

分类号

密级

UDC



北京林业大学

学术型硕士学位论文

基于全基因组重测序的川滇高山栎和刺叶高山栎
生态适应研究

Ecological adaptation study of *Quercus aquifolioides* and *Quercus spinosa*
based on whole-genome resequencing

刘可可

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

二〇二二年 六月 八 日

独创性声明

本人声明所提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： 可可 日期： 2022年6月8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： 可可 导师签名： 杜芳 日期： 2022年6月8日

答辩委员会成员信息

姓名		职称	工作单位
主席	张乃莉	副教授	北京林业大学
委员	李景文	教授	北京林业大学
	侯继华	副教授	北京林业大学

摘要

物种的适应性遗传变异反应了生态适应的模式和过程,理解这种模式和过程有助于我们追溯物种的进化历程并探讨其未来命运。种群和景观基因组学方法可通过检测种群遗传多样性、遗传结构、识别基因组异常位点并确定这些适应性位点的功能,并将基因型和景观特征联系起来揭示适应性选择信号和适应潜力,是现在生态适应研究中的最新方法和手段。

本论文以青藏高原(Qinghai-Tibet Plateau, QTP)及其临近地区和中国东部温暖低地广泛分布且具有重要生态作用的两种常绿栎属植物:川滇高山栎(*Quercus aquifolioides* Rehd. et Wils.)和刺叶高山栎(*Quercus spinosa* David ex Franchet)为研究材料进行适应性遗传变异研究。选取代表其不同进化谱系的24个川滇高山栎样品(青藏高原的7个种群8个个体;横断山脉(HDM)的11个种群16个个体)和25个刺叶高山栎样品(横断山脉的9个种群13个个体;秦岭山脉(QM)和中国东部温暖低地的9个种群12个个体)进行全基因组重测序,测序深度为24.160X-47.371X。经过质控过滤后,最终分别获得358661个和346281个单核苷酸多态性(Single nucleotide polymorphism, SNP)位点。利用种群基因组方法研究两种栎树的遗传多样性和分化,利用景观基因组学手段研究两种栎树遗传变异的响应模式、与生态适应相关的基因组特征以及气候变化下的未来命运。

种群基因组学研究表明川滇高山栎和刺叶高山栎均可根据地理位置的不同分为东西遗传谱系,西部谱系包括川滇高山栎的*Q. aqu-West*谱系和刺叶高山栎的*Q. spi-West*谱系,东部谱系包括川滇高山栎的*Q. aqu-East*谱系和刺叶高山栎的*Q. spi-East*谱系。两个物种西部谱系的遗传分化程度均高于东部谱系,且遗传多样性均低于东部谱系。以上结果与基于中性遗传标记的核微卫星(nSSR)数据结果一致。基于核微卫星(nSSR)数据的种群遗传学研究表明川滇高山栎分化时间早于刺叶高山栎,随后两物种又逐渐发生了种内遗传分化;空间遗传多样性分析揭示了川滇高山栎和刺叶高山栎的优先保护区不同,川滇高山栎优先保护区域应位于横断山地区(HDM, *Q. aqu-East*谱系),而刺叶高山栎优先保护区位于秦岭地区(QM, *Q. spi-East*谱系)。

景观基因组学研究表明川滇高山栎与刺叶高山栎谱系间的遗传变异具有不同响

应模式，在 *Q. spi-East* 谱系发现了显著的地理隔离(IBD)模式，而在 *Q. aqu-East* 和 *Q. spi-West* 谱系发现了显著的环境隔离(IBE)模式，同时年温度范围(bio07)和三月份降水(perc03)分别在塑造 *Q. aqu-East* 和 *Q. spi-West* 谱系种群遗传变异中起着重要作用。基于遗传分化和基因型环境关联的异常位点挑选方法在川滇高山栎和刺叶高山栎的基因组中发现了自然选择的遗传印记，揭示了不同选择压力促使着两种栎树的生态适应。在未来气候变化下，川滇高山栎和刺叶高山栎需要进行优先保护的区域并不相同，川滇高山栎的基因组脆弱性区域位于其分布范围的北部及西藏地区，而刺叶高山栎则在其北部边缘地区。

本研究通过全基因组重测序方法揭示了两种常绿栎树生态适应的遗传机制，为青藏高原地区栎属物种的未来管理和保护策略提供了理论依据，同时种群和景观基因组学方法的结合为理解气候变化下的保护策略和资源管理提供了参考。

关键词：物种保护，生态适应，未来预测，川滇高山栎，刺叶高山栎

Ecological adaptation study of *Quercus aquifolioides* and *Quercus spinosa* based on whole-genome resequencing

Master Candidate: Keke Liu

(Ecology)

Directed by Fang Du

Abstract

The adaptive genetic variation of population reflects the pattern and process of species' ecological adaptation, understanding this pattern and process can help us trace the evolution of species and discuss their future fate. Population and landscape genomics can detect genetic diversity and structure of population, identify outlier loci in genome, determine the function of these adaptive loci, and finally reveal adaptive selection signals and adaptive potential by associating genotypes with landscape characteristics.

In the thesis we selected two evergreen oaks, *Quercus aquifolioides* Rehd. et Wils. and *Quercus spinosa* David ex Franchet, as our studymodel to study adaptive genetic variation. The two species are widely distributed in the Qinghai-Tibet Plateau (QTP) and adjacent regions and play important ecological role in above region .

A total of 24 individuals of *Q. aquifolioides* (eight individuals of seven populations from QTP; 16 individuals of 11 populations from Hengduan Mountains (HDM)) and 25 individuals of *Q. spinosa* (13 individuals of nine populations from Hengduan Mountain; 12 individuals of nine populations from Qinling Mountains (QM) and warm lowlands in east China) were sequenced through whole-genome resequencing. The sequencing depth range was 24.160X-47.371X. After quality control and filter, we identified 358 661 and 346 281 high-quality single nucleotide polymorphisms (SNPs).

Q. aquifolioides and *Q. spinosa* can be divided into West and East lineages through geographical distribution, and showed a high level of intra-species genetic divergence. We found that the genetic differentiation was higher in West lineages than East lineages, genetic diversity was higher in East lineages than West lineages. The above results were consistent with the results of nuclear microsatellite (nSSR) data based on neutral genetic markers. Population genetics research based on nSSR data showed

that the differentiation of *Q. aquifolioides* was earlier than *Q. spinosa*, then intraspecific genetic divergence might have occurred; Spatial genetic diversity analysis revealed that the priority conservation areas were different for two species: a priority area for conservation should be the populations located at HDM (*Q. aqu*-East lineage) for *Q. aquifolioides*, while a priority area for conservation should be the populations located at QM (*Q. spi*-East lineage) for *Q. spinosa*. The genetic diversity was higher in *Q. aquifolioides* than in *Q. spinosa*, the genetic diversity and differentiation among intraspecific lineages were significantly higher in the West lineages than East lineages.

The results of landscape genomics showed different response patterns of genetic variation among intraspecific lineages. We found a significant isolation by distance (IBD) pattern in the *Q. spi*-East lineage, while a significant isolation by environment (IBE) pattern in *Q. aqu*-East and *Q. spi*-West lineage. Temperature and precipitation played an important role on genetic variation in *Q. aqu*-East and *Q. spi*-West lineage populations, respectively. We found genetic imprints of natural selection in genomes based on F_{ST} and genotype-environment association outlier loci detection, suggesting that different selection pressures promote the ecological adaptation of two oak species. Future prediction and spatial diversity analysis revealed that the priority areas for conservation were different for the two oak species, the genomic vulnerability region of *Q. aquifolioides* were in Tibet and the northern margins of distribution range, and *Q. spinosa* was in northern margins of distribution range.

Our study revealed the genetic mechanism of ecological adaptation of two oak species by whole-genome resequencing. Those findings provide preliminary details for future management strategies and also illustrate how the combination of population and landscape genomic approaches can provide a useful tool for conservation and resource management strategies under climate change.

Key Words: species conservation, ecological adaptation, future prediction, *Quercus aquifolioides*, *Quercus spinosa*

目录

1 引言	1
1.1 生态适应.....	1
1.1.1 生态适应研究进展.....	1
1.1.2 种群遗传学揭示生态适应的遗传机理.....	1
1.1.3 景观基因组学揭示生态适应中遗传与环境的关系.....	2
1.1.4 生态适应研究与树木保护.....	4
1.2 川滇高山栎和刺叶高山栎研究概述.....	5
1.2.1 栎属介绍.....	5
1.2.2 川滇高山栎和刺叶高山栎生态适应研究进展.....	6
1.3 本文的研究内容、研究目的及意义.....	8
1.3.1 研究内容.....	8
1.3.2 研究目的及意义.....	8
2 材料与方法	9
2.1 样品采集.....	9
2.2 SNP 数据的获取	9
2.2.1 基因组 DNA 的提取	9
2.2.2 文库构建与测序.....	9
2.2.3 数据比对和变异检测.....	10
2.2.4 SNP 的质控过滤	10
2.3 实验数据.....	14
2.4 种群基因组学分析.....	14
2.4.1 遗传多样性.....	14
2.4.2 遗传分化.....	15
2.4.3 祖先关系.....	15
2.4.4 基于 F_{ST} 的异常位点检测.....	15
2.5 景观基因组学分析.....	15
2.5.1 气候因子获取与筛选.....	15
2.5.2 基因型-环境关联分析	16
2.5.3 基于环境关联的异常位点检测.....	17
2.5.4 未来命运预测.....	17
3 研究结果	19
3.1 测序及 SNP 鉴定结果	19

3.2 遗传多样性.....	24
3.3 遗传分化.....	27
3.4 遗传变异对环境梯度的响应模式.....	32
3.4.1 Mantel 检验	32
3.4.2 冗余分析.....	36
3.4.3 梯度森林.....	39
3.4.4 广义相异建模.....	40
3.5 异常位点检测结果.....	42
3.6 气候变化下的遗传偏移.....	44
3.6.1 梯度森林模型.....	44
3.6.1 非适应性风险分析.....	46
4 讨论	48
4.1 遗传多样性和遗传结构.....	48
4.2 基于基因型-环境关联的遗传变异的响应模式.....	49
4.3 与生态适应相关的基因组特征.....	50
4.4 物种优先保护区的确定.....	50
4.4.1 遗传多样性的空间分布.....	50
4.4.2 未来气候条件下的基因组脆弱性.....	51
5 结论与展望	53
5.1 结论	53
5.2 展望	54
参考文献	55
附录	66
个人简介	75
导师简介	76
获得成果目录清单	77
致谢	78

1 引言

1.1 生态适应

1.1.1 生态适应研究进展

气候变化正在迅速改变物种所经历的一系列生物和非生物条件(Diffenbaugh *et al.*, 2018), 其产生的选择压力可能超过许多物种的适应潜力(Wilczek *et al.*, 2014), 从而增加物种局部灭绝的风险(Willis *et al.*, 2008)。植物在长期演化过程中会通过遗传变异产生与当地环境适合度最高的基因型, 从而适应当地的生物或非生物条件。在没有其他条件限制的情况下, 这些基因型控制的性状可能对其它生境的适合度没有影响, 产生的这种基因型在其本地栖息地的平均相对适合度高于其它栖息地基因型的模式叫生态适应(ecological adaptation), 也称为局地适应(local adaptation), 本论文统一为生态适应表述(Kawecki and Ebert, 2004)。

