model.Scaffold22.95	AT5G14750	<i>WER</i>	response to prolonged cold	Shimizu et al., 2025	A
evm8.model.Scaffold29.88	AT5G43560	<i>TRAF1A</i>	response to environmental stimuli	Qi et al., 2017	A
evm8.model.Scaffold30.114	AT5G48100	<i>LAC15</i>	regulate the seed color	Zhang et al., 2013	A
evm8.model.Scaffold40.109	AT5G23670	<i>LCB2</i>	response to stress	Kimberlin et al., 2013	A
evm8.model.Scaffold48.94	AT5G03795	NA	NA	NA	A
evm8.model.Scaffold65.112	AT1G19630	<i>CYP722A1</i>	NA	NA	A
evm8.model.Scaffold65.57	AT4G36360	<i>BGAL3</i>	NA	NA	A
evm8.model.Scaffold96.15	AT3G61690	<i>NTP8</i>	NA	NA	A
evm8.model.Scaffold97.62	AT1G02400	<i>GA2OX6</i>	regulation of root growth	Kubalová et al., 2025	A
evm8.model.Scaffold111.93	AT1G62870	NA	NA	NA	A
evm8.model.Scaffold120.34	AT4G29920	<i>AtHSPR</i>	response to salt and drought stress	Yang et al., 2015	A

基因 ID Gene ID	拟南芥 ID Arabidopsis ID	基因 Gene name	功能 Function	功能 Reference	谱系 Lineage
evm8.model.Scaffold129.41	AT1G17160	<i>RBSK</i>	NA	NA	A
evm8.model.Scaffold131.116	AT5G06250	<i>DPA4</i>	regulator of seed size and leaf margin development	Schroeder et al., 2017	A
evm8.model.Scaffold154.8	AT4G32730	<i>MYB3R-1</i>	NA	NA	A
evm8.model.Scaffold159.56	AT4G28530	<i>KIR1</i>	regulate flower development	Gao et al., 2018	A
evm8.model.Scaffold169.36	AT4G11090	<i>TBL23</i>	NA	NA	A
evm8.model.Scaffold220.9	AT5G61540	NA	NA	NA	A
evm8.model.Scaffold226.6	AT1G20110	<i>FREE1</i>	enhance salt tolerance	Liu et al., 2025	A
evm8.model.Scaffold312.42	AT4G03420	NA	NA	NA	A
evm8.model.Scaffold342.80	AT3G05675	NA	NA	NA	A
evm8.model.Scaffold406.8	AT3G47000	NA	NA	NA	A
evm8.model.Scaffold406.9	AT5G58900	<i>DIVI</i>	regulator of seed germination in response to various stress	Li et al., 2024	A
evm8.model.Scaffold437.28	AT1G58520	<i>OSCA2.1</i>	responses to drought stress	Yuan et al., 2014	A
evm8.model.Scaffold503.54	AT5G62165	<i>AGL42</i>	regulate flower senescence	Chen et al., 2022	A
evm8.model.Scaffold519.18	AT1G65880	<i>BZO1</i>	NA	NA	A
evm8.model.Scaffold564.1	AT5G21070	NA	NA	NA	A
evm8.model.Scaffold584.30	AT2G26300	<i>GPA1</i>	responses to drought stress	Chen et al., 2025	A
evm8.model.Scaffold620.10	AT2G16760	NA	NA	NA	A
evm8.model.Scaffold734.11	AT2G33450	<i>PRPL28</i>	NA	NA	A
evm.model.Scaffold750.9	AT3G50380	NA	NA	NA	A
evm8.model.Scaffold1040.11	AT4G31820	<i>ENP</i>	NA	NA	A
evm8.model.Scaffold1040.25	AT2G25080	<i>GPX1</i>	response to oxidative stress	Luo et al., 2016	A
evm8.model.Scaffold2.129	AT1G50030	<i>TOR</i>	regulate embryos development, plant and leaf size	McCready et al., 2020	A
evm8.model.Scaffold4.202	AT4G23980	<i>ARF9</i>	NA	NA	A
evm8.model.Scaffold5.23	AT4G02940	<i>ALKBH10B</i>	response to salt stress and regulate flower timing	Shoaib et al., 2021	A
evm8.model.Scaffold12.50	AT5G56880	NA	NA	NA	A
evm8.model.Scaffold38.29	AT3G11440	<i>MYB65</i>	regulate multiple aspects of reproductive organ development	Anand et al., 2023	A