理解环境在推动生物多样性空间格局中的作用是进化生物学和生态学的中心目标, 对环境条件的生态适应是这种空间格局的主要来源。生态适应在很大程度上取决于选择和基因流之间的相互作用(Felsenstein, 1976; García-Ramos and Kirkpatrick, 1997; Lenormand, 2002)。当选择相对于基因流较弱时, 生态适应可能会由于迁移负荷(migration load)而受到抑制(Lenormand, 2002); 相反当选择很强时, 可能会在迁移-选择平衡下发生生态适应(Yeaman and Whitlock, 2011)。已有研究采用种源实验(Wang *et al.* 2010; Wilczek *et al.* 2014)、交互移植试验(Etterson and Shaw, 2001)和原位气候控制(Liancourt *et al.*, 2013; Harte *et al.*, 2015)等评估气候变化下的生态适应能力。例如在种源试验中, 研究人员将一个物种分布范围内众多个种群的基因型种植在多个同质园中, 从而评估在不同条件下基因型对环境变化适应的结果(Wang *et al.*, 2010; Wilczek *et al.*, 2014)。在原位气候控制实验中, 将某一基因型移植到模拟气候预测的地点(例如较温暖的低海拔地区), 揭示气候变化如何改变物种的适应性(Etterson and Shaw, 2001)。这些实验提供了很多关于表型遗传基础的信息(Spitze, 1993; Whitlock, 2008), 但关于生态适应程度和遗传基础的信息较少。近年来各种方法和工具为生态适应研究注入了新思路, 新的 DNA 技术有望提供关于生态适应遗传机制的更多信息, 新的统计方法能够更有效地收集这些信息, 新的理论有助于更好的理解这些结果(Savolainen, 2013; Whitlock, 2015; 王天瑞, 2021; Feng and Du, 2022)。

1.1.2 种群遗传学揭示生态适应的遗传机理

种群遗传学(Population genetics)是研究生物种群遗传结构及其在世代间变化规律的科学, 并由此探讨生物的进化机制, 有助于理解关键进化驱动力(如突变、迁移、

遗传漂变和自然选择)对当前遗传多样性分布的贡献。种群遗传学的基本定律由 Hardy (1908)和 Weinberg (1908)分别提出,即哈迪-温伯格定律。其中心内容是:在随机交配的大种群内,若没有选择、突变或迁移因素的作用,基因频率和基因型频率在世代间保持恒定。20世纪30年代随着 Fisher (1930)的《自然选择的遗传理论》、Wright (1931)的《孟德尔种群内的进化》和 Haldane (1932)的《进化的原因》一系列遗传与进化领域著作的发表,将达尔文的自然选择与孟德尔定律结合起来,形成了种群遗传学基本的理论体系。20世纪60年代起开始利用蛋白质、DNA及RNA等大分子与生物进化及种群变异关系进行探讨,从而把生物大分子的结构变化和种群遗传学理论结合起来,初步形成分子种群遗传学理论体系。20世纪90年代,随着DNA分子标记技术与测序技术的快速发展,分子种群遗传研究也随之发展起来。近年来,越来越多的学者通过DNA序列或大量单核苷酸多态性位点(SNP)去研究不同进化力量对全基因组遗传变异的影响,逐渐形成了一个新的与种群遗传学密切相关的研究自然种群中全基因组变异的数量和原因的领域—种群基因组学(Population genomics)。

传统的检测生态适应的分析方法依赖于种群遗传学(Wright and Gaut, 2004),该方法旨在通过比较不同群体的遗传差异,进而发现异常的单核苷酸多态性位点(Single nucleotide polymorphism, SNP),并假设这些异常位点最有可能受到自然选择的影响(Excoffier *et al.*, 2009; Hohenlohe *et al.*, 2010)。随着基因组测序技术的发展,种群基因组学方法变得越来越普遍。目前,已有多种利用遗传分化特异性位点检测生态适应的研究方法(王天瑞等, 2021),比如 BayeScan (Beaumont and Nichols, 1996)和 Fdist2 (Vitalis *et al.*, 2001)。大多数关于生态适应的种群遗传学分析都集中在候选基因上,因为它们具有假定的表型效应,例如一年生植物的开花时间(Le Corre *et al.*, 2002)、动植物的免疫力(Prugnolle *et al.*, 2005; Moeller and Tiffin, 2008)和高海拔适应(Storz *et al.*, 2007; Du *et al.*, 2020)。这些研究阐明了适应性的分子机制,并确定了许多影响表型的基因中哪些基因与生态适应相关。

1.1.3 景观基因组学揭示生态适应中遗传与环境的关系

景观基因组学(Landscape genomics)通过基于全基因组数据和分子标记,旨在揭示基因组的适应性遗传变异与环境异质性的相关性(Holderegger *et al.*, 2006; Manel *et al.*, 2010; Sork and Waits, 2010)。景观基因组学方法可用于量化空间环境变量对基因组差异的影响,并揭示生态适应的基因组遗传基础(Sork *et al.*, 2013; Bragg *et al.*, 2015; Rellstab *et al.*, 2015),从而为资源管理者提供在气候变化下种群管理所需的信息。

可通过线性和非线性关联分析进行环境与遗传变异的相关性研究。线性关系(linear relationships)分析可以通过整合环境变量和空间遗传结构进行分析以评估地理和环境在驱动遗传变异中的贡献(Feng and Du, 2022)。主要以 Mantel 检验(Van Strien *et al.*, 2015)和冗余分析(Redundancy analyses, RDA; Legendre and Legendre, 2012)为主。

Mantel 检验中的距离隔离(Isolation by distance, IBD; Van Strien *et al.*, 2015)描述了当种群之间的扩散受到地理限制时, 遗传变异会发生局部积累产生生态适应(Slatkin, 1993)。同时由距离隔离产生的生态适应减少了自然种群之间的基因流动, 这种基因流动有效速率的适应性降低有助于形成环境隔离(Isolation by environment, IBE)模式(Wang and Summers, 2010; Wang and Bradburd, 2014; Manthey and Moyle, 2015)。为了控制地理因素和环境因素的影响, 研究者也使用偏 Mantel 检验以评估不同生态和进化因素对遗传变异的相对影响。但是鉴于 Mantel 检验的统计能力较低, 错误率较高(Guillot and Rousset, 2013), 研究者通常会使用冗余分析(RDA)来进一步研究特定环境与遗传变异的相关性。冗余分析将遗传数据与特定环境因素相结合进行多变量建模, 并且可以讨论特定环境梯度上的遗传变异分布(Sork *et al.* 2010)。线性关系分析的局限性是无法拟合成分周转率沿环境梯度的变化以及遗传距离与环境 and 地理距离之间的曲线关系。因此, 在使用线性关系评估遗传变异与环境梯度之间的关系时会存在一定误差。非线性关系(Non-linear relationships)的梯度森林(Gradient forest, GF)和广义相异建模(Generalized dissimilarity modeling, GDM)可以将周转函数应用于每个映射的环境变量, 进一步评估环境和空间变量在构建遗传变异中的贡献(Fitzpatrick and Keller, 2015)。梯度森林(GF)作为随机森林的扩展, 是一种能探索空间、环境、等位基因变量之间非线性相关性的非参数、机器学习回归树的方法(Ellis *et al.*, 2012)。该方法使用机器学习算法将生物数据划分为不同的等位基因频率值, 并在每个环境变量的多个分割值处进行划分(Ellis *et al.*, 2012)。梯度森林模型在拟合过程中是隐式的, 不需要变换预测因子, 并且考虑了预测因子之间的复杂交互作用(Fitzpatrick and Keller, 2015)。广义相异建模(GDM)是一种基于距离的矩阵回归模型的非线性扩展方法, 它可以用拟合样条曲线来说明遗传距离与环境、地理距离的非线性关系(Ferrier *et al.*, 2007)。该模型将所有位点对之间物种或遗传组成的差异(生物距离)与它们的环境条件(环境距离)的差异程度联系起来, 进而研究它们之间相关性。

景观基因组学试图通过基因型-环境关联(Genotype-environment associations, GEAs)方法检测选择下的基因组的多态性适应性位点, 并揭示潜在的环境驱动因素(Li *et al.*, 2017)。基因型环境关联分析基于混合效应模型, 可以通过考虑成对遗传距离和种群结构的影响来减少假阳性结果。在混合模型中, 遗传结构被纳入随机因素, 等位基因频率被定义为响应变量, 环境因素被用作固定因素(Feng and Du, 2022)。研究可以通过 BAYENV 或潜在因素混合模型(Latent factor mixed models, LFMM)分析遗传数据和假设驱动选择的环境变量之间的关联来识别适应性位点。基于贝叶斯广义线性混合模型的 BAYENV 被用于在校正种群结构和大小后测试等位基因频率和环境变量之间的相关性, 通过贝叶斯因子(Bayes factor, BF)来识别与环境变量的相关强度(Coop *et al.*, 2010)。LFMM 使用基于主成分分析(Principal component analysis, PCA)变量的分层

贝叶斯混合模型实现了快速算法,通过引入种群结构检测环境和遗传变异之间的相关性(Frichot *et al.*, 2013)。这些方法使用基于个体(而非基于群体)的采样,允许检测不会在种群中产生高度遗传分化的选择性事件,同时为控制中性遗传结构的影响提供了统一的统计框架(Rellstab *et al.*, 2015; Forester *et al.*, 2018)。

快速的气候变化正日益影响全球生态系统,了解物种的适应性遗传多样性对评估物种在未来气候变化下的脆弱性和制定有效的保护策略是必要的。物种适应气候变化的能力既取决于选择所依赖的遗传多样性,也取决于遗传多样性的具体组成,即在未来气候条件下存在有益的等位基因或基因型(Hughes *et al.*, 2008; Wernberg *et al.*, 2018)。尽管许多物种能够适应不断变化的环境,但适应速度依旧被认为远远落后于预测的气候变化速度,从而导致一些物种濒危甚至灭绝(Shaw and Etterson, 2012; Browne *et al.*, 2019),因此迫切需要评估物种在气候变化下的适应能力。如果当前的生态适应性遗传多样性与未来气候条件不匹配,种群会表现出高度的基因组脆弱性(Bay *et al.*, 2018; Gugger *et al.*, 2018; Ruegg *et al.*, 2018; Jia *et al.*, 2020)。适应不匹配或基因组脆弱性是衡量在不断变化的气候条件下基因组变化或进化的程度,这种当前和潜在未来气候之间基因-环境关系的不匹配程度即为遗传偏移(Genetic offset),是对当前和未来气候之间为维持当前观察到的遗传变异和环境变化之间的关系所需的遗传变化幅度的度量(Fitzpatrick and Keller, 2015)。在没有种群迁移的情况下,基因组脆弱性最高区域的种群需要更快的适应性进化以应对快速的气候变化,否则将会面临种群数量减少的风险(Bay *et al.*, 2018)。因此,了解物种中的基因组脆弱性分布对于及时识别最脆弱的种群和适应气候的种群至关重要,有助于确定优先保护的目标及措施(Breed *et al.*, 2019)。梯度森林(GF)和广义相异建模(GDM)可以计算未来气候情景下的遗传偏移。同时为了可视化建模的结果,模型使用主成分分析(PCA)将多个转换后环境变量(即遗传重要性值)的输出减少为多元综合变量。模型中的转换函数可以将当前和未来的气候变量转化为遗传重要性,然后对于景观中的每个网格单元(Grid cell),计算了当前和未来遗传重要性值之间的欧几里德距离(Ellis *et al.*, 2012),该距离用作遗传偏移的度量。为了进一步提高预测的准确性和识别在未来气候变化下易受到威胁的种群,线性回归的非适应性风险分析(Risk of nonadaptedness, RONA)可用于评估环境相关基因座的等位基因频率的平均变化,该等位基因频率将需要对预测的未来环境条件做出反应,根据等位基因频率的变化来评估基因型和环境之间的相关关系(Rellstab *et al.*, 2016)。

1.1.4 生态适应研究与树木保护

森林覆盖了北半球约 30%的陆地表面,包含了约四分之三的全球陆地生物量,与全球碳循环紧密相连(Holliday *et al.*, 2017; Isabel *et al.*, 2020)。固着生长的树木具有世代周期长、较高的遗传多样性和经历快速的微进化过程的特征,通常表现出强烈的生态适应信号(Hamrick *et al.*, 1992; Petit and Hampe, 2006; Neale and Kremer, 2011)。树木

对环境变化的适应主要来自持续的遗传变异，与新突变的产生相比，这种变异更有助于适应气候变化。由于树木生态适应的遗传基础与气候变化、遗传资源保护密切相关，因此，深入研究树木生态适应的遗传基础尤为重要。

最早应用于生态适应研究的种群遗传学方法是通过揭示树木的遗传结构和遗传多样性的分布模式来提供相应的保护措施。后来，随着景观基因组学的发展，学者们利用大量基因位点并通过结合地理和环境信息来了解过去气候对遗传变异的影响程度(Sork *et al.*, 2013), 为物种保护提供了更多的细节。比如, 基于遗传分化(F_{ST})和 GEAs 的异常位点挑选方法, 可以将物种的适应性性状与 SNP 联系起来。Wang 等(2016)对两种杨树的进行 F_{ST} 全基因组扫描, 发现在明显低遗传分化区域有 60 个显著富集的功能类别基因, 大多数与免疫和防御反应、信号转导或凋亡相关。Sork 等(2016)发现了北美峡谷栎(*Quercus lobata*)中 5 个与温度胁迫相关的异常 SNP。McKown 等(2018)发现了毛果杨(*Populus trichocarpa*)中的异常位点与冬季低温和热量累积的不同选择有关, 两种选择都可能受到气候变暖的影响, 这为识别毛果杨适应性遗传变异的空间结构以及制定保护策略奠定了基础。Zhao 等(2019)通过检测银杏(*Ginkgo biloba*)基因组中的选择特征区域, 发现了一些参与昆虫和真菌防御以及对脱水、低温和高盐等非生物胁迫反应的重要基因。随着梯度森林(GF)和非适应性风险分析(RONA)未来预测功能的应用, 可以进一步分析树木在未来环境变化之下遗传变异对环境梯度的响应。Gugger 等(2018)利用 GF 的预测功能发现夏威夷金合欢(*Acacia koa*)在未来气候变化情景下在夏威夷岛的东部表现出最高的遗传偏移, 这可能与该地区的降雨量的减少相关。Jia 等(2020)利用 GF 建模预测了在未来气候条件下侧柏(*Platycladus orientalis*)分布的南部区域将显示出更高的遗传偏移。Du 等(2020)基于 RONA 分析确定了旱季降水(prec01)是影响川滇高山栎(*Quercus aquifolioides*)未来命运最具代表性的环境因素, 同时隔离在分布区东部边缘的川滇高山栎种群在未来气候下可能面临更高的灭绝风险。以上未来预测分析发现在未来气候变化的影响下, 树木对环境的适应能力将发生变化且一些分布范围存在潜在风险。综上, 种群遗传学和景观基因组学的发展为树木进化和适应过程的遗传基础研究提供了实用工具, 为树木未来管理和保护策略提供了理论依据。

1.2 川滇高山栎和刺叶高山栎研究概述

1.2.1 栎属介绍

栎属(*Quercus* L.)是北半球亚热带森林和温带森林中最具多样性和生态重要性的属之一, 也是陆地生物多样性的主要驱动力之一, 大约包括 430-435 种乔木和灌木, 同时也是壳斗科(Fagaceae)中最大的属(Denk *et al.*, 2018; Kremer and Hipp, 2020)。栎属在北半球的分布范围从赤道一直延伸到欧洲北纬 60°地区, 海拔范围从海平面一直延

伸到 4000 米左右(Ferguson, 2006; De Beaulieu and Lamant, 2010)。最新的分类系统将栎属分为八个组,分别为活栎组(*Quercus* sect. *Virentes*)、白栎组(*Quercus* sect. *Quercus*)、红栎组(*Quercus* sect. *Lobatae*)、青冈栎组(*Quercus* sect. *Cyclobalanopsis*)、冬青栎组(*Quercus* sect. *Ilex*)、麻栎组(*Quercus* sect. *Cerris*)、中间栎组(*Quercus* sect. *Protobalanus*)和本都栎组(*Quercus* sect. *Ponticae*) (Denk *et al.*, 2018)。中国作为栎类多样性中心之一(Oldfield and Eastwood, 2007, Denk *et al.*, 2018), 包含有麻栎组、白栎组、冬青栎组和青冈栎组共计约 130 种。据中国国家林业局统计, 栎林占中国天然林总面积的 13%, 分布西起喜马拉雅东至台湾岛, 从中国东北一直延伸到海南岛的亚热带和温带地区(Zhou *et al.*, 1992)

栎属植物之间弱的生殖隔离、频繁的种间基因交换和高表型变异导致了属内物种边界的混乱, 引起了关于物种概念的许多争论(Burger, 1975; Rushton, 1993)。栎属物种间存在频繁的杂交和基因渐渗, 其叶片、壳斗以及种子形态多样可能是对环境变化的适应(Aldrich and Cavender-Bares, 2011)。研究者已经通过对部分栎属植物的杂交渐渗、形态变异等方面研究栎属对环境变化的适应情况。比如 Petit 等(2004)发现无柄栎(*Q. petraea*)可通过长距离花粉扩散与本地种有柄栎(*Q. robur*)种群进行杂交, 杂交个体与无柄栎的反复回交使得无柄栎在当地成功定殖, 表明杂交可以促进后来种的扩散。Leroy 等(2020)通过全基因组测序检测到了无柄栎受到有柄栎的与温度耐受性相关基因的渐渗, 表明适应性渐渗有助于有柄栎对环境的适应。栎属植物叶片形态多样, 叶片形态差异的形成是环境和遗传共同作用的结果(Hipp *et al.*, 2018; Cavender-Bares, 2019; Skelton *et al.*, 2021; Li *et al.*, 2021)。栎属植物对各种生态环境有广泛的适应能力, 研究其分子机制能够为生态适应的机理提供独特见解。比如 Rellstab 等(2016)基于混池测序(pool-seq)在三种欧洲白栎(*Q. petraea*, *Q. pubescens* and *Q. robur*)中发现了 7 个与相同环境因子相关的基因, 揭示了它们生态适应的模式。Sork 等(2016)利用候选基因方法在北美峡谷栎(*Q. lobata*)中发现了与气候相关的功能基因, 揭示了峡谷栎种群气候适应的模式。最近发布的高质量组装的栎属基因组序列(Sork *et al.*, 2016; Plomion *et al.*, 2018)可能为研究栎树适应的基因组印记提供了新的工具。Zhou 等(2022)基于全基因组重测序技术在槲栎(*Q. dentata*)种群中检测到 108 个单核苷酸多态性(SNP)与环境因素显著相关, 其中 12 个 SNP 聚集在约 500 kb 的区域, 同时在该区域中发现了多个正选择信号。Liang 等(2022)利用全基因组重测序技术在三个栎树物种(*Q. acutissima*, *Q. variabilis* and *Q. chenii*)中发现了正选择的信号, 揭示了连锁选择如何促进基因组变异的异质性。

1.2.2 川滇高山栎和刺叶高山栎生态适应研究进展

川滇高山栎(*Quercus aquifolioides* Rehd. et Wils.)和刺叶高山栎(*Quercus spinosa* David ex Franchet)属于栎属冬青栎组(*Quercus* sect. *Ilex* Loudon)植物, 是高山栎类植物

中分布范围最为广泛的两个种。川滇高山栎主要分布在青藏高原(Qinghai-Tibet Plateau, QTP)和横断山脉(Hengduan mountains, HDM), 刺叶高山栎主要分布在横断山脉、秦岭山脉(Qinling mountains, QM)以及中国东部的温暖低地。川滇高山栎和刺叶高山栎为阳性树种, 雌雄同株, 风媒传粉, 具有耐干旱、贫瘠、寒冷的能力; 根系发达, 萌生能力强, 是高海拔地区优良的生态修复树种(陈焕镛和黄成就, 1998; 端木炘, 1994)。由于这两种栎树对各种生态环境和栖息地有广泛的耐受性, 理解其生态适应的分子机理可以帮助我们更好的了解物种的未来命运并为物种的保护和利用奠定理论基础。根据以往研究这两个物种在其分布范围内可根据不同的地理位置分为不同的种内谱系: 刺叶高山栎分为 West 和 East 谱系(Feng *et al.*, 2016; Ju *et al.*, 2019; 王天瑞, 2020), 而川滇高山栎可分为西藏(Tibet)和横断山-川西高原(HDM-WSP)谱系(Du *et al.*, 2017)。最近的研究进一步表明, 气候震荡引发了这两种栎树在寒冷的高地和温暖的低地之间的分化事件, 冷暖谱系的分化时间为 9.3 百万年前(Million years ago, Mya), 此时 HDM 经历了显著的抬升, HDM 的持续抬升和气候变冷触发了栎树的分化, 促使高山栎类植物对高寒环境产生适应(Meng *et al.*, 2017)。

生态位模型(Ecological niche modeling, ENMs)的研究表明川滇高山栎和刺叶高山栎的分布范围相对稳定, 第四纪冰期和间冰期的寒冷气候对高山栎植物的影响较小, 表明这两物种在青藏高原地区的长期生存并不受低温的限制(Feng *et al.*, 2016; Du *et al.*, 2017; Meng *et al.*, 2017; Ju *et al.*, 2019); 但由于土地使用/覆盖和人类影响等空间限制, 其分布范围可能会出现收缩(Liao *et al.*, 2021)。Du 等(2020)运用非适应性风险(RONA)分析的结果表明, 与 HDM-WSP 谱系中的所有其他种群相比, 位于横断山脉最东端的川滇高山栎边缘种群(PW)对旱季降水(prec01)的适应潜力较低, 揭示了隔离在东部边缘的川滇高山栎种群在未来气候下可能面临更高的灭绝风险。另外基于异常位点挑选的方法, 在川滇高山栎的 HDM-WSP 谱系中发现两个基因中的三个 SNP 位点对旱季降水有响应(王玉垚, 2018; Du *et al.*, 2020)。Liu 等(2022)基于核微卫星数据发现了川滇高山栎和刺叶高山栎从青藏高原由西向东迁移的证据, 同时揭示了种内谱系间对环境梯度不同的响应模式。虽然目前对川滇高山栎和刺叶高山栎的进化历史、系统地理模式和潜在分布已有大量研究, 但对其基因组范围内的生态适应分子机制的研究较少。

1.3 本文的研究内容、研究目的及意义

1.3.1 研究内容

本文通过在川滇高山栎和刺叶高山栎的分布范围内进行系统野外调查取样,利用全基因组重测序技术获得大量的基因组数据。通过线性和非线性分析评估遗传变异和环境变量之间的相互关系,以揭示这两种栎树的地理隔离或环境隔离模式,证实地理和气候因素在多大程度上塑造了这两种栎树当前的基因组遗传变异;使用基于遗传分化和基于基因型环境关联的方法进行全基因组扫描寻找具有适应潜力性的异常位点;结合梯度森林(GF)以及非适应风险分析(RONA)的预测功能,评估川滇高山栎和刺叶高山栎种群在未来气候变化下的适应潜力。