基因 ID Gene ID	拟南芥 ID Arabidopsis ID	基因 Gene name	功能 Function	功能 Reference	谱系 Lineage
evm8.model.Scaffold53.36	AT4G12080	<i>AHL1</i>	response to oxidative stress	Shi et al., 2024	A
evm8.model.Scaffold58.54	AT1G80340	<i>GA3OX2</i>	modify the appearance of abaxial trichomes as well as leaf shape	Werner et al., 2025	A
evm8.model.Scaffold113.79	AT1G07870	<i>PBL5</i>	NA	NA	A
evm8.model.Scaffold120.4	AT5G57150	<i>BHLH35</i>	NA	NA	A
evm8.model.Scaffold136.11	AT1G07650	<i>LMK1</i>	NA	NA	A
evm8.model.Scaffold225.6	AT1G15750	<i>TPL</i>	involved in defense responses	Causier et al., 2012	A
evm8.model.Scaffold227.66	AT1G76490	<i>HMG1</i>	NA	NA	A
evm8.model.Scaffold279.34	AT3G51620	<i>NTP6</i>	NA	NA	A
evm8.model.Scaffold360.36	AT1G33240	<i>GTL1</i>	regulate drought resistance	Chen et al., 2025	A
evm8.model.Scaffold386.39	AT3G50120	<i>DUF247-1</i>	regulate plant size	Wannitikul et al., 2023	A
evm8.model.Scaffold401.21	AT5G61910	NA	NA	NA	A
evm8.model.Scaffold437.9	AT2G22840	<i>GRF1</i>	responses to cold stress	Lantzouni et al., 2020	A
evm8.model.Scaffold558.30	AT4G27310	<i>BBX28</i>	affect flowering on long days	Liu et al., 2020	A
evm8.model.Scaffold714.14	AT2G02060	NA	NA	NA	A
evm8.model.Scaffold829.17	AT2G02850	<i>ARPN</i>	response to drought stress	Yang et al., 2024	A
evm8.model.Scaffold1003.1	AT4G32280	<i>IAA29</i>	NA	NA	A
evm8.model.Scaffold1.215	AT3G17700	<i>CNGC20</i>	regulate freezing tolerance	Peng et al., 2024	B
evm8.model.Scaffold40.74	AT3G26990	NA	NA	NA	B
evm8.model.Scaffold111.22	AT4G11860	NA	NA	NA	B
evm8.model.Scaffold131.116	AT5G06250	<i>DPA4</i>	regulator of seed size and leaf margin development	Schroeder et al., 2017	B
evm8.model.Scaffold256.32	AT5G11160	<i>APT5</i>	NA	NA	B
evm8.model.Scaffold512.9	AT5G05340	<i>PRX52</i>	NA	NA	B
evm8.model.Scaffold756.36	AT3G28345	<i>ABCB15</i>	NA	NA	B
evm8.model.Scaffold822.19	AT4G18470	<i>SNII</i>	promote the DNA damage responsive, including radiation sensitive	Yin et al., 2023	B
evm8.model.Scaffold19.39	AT1G11790	<i>ADT1</i>	NA	NA	B
evm8.model.Scaffold63.132	AT2G46870	<i>NGAI</i>	regulate organ formation and cell identity	Alvarez et al., 2006	B