1.3.2 研究目的及意义

本研究使用全基因组重测序技术,通过种群遗传学和景观基因组学方法研究川滇高山栎和刺叶高山栎的生态适应的遗传机制并制定相应的保护措施,旨在达到以下目的:

- (1) 研究川滇高山栎和刺叶高山栎遗传多样性和遗传结构;
- (2) 确定地理和气候因素在多大程度上塑造了川滇高山栎和刺叶高山栎当前的基因组遗传变异,以及影响遗传变异最相关的环境变量;
- (3) 鉴定与遗传分化和环境变量相关的异常位点;
- (4) 预测两种常绿栎遗传-环境关联最易受到未来气候变化影响的种群或区域,即在未来气候变化影响下最脆弱的种群或区域。

2 材料与方法

2.1 样品采集

以中国各标本馆的标本记录和各类森林调查结果以及地方植物志为依据,记录川滇高山栎和刺叶高山栎在中国的详细地点。根据地点信息对所有地点进行野外调查采样,记录调查样地内种群大小、结实状况及生长状况。每个种群至少相距 30 km,个体之间的间隔至少为 100 m。为了保证样品可以提取高质量的 DNA,通常选取无病虫害、完整的栎树叶片,统一编号后放入硅胶干燥剂进行干燥处理并置于实验室干燥通风处保存。

本研究采集了 60 个川滇高山栎种群共 996 个个体和 46 个刺叶高山栎种群共 661 个个体,基于 15 对核微卫星(Nuclear microsatellite, nSSR)的基因分型后选取代表其不同进化谱系的 24 个川滇高山栎样品 (*Q. aqu-West*: 8 个, *Q. aqu-East*: 16 个)和 25 个刺叶高山栎 (*Q. spi-West*: 13 个, *Q. spi-East*: 12 个)样品用于后续分析。选取的川滇高山栎和刺叶高山栎样品的分布以及采样点的详细位置信息见图 2.1,黑色和红色虚线分别代表川滇高山栎和刺叶高山栎的分布区域。采样点详细的地点和经纬度信息见表 2.1。

2.2 SNP 数据的获取

2.2.1 基因组 DNA 的提取

本研究对经变色硅胶干燥处理的川滇高山栎和刺叶高山栎种群叶片使用 TIANGEN 植物基因组 DNA 提取试剂盒(天根,中国)提取基因组 DNA,对提取失败或者质量不好的样品使用 CTAB (cetyltrimethylammonium bromide)法提取 DNA (Richards *et al.*, 1994)。将 DNA 样品在 1%琼脂糖凝胶 130V, 200mA 条件下电泳 30 分钟,用凝胶成像仪鉴定 DNA 质量。所有样品的 DNA 在本实验室-20℃下保存。

2.2.2 文库构建与测序

DNA 提取后,根据二代测序(Next generation sequencing, NGS)文库构建要求首先检测 DNA 的浓度, DNA 总量大于 1μg 的样品在 DNBSEQ 平台进行 DNA 小片段文库构建。在 BGISEQ-500 平台对每个合格的文库进行高通量测序,我们采用双端(pair-end) 150bp 的测序策略,测序深度为 30X。测序得到的原始图像数据经碱基识别软件转化为原始序列数据(raw reads),数据以 FASTQ 文件格式存储,称之为 raw data。

使用 SOAPnuke ver. 1.5.4 对原始数据进行质量控制(Chen *et al.*, 2018),具体质量控制原则如下:

- (1) 去除包含测序接头的 reads;
- (2) 去除低质量碱基比例 (base 质量小于或等于 5) 大于 50% 的 reads;
- (3) 去除未知碱基 (N 碱基) 比例大于 10% 的 reads。

2.2.3 数据比对和变异检测

我们将质量控制后得到无错数据(clean data), 应用短序列比对软件 BWA (Burrows-Wheeler Aligner, <http://bio-bwa.sourceforge.net/bwa.shtml>) (Li and Durbin, 2010) 的“mem”算法比对到实验室组装的川滇高山栎参考基因组上(未发表), 比对参数为: -t 4 -M -R, 生成比对结果.sam 文件。然后再用 samtools (<http://samtools.sourceforge.net/>)软件的 sort 工具将.sam 文件转变为.bam 文件。比对结束后, 在进行后续变异检测之前, 需要对.bam 文件进行处理, 标记或删除其中的重复 reads, 同时去除 mapQ 值小于 30 的 reads (过滤参数为: samtools view -q30)。使用 GATK 工具包中的 MarkDuplicates 功能过滤多重比对的 reads, 以减少后续变异检测的错误率。

我们采用 GATK (<https://software.broadinstitute.org/gatk/>; McKenna *et al.*, 2010) 的 UnifiedGenotyper 功能对种群进行 SNP 变异检测, 参数为: -glm BOTH -T -stand_call_conf 50.0 -dcov 1000 -A Coverage -A AlleleBalance。最后生成一个包含所有样本的变异位点和基因型信息的.vcf 文件, 如单核苷酸多态性(Single nucleotide polymorphisms, SNPs)和插入缺失(insertions-deletions, InDels)。我们对所有的变异位点进行硬过滤(hard filter), 参数设置为: -T VariantFiltration -R genome.fa -V AllSample.genotype.raw.snp.vcf.gz --filterExpression "MQ0 >= 4 andand ((MQ0 / (1.0 * DP)) > 0.1)" --filterName "HARD_TO_VALIDATE" --filterExpression "DP < 5 " --filterName "LowCoverage" --filterExpression "QUAL < 30.0 " --filterName "VeryLowQual" --filterExpression "QUAL > 30.0 andand QUAL < 50.0 " --filterName "LowQual" - --filterExpression "QD < 1.5 " --filterName "LowQD" -o AllSample.genotype.filter.snp.vcf.gz。随后使用 vcftools (<https://vcftools.github.io/examples.html>, Danecek *et al.*, 2011)工具将.vcf 文件中的突变拆分成 SNP 和 InDels。

2.2.4 SNP 的质控过滤

低质量的 SNP 会影响后续数据分析结果, 因此在分析之前需要对数据进行过滤。我们采用以下过滤标准:

- (1) 去除 SNP 检出率 > 10%和最小等位基因频率(Minor Allele Frequency, MAF) < 0.05 的 SNP;
- (2) 去除每个个体中偏离哈代-温伯格平衡($P < 0.001$)的 SNP;
- (3) 去除基因型的质量得分(GQ) < 30 的 SNP;

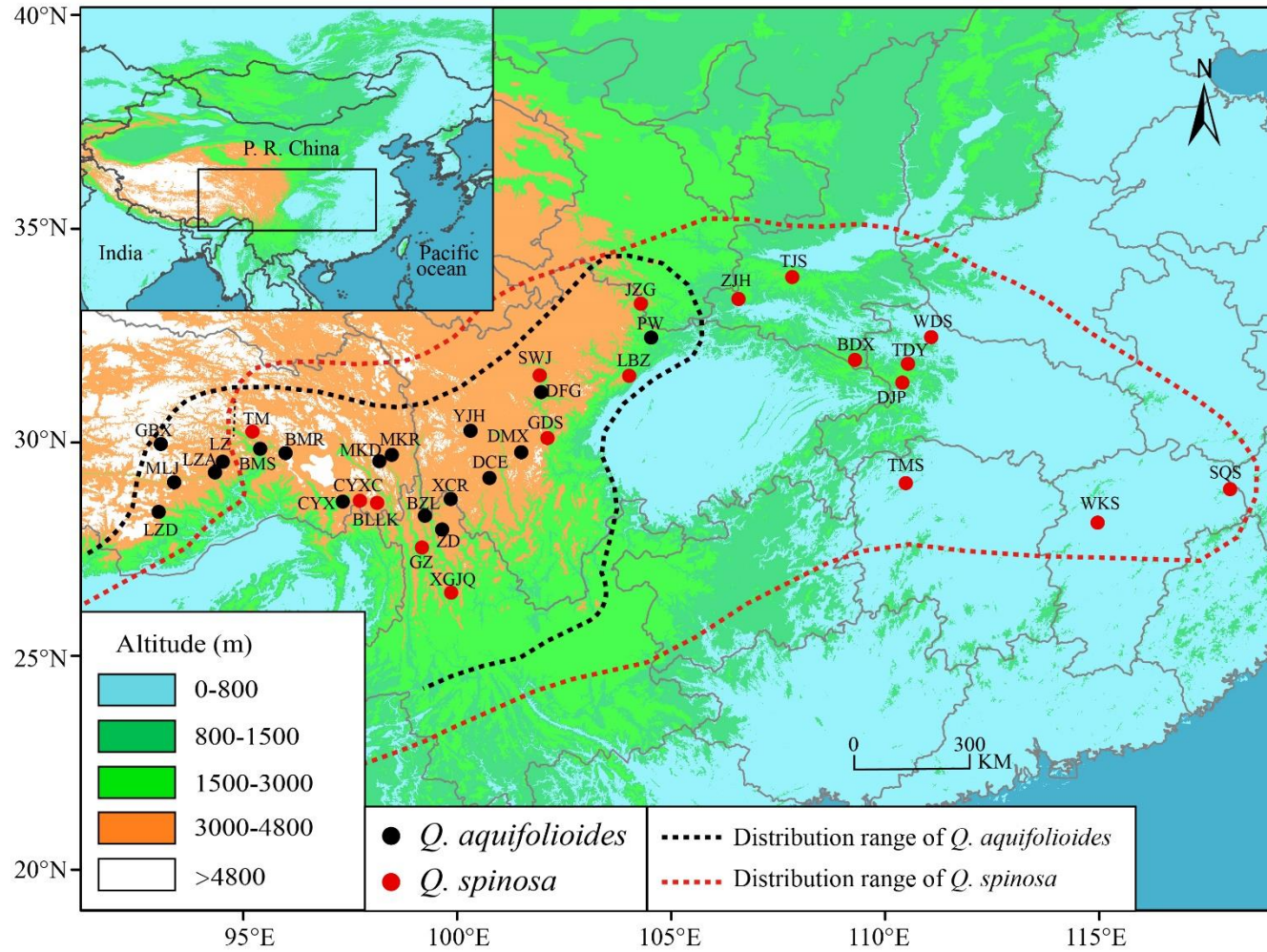


图 2.1 川滇高山栎和刺叶高山栎地理分布区和采样点

Figure 2.1. Geographical distribution and sampling sites of *Q. aquifolioides* and *Q. spinosa*

表 2.1 川滇高山栎和刺叶高山栎样品采集信息表
Table 2.1 Sample information of *Q. aquifolioides* and *Q. spinosa*

种群编号 Pop code	分布地点 Site	样品量 Number	年温度范围(°C) bio07	最潮湿季节的平均 温度(°C) bio08	三月份降水量(mm) prec03	九月份降水量(mm) prec09
川滇高山栎						
LZD	西藏隆子县斗玉乡	1	27.60	15.00	14	88
GBX	西藏工布达江县江达乡	2	29.20	12.90	7	81
MLJ	西藏米林县甲格村	1	27.70	15.55	9	85
LZA	西藏米林县林芝机场	1	28.00	16.40	18	116
LZ	西藏林芝县	1	27.70	11.67	16	102
BMS	西藏波密县索通村	1	28.20	17.28	22	142
BMR	西藏波密然乌镇	1	28.50	12.98	24	125
CYX	西藏察隅县雄久村	2	27.60	18.07	62	207
MKD	西藏芒康县登巴村	1	29.10	7.85	24	81
MKR	西藏芒康县如美镇	1	30.90	10.37	22	81
BZL	云南奔子栏乡	1	25.60	12.02	51	71
ZD	云南中甸	1	27.80	13.13	57	82
XCR	四川乡城县然乌温泉	1	27.30	9.72	28	86
DCE	四川稻城县俄洛村	1	30.90	9.90	20	98
YJH	四川雅江红龙乡	2	31.10	9.52	9	111
DMX	四川康定县明希乡	2	29.50	12.70	11	130
JCK	四川金川县卡撒乡	3	29.90	19.33	17	105
PW	四川平武县	1	29.30	21.15	20	124
刺叶高山栎						
TM	西藏波密县通麦	2	27.60	18.07	21	142
CYXC	西藏察隅县县城	1	27.20	17.22	59	181

材料方法

BLLK	西藏察隅县白隆拉卡	1	27.40	9.03	43	87
GZ	云南省香格里拉高山植物园	1	27.50	12.85	57	83
XGJQ	云南省丽江雄谷景区	3	23.60	17.82	79	92
SWJ	四川金川县撒瓦脚	1	29.30	14.27	16	111
GDS	四川康定县郭达山	1	27.40	17.38	13	127
LBZ	四川汶川县萝卜寨	2	29.10	18.12	26	122
JZG	四川九寨沟县	1	33.30	16.43	21	103
ZJH	陕西勉县张家河林场	2	30.10	17.07	28	137
TJS	陕西周至县铁甲树	2	32.10	16.10	31	131
BDX	陕西平利县八道乡	1	29.80	17.08	49	159
DJP	湖北神农架断疆坪	1	29.50	15.40	63	161
TMS	湖南张家界天门山	2	29.50	22.27	92	111
TDY	湖北神农架土地垭	1	29.10	16.38	60	157
WDS	湖北十堰市武当山	1	29.50	15.40	63	161
WKS	江西武宁县钨矿山	1	32.10	21.88	160	88
SQS	江西玉山县三清山	1	26.70	14.67	193	192

(4) 为了减少强连锁不平衡(LD)区域对 SNP 的影响, 在 PLINK v.2.0 中 (<http://www.cog-genomics.org/plink2/>, Purcell *et al.*, 2007)使用参数“-indep 50 10 0.2”过滤强连锁位点。

2.3 实验数据

本研究使用全基因组重测序数据集为主要研究数据集, 以实验室已有的 60 个川滇高山栎种群共 996 个体和 46 个刺叶高山栎种群共 661 个体的 15 对核微卫星数据为对照数据集(Du *et al.*, 2017; Li *et al.*, 2021)。使用 nSSR 数据集的分析均在文章中特殊说明, 未特殊说明的分析均采用全基因组重测序数据集。

2.4 种群基因组学分析

2.4.1 遗传多样性

我们使用 STACKS (Catchen *et al.*, 2013)软件中的 POPULATIONS 程序计算川滇高山栎和刺叶高山栎的杂合度和多样性参数, 选取多态基因位点百分比(Proportion of polymorphic loci, p)、观测杂合度(Observed heterozygosity, Obs Het)、期望杂合度(Expected heterozygosity, Exp Het)以及香农指数(P_i)来描述种群遗传多样性。使用独立样本 T 检验检验参数的显著性。我们使用 vcftools 基于滑动窗口法计算了川滇高山栎和刺叶高山栎的全基因组范围内核苷酸多态性(π)和 Tajima's D 值, 窗口大小设置为 100 kbp, window-step 参数设置为 10 kbp。

关于 SSR 数据集, 我们采用空间分析方法分别对川滇高山栎和刺叶高山栎分布范围内的空间遗传多样性进行了评估。根据 Van Zonneveld 等(2012)描述的圆形邻域方法(Circular neighborhood methodology)计算并绘制等位基因丰富度和局部共有等位基因的空间分布。等位基因丰富度(Allelic richness, AR)也称为每个基因位点的平均等位基因数, 是旨在选择保护种群的遗传多样性的直接测量(Frankel *et al.*, 1995; Petit *et al.*, 1998)。局部共有等位基因(Locally common alleles, LCA)是在 25% (或更少) 网格单元中频率高于 5% 的等位基因。首先使用 ArcGIS 10.2 在川滇高山栎和刺叶高山栎采样点建立直径为 1° (约 111 m)的缓冲区, 在每一个缓冲区内生成大小为 10 min (相当于研究区域内约 18 m) 的网格单元(grid), 每一个圆形领域大约产生 32 个网格单元, 随后将每一棵树的等位基因组成重新取样(Re-sampled)到所有周围的网格单元, 在 GenAlEx 6 (Peakall and Smouse, 2006)中计算每一个网格单元的 AR 和 LCA 值。圆形邻域方法本身并没有消除原始数据集中已经存在的采样偏差, 这种采样偏差反而对 AR 的影响更为明显。因此, 我们使用 HP-RARE 软件(Kalinowski, 2005)采用稀疏法(Rarefaction methodology)校正等位基因丰富度(AR)计算中的样本偏差, 将 AR 通过稀疏校正到每个网格单元最小样本大小为 10, 删除重新取样后树的数量不足 10 的网格

单元，重新计算每一个网格的 AR 值。

2.4.2 遗传分化

我们基于遗传固定指数(F_{ST})来衡量种群的遗传分化程度。首先在 *vcftools* 中基于滑动窗口法计算了川滇高山栎和刺叶高山栎全基因组范围内的遗传分化差异。其次使用 *ADMIXTURE* 利用马尔可夫链蒙特卡罗(MCMC)算法模拟后验分布，同时估计种群等位基因频率和祖先比例(Alexander *et al.*, 2009)。使用 *vcftools* 将.vcf 格式文件转换为 *ADMIXTURE* 分析需要的.bed 或.ped 格式文件。运行 *ADMIXTURE* 之前使用 *P LINK 2.0* 处理缺失数据，生成 QC 结果文件。定义假定的分支数量 K 为 1-10，交叉验证(CV error)误差最小时对应的 K 值为最优进化枝，最后在 R 语言中作图。*AMOV A* 分析使用 *Arlequin 3.5* (<http://cmpg.unibe.ch/software/arlequin35>, Excoffier and Lischer, 2010)基于 *ADMIXTURE* 的聚类结果计算不同谱系间和种群间的遗传分化程度。

2.4.3 祖先关系

本研究运用 *DIYABC 2.1.0* (Cornuet *et al.*, 2014) 进行近似贝叶斯计算 (Approximate Bayesian computation, ABC) 以确定川滇高山栎和刺叶高山栎种内谱系间的祖先关系。首先根据种内谱系的分化时间先后顺序设置了 8 种不同的进化假说情景 (附录 A)，接着使用与观测数据最接近的 1% 的模拟数据来估计每个场景中参数的后验分布。在对每个情景进行 100 000 次模拟后，进行主成分分析以预先评估模拟数据集和观察到的数据集之间的相似性。然后我们使用对 500 个模拟数据集的直接估计 (Direct estimate) 和对 10 000 个模拟数据集的逻辑回归 (Direct estimate) 来评估每个情景的后验概率。对于最接近的模拟数据的 10 个子集，最合适的情景具有最高的置信区间 (CI) 为 95% 的后验概率。

2.4.4 基于 F_{ST} 的异常位点检测

我们使用 *BayeScan 2.0* 中实现的贝叶斯方法，通过群体之间等位基因频率的不同，从遗传数据中识别出由自然选择的特定位点 (Beaumont and Balding, 2010)。本研究把中性模型的先验概率 prior odds 设置为 10，将后验概率 (Posterior Odds, PO) 设置为 10，后验概率大于 10 的位点意味着受到选择，这种与中性理论下预期的遗传分化 (F_{ST}) 显著不同的位点被作为遗传分化特异性位点 (Foll, 2012)。对川滇高山栎和刺叶高山栎以及各自的种内谱系分别运行 *BayeScan* 已检验与遗传分化相关的异常位点，比较谱系间异常位点的差异，寻找在谱系间共同受到选择的异常位点。

2.5 景观基因组学分析

2.5.1 气候因子获取与筛选

从 WorldClim (<http://worldclim.org>, Fick and Hijmans, 2017)数据库中下载 1970-2000 年分辨率为 30 seconds 的气候因子栅格数据, 31 个气候数据包括 19 个生物气候变量以及 12 个月平均降水变量代表近期采样点的气候数据。IPCC 第五份评估报告采用了四个温室气体浓度情景, 不同代表路径浓度(RCP)按低至高排列分别为 RCP26、RCP45、RCP60 和 RCP85。为了对未来气候作出评估, 本研究在 WorldClim 数据库中下载了 2050 年和 2070 年 RCP26 和 RCP85 碳排放情境下的气候数据, 分别反映了两种不同的极端气候变化条件(Adopted, 2014)。随后在 ArcGIS 中根据每个采样点的经纬度提取现在和未来的气候数据, 然后在 R 语言中使用“usdm” (Naimi *et al.*, 2014)包中的方差膨胀因子(VIF)测试排除了共线性阈值大于 0.7 的气候变量, 删除高度相关的气候因子后共保留四个气候因子用于后续分析, 分别是年温度范围(bio07)、最潮湿季节的平均温度(bio08)、三月份降水量(prec03)和九月份降水量(prec09)。川滇高山栎和刺叶高山栎四个气候因子的详细信息见表 2.1。

2.5.2 基因型-环境关联分析

(1)线性关系分析

我们使用 Mantel 检验量化环境和空间预测因子对遗传变异的影响, 以及评估地理隔离(Isolation by distance, IBD; Van Strien *et al.*, 2015)和环境隔离(Isolation by environment, IBE; Manthey and Moyle, 2015)对遗传变异的影响。首先使用 PGDSpider_2.1.1.0 (<http://cmpg.unibe.ch/software/PGDSpider/>, Lischer *et al.*, 2012)软件将.vcf 文件转成 GENEPOP 格式的文件, 在 genepop 程序中进行转化并生成成对种群间的遗传分化矩阵(F_{ST} 矩阵), 以成对种群间 $F_{ST} / (1-F_{ST})$ 的值衡量遗传距离。通过 R 语言“geosphere”程序包(Hijmans, 2014)将样本经纬度数据转化为群体之间的地理距离矩阵, 以此来衡量地理距离。然后在 R 语言“ecodist”函数包(Goslee *et al.*, 2007)中将下载到现在气候数据整合为一个 Bray-Cutis 矩阵, 以此来衡量环境距离。最后在 R 语言“ecodist”函数包中对地理距离/环境距离矩阵和遗传分化矩阵进行 Mantel 检验以衡量地理隔离/环境隔离的影响。同时为了在衡量地理隔离/环境隔离影响时避免环境/地理距离的影响, 还需要偏 Mantel 检验(Smouse *et al.*, 1986)以控制地理/环境距离的影响。多元回归(MRM)是偏 Mantel 检验分析的扩展, MRM 涉及对任意数量解释矩阵的响应矩阵的多重回归, 可用于研究多元响应距离矩阵和任意数量的解释距离矩阵之间的线性、非线性度量关系(Lichstein, 2007)。本文采用“ecodist”程序包(Goslee and Urban, 2007)进行了距离矩阵的多元线性回归, 以检验遗传距离矩阵与气候距离的多元相关性。Mantel 检验和 MRM 的显著性通过 10 000 次置换检验评估, 显著性水平设置为 0.05。

本研究通过 R 语言中的“LEA”函数包(Frichot and François, 2015)将.vcf 文件转换

成个体水平的等位基因频率数据,将提取的4个气候变量和采样点经纬度整合起来作为冗余分析的环境数据文件。用R语言中的“vegan”函数包(Oksanen *et al.*, 2017)计算一个或多个拟合模型对象的方差(或偏差)并分析每个环境变量的显著性。同时考虑到气候数据可能存在明显的空间差异性,可能与经纬度数据存在强烈的共线性现象,因此需要消除经纬度的影响来进行偏RDA分析(Legendre and Anderson, 1998)。冗余分析和偏冗余分析的置换排列数量设置为999,假设检验显著水平设置为0.05。