基因 ID Gene ID	拟南芥 ID Arabidopsis ID	基因 Gene name	功能 Function	功能 Reference	谱系 Lineage
evm8.model.Scaffold77.115	AT3G01470	<i>ATHB1</i>	regulate leaf margin development	Miguel et al., 2019	B
evm8.model.Scaffold87.41	AT4G14540	<i>NF-YB3</i>	regulate disease resistance	Lin et al., 2024	B
evm8.model.Scaffold154.2	AT2G20120	<i>COV1</i>	NA	NA	B
evm8.model.Scaffold174.94	AT4G24690	<i>NBR1</i>	response to heat and UVB stresses	Wan et al., 2023	B
evm8.model.Scaffold360.37	AT1G33230	NA	NA	NA	B
evm8.model.Scaffold418.41	AT5G62680	<i>NPF2.11</i>	NA	NA	B
evm8.model.Scaffold429.57	AT5G22770	<i>AP2A1</i>	NA	NA	B
evm8.model.Scaffold484.30	AT3G01150	<i>PTB1</i>	involved in pollen germination	Wang & Okamoto, 2009	B
evm8.model.Scaffold779.10	AT3G23410	<i>FAO3</i>	NA	NA	B
evm8.model.Scaffold981.12	AT3G24503	<i>ALDH2C4</i>	improve the drought resistance	Liu et al., 2024	B
evm8.model.Scaffold1161.1	AT4G32030	NA	NA	NA	B
evm8.model.Scaffold1188.2	AT1G71220	<i>UGGT</i>	response to heat and salt stress	Blanco-Herrera et al., 2015	B
evm8.model.Scaffold3.119	AT2G16950	<i>TRN1</i>	NA	NA	B
evm8.model.Scaffold4.172	AT3G61860	<i>RS31</i>	response to cold and red light	Köster et al., 2025	B
evm8.model.Scaffold61.37	AT2G27900	NA	NA	NA	B
evm8.model.Scaffold167.31	AT3G18520	<i>HDA15</i>	response to salt stress	Tilak et al., 2023	B
evm8.model.Scaffold270.8	AT4G16760	<i>ACX1</i>	response to drought tolerance	Ebeed et al., 2018	B
evm8.model.Scaffold759.27	AT2G36830	<i>GAMMA-TIP</i>	NA	NA	B
evm8.model.Scaffold968.16	AT4G16760	<i>ACX1</i>	response to drought tolerance	Ebeed et al., 2018	B
evm8.model.Scaffold42.82	AT1G67260	<i>TCPI</i>	NA	NA	C
evm8.model.Scaffold120.4	AT5G57150	<i>BHLH35</i>	NA	NA	C
evm8.model.Scaffold968.16	AT4G16760	<i>ACX1</i>	response to drought tolerance	Ebeed et al., 2018	C
evm8.model.Scaffold4.126	AT2G46770	<i>NST1</i>	regulate the secondary cell wall thickenings in the anther endothecium	Mitsuda et al. 2005	C
evm8.model.Scaffold97.121	AT5G30495	NA	NA	NA	C
evm8.model.Scaffold427.22	AT3G26935	NA	NA	NA	C
evm8.model.Scaffold1175.19	AT5G56180	<i>ARP8</i>	NA	NA	D

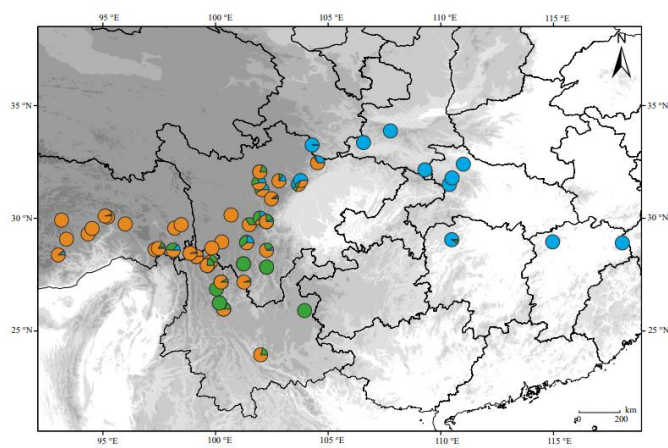
基因 ID Gene ID	拟南芥 ID Arabidopsis ID	基因 Gene name	功能 Function	功能 Reference	谱系 Lineage
evm8.model.Scaffold11.6	AT1G20160	<i>ATSBT5</i>	NA	NA	D
evm8.model.Scaffold14.28	AT3G21350	<i>MED6</i>	NA	NA	D
evm8.model.Scaffold15.68	AT4G23060	<i>IQD22</i>	response to plant hypoxia	Zhao et al., 2025	D
evm8.model.Scaffold48.108	AT5G03770	<i>KDTA</i>	NA	NA	D
evm8.model.Scaffold107.32	AT2G20900	<i>DGK5</i>	response to water and salt stress	Li et al., 2024	D
evm8.model.Scaffold116.18	AT1G22430	NA	NA	NA	D
evm8.model.Scaffold120.34	AT4G29920	<i>AtHSPR</i>	response to salt and drought stress	Yang et al., 2015	D
evm8.model.Scaffold120.47	AT4G33170	NA	NA	NA	D
evm8.model.Scaffold149.54	AT1G32240	<i>KAN2</i>	NA	NA	D
evm8.model.Scaffold246.56	AT2G20470	<i>NDR5</i>	regulate pollen development	Yoon et al., 2021	D
evm8.model.Scaffold258.43	AT4G30996	<i>NKS1</i>	NA	NA	D
evm8.model.Scaffold673.13	AT4G34540	<i>GVS1</i>	regulation of leaf senescence	Lyu et al., 2018	D
evm8.model.Scaffold981.12	AT3G24503	<i>ALDH2C4</i>	improves the drought resistance	Liu et al., 2024	D
evm8.model.Scaffold11.119	AT1G47670	<i>LHT4</i>	involved in reproduction, specifically in anther and pollen development	Foster et al., 2008	D
evm8.model.Scaffold795.13	AT5G12920	NA	NA	NA	D



附图 1 高山栎组植物在 K = 2 时的群体遗传结构

Appendix Figure 1: Population genetic structure and geographical distribution of *Quercus* sect.