(2)非线性关系分析

本研究在“gradientForest”和“extendedForest”程序包下(Ellis *et al.*, 2012)运用梯度森林模拟环境、地理因子对遗传变异的非线性影响。通过综合单因素随机森林分析及交叉验证精度重要性和 R^2 加权重要性值,GF可以得到每一个预测变量单一的功能,代表了沿着预测变量梯度每一个预测变量组成的变化,也可用来检测单个遗传位点对环境梯度的响应模式(Breiman, 2001)。

随后我们使用R语言中的“gdm”函数包(Manion *et al.*, 2018)通过广义相异建模分析检测川滇高山栎和刺叶高山栎遗传变异与气候、地理因子之间的非线性关系。我们还通过拟合样条曲线评估了等位基因组成变化率沿环境梯度的变化(Fitzpatrick and Keller, 2015)。个体间的遗传距离基于等位基因频率计算,地理距离基于采样点经纬度的欧几里得距离。我们通过随机化测试评估变量显著性,并模拟1000次自举迭代来评估由于抽样误差导致的不确定性(Ferrier *et al.*, 2007; Fitzpatrick *et al.*, 2013)。

2.5.3 基于环境关联的异常位点检测

本研究采用潜在因素混合模型(Latent factor mixed models, LFMM, Frichot *et al.*, 2013)检测预测因子和等位基因频率之间的关联,同时估计由于模型混杂效应(Confounding effects)导致的未观察到的潜在因子。我们使用R语言中的“LEA”程序包(Frichot and François, 2015)将.vcf格式的遗传数据转换成后续分析需要的等位基因频率数据,对缺少基因型的数据用“LEA”包中的snmf()函数计算,该计算所涉及的基因型矩阵基于估计的祖先系数和祖先基因型频率。并使用基于贝叶斯马尔科夫链蒙特卡罗(Bayesian Markov chain Monte Carlo, MCMC)的LFMM算法运行十次独立运算,并将iterations设置为100000, burnin设置为50000。使用“LEA”中的lfmm.pvalues()函数计算调整后的 p 值,查看 $k=2$ 是能否控制混杂效应。我们将计算每一个SNP位点的 p 值,将 $p > 0.001$,即 $-\log_{10}p > 3$ 的SNP位点当做异常位点。

2.5.4 未来命运预测

本研究采用两种模型梯度森林(GF)和非适应性风险分析(RONA)预测未来气候变化影响下川滇高山栎和刺叶高山栎脆弱程度分布区域或种群,这些区域或种群由于缺

少迁移和适应性变异而不适应未来气候。

(1) 梯度森林

我们使用梯度森林(GF)预测未来气候变化影响下物种脆弱程度分布区域。首先使用 Maxent 最大熵模型预测了川滇高山栎和刺叶高山栎潜在分布区, 基于物种分布数据和气候数据进行 10 次重复运算, 选择 ROC 值最大的一次运算结果进行预测; 在 ArcGIS 中将 Maxent 中生成的栅格文件去阈值, 将阈值大小设置为 0.5; 将栅格文件转成网格单元数据(grid), 去除不适宜该物种生存的点, 随后在栅格计算器中计算每一个 grid 的经纬度; 最后在 ArcGIS 中提取每一个点的现在和未来情境下的气候数据。我们使用 GF 的 predict()函数来预测川滇高山栎和刺叶高山栎分布区内上所有网格单元的遗传变异, 使用转换函数将当前的气候数据装成遗传重要性值。为了可视化 GF 建模的结果, 我们使用 PCA 将即遗传重要性值的输出结果减少为多变量合成变量。主成分分析以环境变量为中心, 但不按比例进行, 以保持环境变量之间遗传重要性值的差异。最后前三个 PC 分别被分配到红-绿-蓝三种颜色中, 并在地理空间中可视化, 颜色相似性对应于遗传组成预期模式的相似性(Martins *et al.*, 2018)。为了估计气候变化的脆弱性, 我们使用 GF 转换函数将 2050 和 2070 年的两种未来气候情景的数据转换为遗传重要性值; 然后, 我们使用 GF 中的预测函数计算了当前和未来基因组成之间的欧几里得距离, 绘制了由川滇高山栎和刺叶高山栎的预期分布内整个景观的预测结果, 以确定由于气候变化基因与环境关系将受到最严重破坏的地理区域(Fitzpatrick and Keller, 2015)。

(2) 非适应性风险分析

RONA 分析基于 LFMM 的运算产生的等位基因频率数据, 首先基于当前等位基因频率和当前的环境变量建立位点-环境关联的线性模型, 然后使用推断出的线性模型预测 2050 年和 2070 年两种对比代表性浓度途径(RCP26 和 RCP85)情景下未来环境变量的预期等位基因频率, 最后我们计算了当前和未来预期等位基因频率之间的平均差值(RONA), 它估计了种群适应未来气候所需的等位基因频率的平均变化。我们确定 RONA 值大的种群由于缺少适应性遗传变异和迁移而面临着较大的灭绝风险, 并确定了影响川滇高山栎和刺叶高山栎未来命运的最重要的气候因子。

3 研究结果

3.1 测序及 SNP 鉴定结果

经统计川滇高山栎和刺叶高山栎 49 个样品全基因组重测序的测序深度在 24.160X-47.371X 之间, 平均测序深度为 34.330X, 超过 85%的个体的测序深度在 30X 以上(图 3.1)。测序过程中碱基正确识别率在 99%以上(Q20)的碱基数目百分比在 94.23%-97.33%之间, 平均值为 97.71%; 测序数据的 GC 含量在 36.57%-41.12%之间。高测序深度以及高碱基正确识别率保证了测序数据的正确性和可信度, 确保了下游分析结果的可靠性。每个个体产生的原始下机读数(raw reads)的个数在 195 254 690- 368 841 690 条, 每个个体产生的下机原始数据(raw data)在 29.28G-51.53G 之间, 一共产生了 1851.78G 的下机原始数据。下机原始数据经过 SOAPnuke 质量控制后得到的每个个体的无错 reads 个数在 184 831 208-362 399 486 条之间, 每个个体产生的无错数据(clean data)在 25.82G-50.63G 之间, 一共产生了 1797.78G 的 clean data。随后我们将所有的无错数据比对到参考序列, 每个个体参与比对的 reads 在 189 526 594-370 248 570 之间, 比对率在 84.99%-99.27%之间, 除极少数个体的比对率在 90%以下, 其余个体的比对率均高于 90%(表 3.1)。

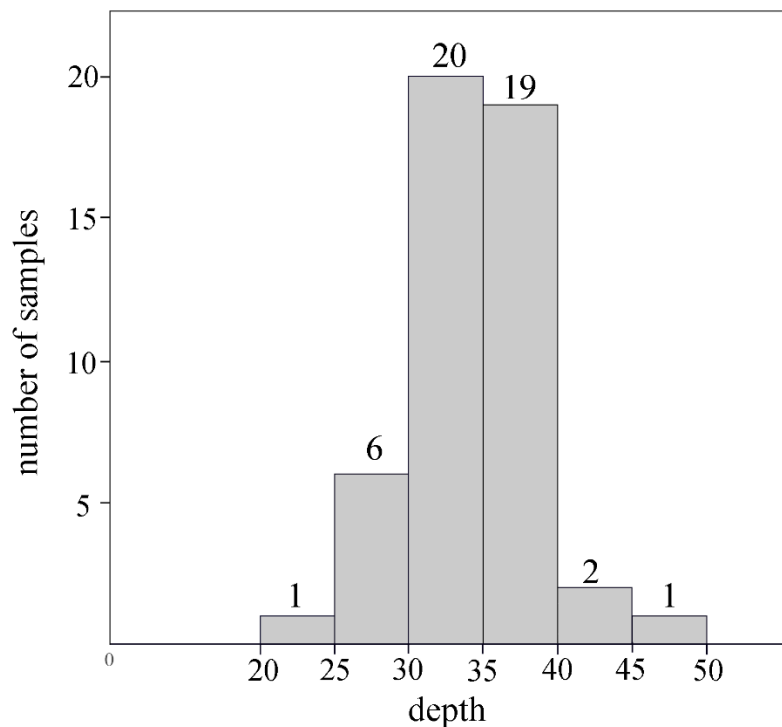


图 3.1 川滇高山栎和刺叶高山栎测序深度统计分布图

Figure 3.1 The statistic distribution of sequencing depth of *Q. aquifolioides* and *Q. spinosa*

表 3.1 川滇高山栎和刺叶高山栎测序和比对结果统计

Table 3.1 Statistics of sequencing and mapping results of *Q. aquifolioides* and *Q. spinosa*

样品名称 Sample ID	测序深度 Depth	原始读数 Raw reads	无错读数 Clean reads	原始数据(bp) Raw data	无错数据(bp) Clean data	Q20 (%)	GC (%)	比对 read Mapping read	比对率 (%) Mapping ratio
Qa_LZD_17	27.819X	219383168	212825472	32907475200	31923820800	95.13	38.38	214637501	97.71
Qa_GBX_2	37.860X	297204112	289635454	44580616800	43445318100	95.41	37.24	293953794	98.34
Qa_GBX_21	39.270X	308811744	300426938	46321761600	45064040700	95.48	38.93	288732832	93.21
Qa_MLJ_5	32.874X	259578274	251491496	38936741100	37723724400	94.88	37.60	255304548	98.39
Qa_LZA_26	26.461X	209963332	202433570	31494499800	30365035500	94.85	37.73	205774543	98.52
Qa_LZ_12	35.861X	297473652	274349092	44621047800	41152363800	95.44	37.64	279542837	98.72
Qa_BMR_9	26.694X	209964678	204214316	31494701700	30632147400	94.87	37.83	208253034	98.84
Qa_BMS_3	33.633X	266201314	257300196	39930197100	38595029400	95.99	37.64	260739646	97.98
Qa_CYX_14	37.016X	293615644	283183644	44042346600	42477546600	95.35	37.61	286702832	97.98
Qa_CYX_15	34.572X	274984176	264487044	41247626400	39673056600	95.42	37.64	268900769	98.39
Qa_MKD_14	30.841X	244212544	235942986	36631881600	35391447900	94.91	37.53	238883057	98.05
Qa_MKR_15	35.711X	284800978	273197366	42720146700	40979604900	95.07	37.46	278985092	99.03
Qa_BZL_19	27.447X	216812616	209978532	32521892400	31496779800	94.72	37.68	213556436	98.18
Qa_ZD_3	43.160X	351915756	330187906	52787363400	49528185900	95.64	37.87	339307657	99.27
Qa_XCR_6	24.160X	195254690	184831208	29288203500	27724681200	95.12	37.26	189526594	99.09
Qa_DCE_8	28.223X	224094552	215914040	33614182800	32387106000	94.98	37.28	221203731	99.12
Qa_YJH_2	38.813X	307237232	296933602	46085584800	44540040300	95.61	37.40	303603641	99.08