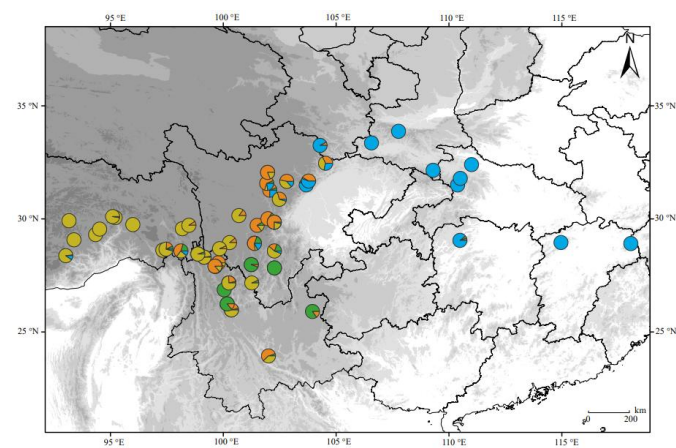
Heterobalanus at K = 2



附图 2 高山栎组植物在 K = 3 时的群体遗传结构

Appendix Figure 2: Population genetic structure and geographical distribution of *Quercus* sect.

Heterobalanus at K = 3



附图 3 高山栎组植物在最优值 K = 4 时的群体遗传结构

Appendix Figure 3: Population genetic structure and geographical distribution of *Quercus* sect.

Heterobalanus at K = 4

导师简介

杜芳，女，1981年11月生，甘肃甘南人。2010年6月获得兰州大学生态学博士学位，2010年12月获得法国波尔多第一大学生态学博士学位。现任北京林业大学生态与自然保护学院教授。

主要致力于森林树种遗传多样性及物种形成的机理研究。主持国家自然科学基金项目五项，参与国家自然科学基金重点项目和科技部重点研发计划各一项。以第一作者或通讯作者发表SCI论文41篇，发明专利1项，论文总计被引用1400余次（谷歌学术统计）。2014年在北京林业大学搭建了首个分子生态实验室，为从微观角度解决宏观森林生态学问题搭建了良好的平台。

2009年获法国农业科学学院三大科研进展奖励；2014年获得教育部高等学校科学研究优秀成果奖自然科学奖一等奖（排名第四）；2014、2018年获得梁希青年论文奖；2015年当选北京市科技新星；2019年起担任中国植物学会女科学家协会委员；2023年起担任中国植物学会委员。2021年开始担任SCI期刊*Ecology and Evolution*的Associate Editor；2022年联合华东师范大学的王嵘教授与日本和加拿大的学者一起组织了“森林树种保护基因组学（Forest Tree Conservation Genomics）”的专刊；2023年开始担任*Plant Diversity*的Associate Editor。

获得成果目录清单

发表论文:

- Luo, Y., Qin, W., Yan, Y., Yin, K., Zang, R., & Du, F. K. (2024). Climate change vulnerability and conservation strategies for tertiary relict tree species: Insights from landscape genomics of *Taxus cuspidata*. *Evolutionary Applications*, 17(9), e13686.
- Luo, Y., Qin, W., Lan, B., Yin, K., Ru, D., & Du, F. K. Xu. Out of Tibet: How adaptive introgression facilitates evergreen oaks' migration. (*Resubmit to New phytologist*)
- Chen, X., Luo, Y., Wang, R., & Du, F. K. (2023). The distinct fruit size and physical defense promote divergent secondary seed dispersal strategies of three oak species. *Forest Ecology and Management*, 529, 120642.
- Zhang, Y., Luo, Y., Qi, M., Li, Y., & Du, F. K. (2025). Leaf morphological trait integration and modularity provide insights into ecological adaptation in evergreen oaks. *Forest Ecosystems*, 14, 100350.
- Wang, R., Luo, Y., Qi, M., Zhang, Y., Zhang, J., Luo, Y., & Du, F. K. (2026). Fine-scale spatial genetic structure and leaf shape variation in five Fagaceae species: insights into conservation and adaptation. *Ecology and Evolution*, 16(2), e72863.