研究结果

Qa_YJH_5	32.843X	257969890	251259264	38695483500	37688889600	95.11	37.56	254219146	98.03
Qa_DMx_7	31.492X	247367892	240920886	37105183800	36138132900	94.93	37.53	246823137	99.11
Qa_DMx_8	31.960X	250895674	244506206	37634351100	36675930900	95.44	37.37	249906247	98.85
Qa_JCK_7	36.657X	288653828	280434602	43298074200	42065190300	95.75	37.47	287274376	98.99
Qa_JCK_12	37.912X	300066766	290033146	45010014900	43504971900	95.70	37.71	296268091	98.81
Qa_JCK_14	38.966X	306398242	298097130	45959736300	44714569500	95.15	37.78	304303485	98.66
Qa_PW_6	33.361X	261721410	255220584	39258211500	38283087600	95.09	38.54	236103940	89.83
Qs_TM_1	34.543X	271561268	264265746	40734190200	39639861900	95.16	41.11	230946538	84.99
Qs_TM_3	31.324X	243637638	239637564	36545645700	35945634600	96.20	37.26	243482023	98.21
Qs_CYXC_2	32.384X	254548784	247746774	38182317600	37162016100	95.88	37.66	252158298	98.50
Qs_BLLK_9	37.216X	290214906	284711856	43532235900	42706778400	96.27	37.26	290648860	98.77
Qs_GZ_4	32.845X	256654216	251272332	38498132400	37690849800	96.69	37.21	256704476	98.88
Qs_XGJQ_1	33.559X	265015824	256733246	39752373600	38509986900	96.04	37.42	261197630	98.56
Qs_XGJQ_3	37.341X	293065508	285670630	43959826200	42850594500	96.06	37.43	289949375	98.21
Qs_XGJQ_2	34.333X	270003264	262654124	40500489600	39398118600	95.96	37.20	268489177	98.82
Qs_SWJ_2	40.212X	312399110	307630914	46859866500	46144637100	95.43	36.79	314182966	98.68
Qs_GDS_8	47.371X	368841690	362399486	55326253500	54359922900	97.33	36.97	370248570	98.48
Qs_LBZ_6	37.325X	288579388	285544338	43286908200	42831650700	94.88	36.57	291006406	98.50
Qs_LBZ_8	28.309X	221849504	216571112	33277425600	32485666800	95.43	37.48	220469770	98.51
Qs_JZG_2	35.397X	278015408	270799202	41702311200	40619880300	95.77	37.42	276390854	98.53
Qs_ZJH_9	38.679X	301547228	295903288	45232084200	44385493200	95.60	38.16	295977651	96.71

基于全基因组重测序的川滇高山栎和刺叶高山栎生态适应研究

Qs_ZJH_1	30.974X	241362366	236960106	36204354900	35544015900	94.23	41.12	218117300	89.33
Qs_TJS_3	36.906X	288389614	282344292	43258442100	42351643800	94.88	37.98	282522152	96.65
Qs_TJS_5	36.213X	282522938	277038418	42378440700	41555762700	94.86	37.55	280101984	97.62
Qs_BDX_9	30.327X	237466610	232013088	35619991500	34801963200	94.77	37.99	235280539	97.96
Qs_DJP_22	33.150X	260376016	253609760	39056402400	38041464000	95.00	37.28	259259378	98.56
Qs_TDY_17	33.938X	268269350	259634468	40240402500	38945170200	95.24	37.46	264378723	98.23
Qs_WDS_1	37.927X	300314080	290150614	45047112000	43522592100	95.38	37.31	297886509	98.98
Qs_TMS_3	30.385X	237348352	232455222	35602252800	34868283300	94.76	37.30	237607961	98.62
Qs_TMS_8	36.835X	288423978	281800222	43263596700	42270033300	95.14	37.13	288664087	98.60
Qs_WKS_15	36.872X	291287888	282084352	43693183200	42312652800	95.13	37.60	287862925	98.49
Qs_SQS_3	34.191X	269303486	261569978	40395522900	39235496700	95.34	37.43	266046498	98.08

注 a: Q20: 碱基正确识别率在 99%以上;

注 b: GC: 鸟嘌呤和胞嘧啶所占的比率。

经过变异检测(SNP calling)一共得到 18 039 605 个 SNP 和 2 170 823 个 Indel。在所有的 Indel 中共有 1 073 905 个插入(insertion)和 1 087 088 个缺失(deletion)。在进行严格 SNP 位点过滤后,川滇高山栎和刺叶高山栎分别共保留 358 661 个和 346 281 个 SNP。我们在全基因组水平上对 Indel 和 SNP 位点在每条染色体上的分布进行了统计,外圈的折线图是 SNP 密度分布,内圈代表着 InDel 密度分布,发现 Indel 和 SNP 位点均匀分布在染色体中,红色和绿色分别代表刺叶高山栎和川滇高山栎(图 3.2)。

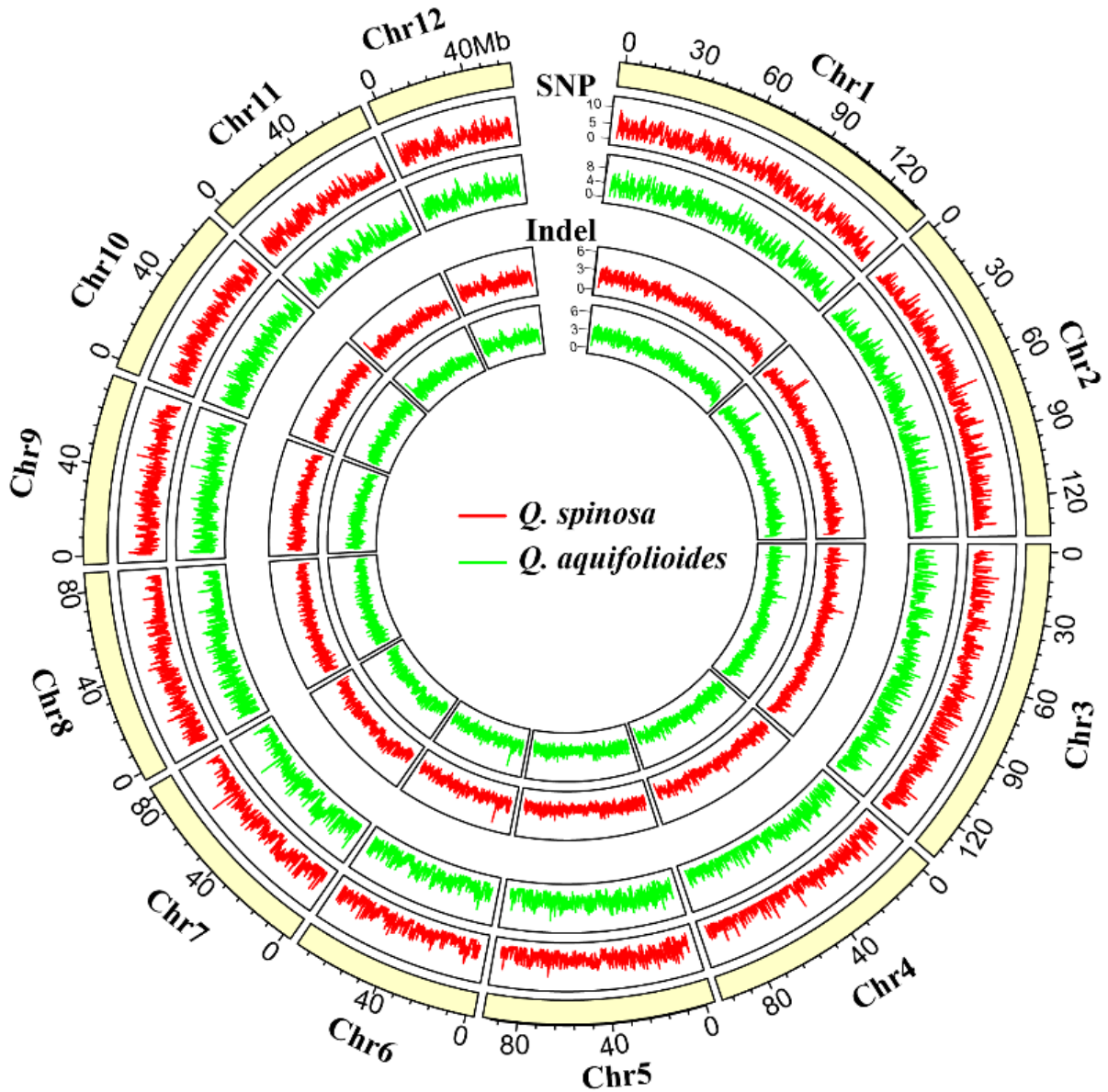


图 3.2 川滇高山栎和刺叶高山栎单核苷酸多态性(SNP) 和 InDels 的密度分布
 Figure 3.2 The density distribution of single nucleotide polymorphisms (SNPs) and InDels for *Q. aquifolioides* and *Q. spinosa*

注 a: Chr1-Chr12 代表栎树假定的 12 条染色体。

3.2 遗传多样性

Populations 程序基于川滇高山栎和刺叶高山栎所有 SNP 数据分别计算的两物种遗传多样性结果显示(表 3.2): 川滇高山栎的遗传多样性水平高于刺叶高山栎(H_O : 0.30 vs. 0.26; H_E : 0.17 vs. 0.15)。我们同样发现 *Q.aqu*-East 谱系的遗传多样性水平高于 *Q.aqu*-West 谱系(H_O : 0.31 vs. 0.28; H_E : 0.18 vs. 0.15); 同时 *Q.spi*-East 谱系的遗传多样性水平高于 *Q.spi*-West(H_O : 0.27 vs. 0.25; H_E : 0.15 vs. 0.14)。基于滑动窗口法计算的川滇高山栎和刺叶高山栎的全基因组范围内核苷酸多态性(π)和 Tajima's D 值也发现了类似的结果(图 3.3): 川滇高山栎的核苷酸多态性(π)和 Tajima's D 值均高于刺叶高山栎。

表 3.2 川滇高山栎和刺叶高山栎 SNP 数据集的遗传多样性估计结果
Table 3.2 Genetic diversity estimates for SNP datasets of *Q. aquifolioides* and *Q. spinosa*

种群编号 Pop code	多态位点比例 p	核苷酸多态性 π	观测杂合度 H_O	期望杂合度 H_E
<i>Q. aquifolioides</i>				
<i>Q.aqu</i> -West lineage				
LZD	0.86	0.28	0.28	0.14
GBX	0.83	0.27	0.29	0.20
MLJ	0.86	0.29	0.29	0.14
LZA	0.86	0.28	0.28	0.14
LZ	0.85	0.30	0.30	0.15
BMS	0.86	0.27	0.27	0.14
BMR	0.87	0.27	0.27	0.13
mean	0.86	0.28	0.28	0.15
<i>Q.aqu</i> -East lineage				
CYX	0.84	0.27	0.28	0.20
MKD	0.83	0.35	0.35	0.17
MKR	0.85	0.31	0.31	0.15
BZL	0.83	0.34	0.34	0.17
ZD	0.85	0.31	0.31	0.15
XCR	0.86	0.29	0.29	0.14
DCE	0.85	0.30	0.30	0.15
YJH	0.82	0.29	0.31	0.22
DMX	0.83	0.28	0.30	0.21
JCK	0.83	0.26	0.27	0.22
PW	0.85	0.30	0.30	0.15
mean	0.84	0.30	0.31	0.18
P	0.021	0.080	0.031	0.008
mean total	0.85	0.29	0.30	0.17

研究结果

<i>Q. spinosa</i>				
<i>Q. spi</i> -West-lineage				
TM	0.87	0.21	0.22	0.16
CYXC	0.88	0.23	0.23	0.12
BLLK	0.88	0.25	0.25	0.12
GZ	0.89	0.22	0.22	0.11
XGJQ	0.87	0.20	0.21	0.16
SWJ	0.87	0.27	0.27	0.13
GDS	0.86	0.28	0.28	0.14
LBZ	0.83	0.28	0.29	0.21
JZG	0.85	0.30	0.30	0.15
mean	0.87	0.25	0.25	0.14
<i>Q. spi</i> -East lineage				
ZJH	0.84	0.26	0.27	0.19
TJS	0.85	0.25	0.28	0.18
BDX	0.86	0.28	0.28	0.14
DJP	0.86	0.28	0.28	0.14
TMS	0.83	0.28	0.29	0.21
TDY	0.86	0.28	0.28	0.14
WDS	0.86	0.27	0.27	0.14
WKS	0.88	0.23	0.23	0.12
SQS	0.90	0.20	0.20	0.10
mean	0.86	0.26	0.27	0.15
<i>P</i>	0.300	0.050	0.040	0.020
mean total	0.86	0.26	0.26	0.15

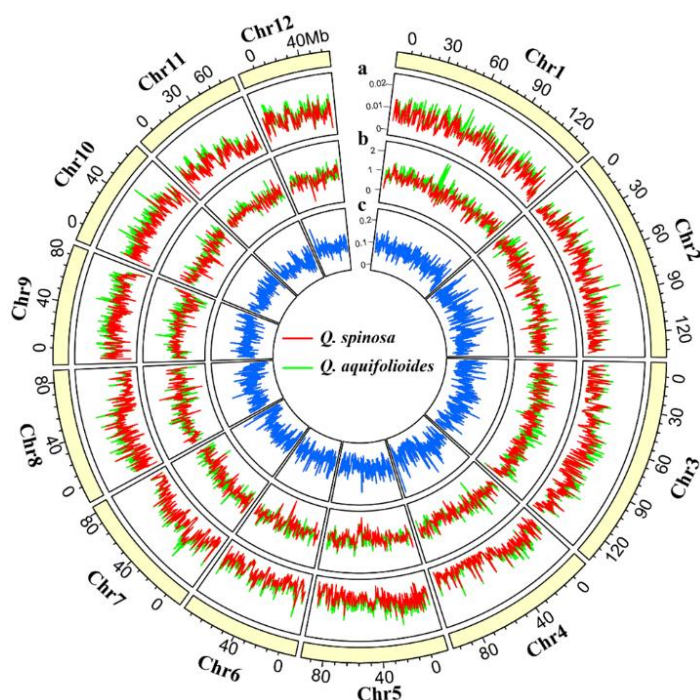


图 3.3 川滇高山栎和刺叶高山栎全基因组多态性模式的圆圈图。圆圈从最外到最内(a): 核苷酸多样性(π), (b) Tajima's D, (c) 固定指数(F_{ST})

Figure 3.3 Circos plot of the genome-wide patterns of polymorphism among *Q. aquifolioides* and *Q. spinosa*. The circles from outermost to innermost (a):Nucleotide diversity (π), (b) Tajima's D, (c) Fixation index (F_{ST})

注 a: Chr1-Chr12 代表栎树假定的 12 条染色体。

我们基于 SSR 数据集采用循环邻域重采样(re-sampling)技术, 以确保空间多样性分析的数据点充分且分布更均匀。重新采样后川滇高山栎和刺叶高山栎分别产生了 31 872 和 21 152 棵树的数据集用于进一步的等位基因丰富度(AR)和局部共有等位基(LCA)的空间分析(附录 B)。空间分析结果表明两种栎树的 AR 和 LCA 富集区域不同。对于川滇高山栎而言, 位于四川省西南部和云南省西北部的横断山脉(HDM)地区的种群(*Q. aqu-East lineage*)的 AR 和 LCA 高于位于青藏高原(QTP)地区的种群(*Q. aqu-West lineage*)。位于横断山脉最东端的川滇高山栎边缘种群 PW, 其 AR 值明显低于横断山脉地区的其他种群(图 3.4)。对于刺叶高山栎而言, 位于秦岭山脉(QM)的种群(*Q. spi-East lineage*)显示出比位于 HDM 的种群(*Q. spi-West lineage*)更高的 AR 和 LCA 值(图 3.4)。

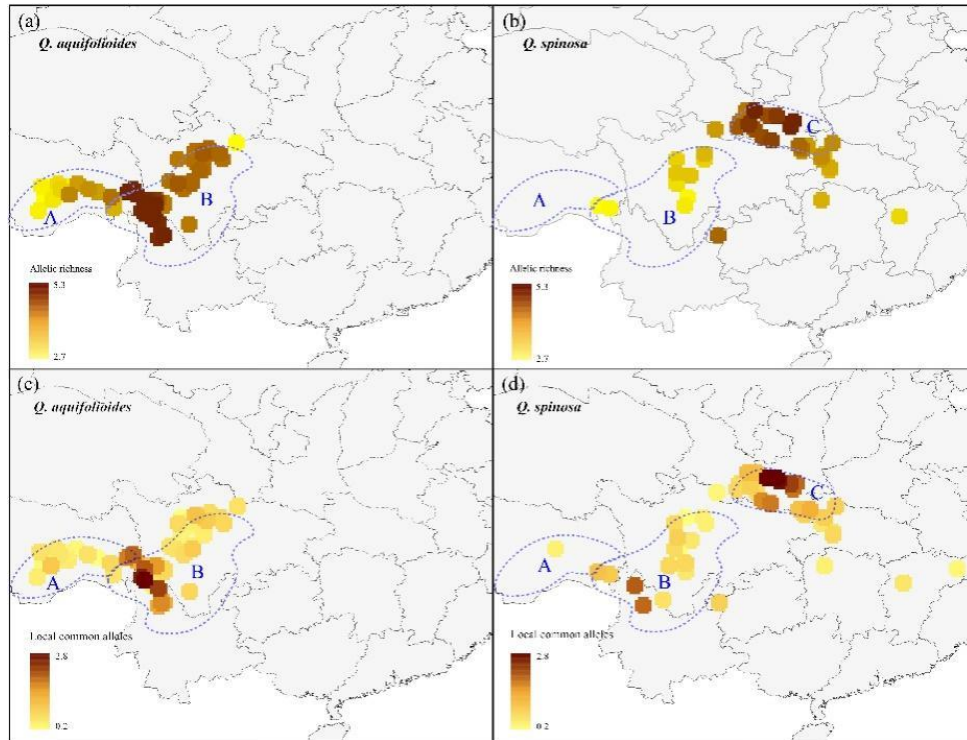


图 3.4 川滇高山栎和刺叶高山栎等位基因丰富度和局部共有等位基因图。

Figure 3.4 The allelic richness and locally common alleles map of *Q. aquifolioides* and *Q. spinosa*.

注：浅蓝色虚线代表三个主要的研究领域：A：青藏高原(QTP)；B：横断山脉(HDM)；C：秦岭(QM)。

3.3 遗传分化

我们用 ADMIXTURE 检测两种栎树的种群遗传结构，并比较不同 K 值的交叉验证错误率(Cross-validation error)，结果显示 K=2 和 K=4 时交叉验证错误率最低，因此 K=2 和 K=4 时 ADMIXTURE 模型的可信度最高。我们绘制了 K=2-4 时的模拟结果(图 3.5)，结果显示：K=2 时所有的个体聚为两支，分别对应着川滇高山栎和刺叶高山栎。K=3 时，刺叶高山栎进一步发生分化为两个谱系，我们根据两个谱系的地理位置命令为 *Q. spi-West* 和 *Q. spi-East* 谱系，其中 *Q. spi-West* 谱系主要分布在横断山脉地区，有 8 个种群 13 个个体；*Q. spi-East* 谱系主要分布在秦岭山脉和中国东部的温暖低地，有 8 个种群 12 个个体。当 K=4 时，川滇高山栎发生了分化，同样也按照地理位置划分为东西两个谱系(*Q. aqu-West* 和 *Q. aqu-East*)，其中 *Q. aqu-West* 谱系主要分布在青藏高原地区，有 7 个种群 8 个个体；*Q. aqu-East* 谱系主要分布在横断山脉地区，有 11 个种群 16 个个体。

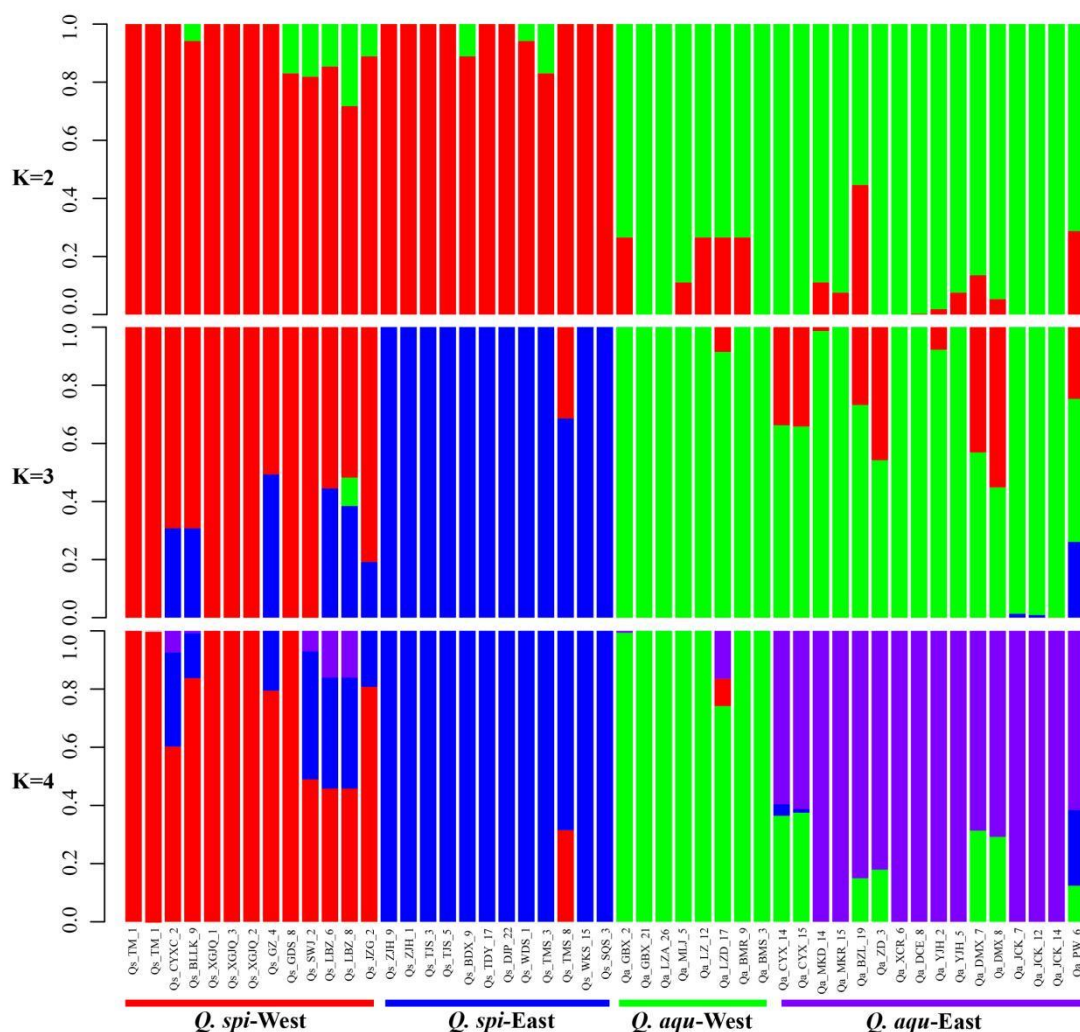


图 3.5 ADMIXTURE 种群遗传结构图。每个条形代表一个个体，条形部分的颜色取决于假设 K=2、K=3 或 K=4 估计的祖先比例

Figure 3.5 Model-based population assignment by ADMIXTURE analysis for K=2 to 4. Each bar represents a single individual, with portions of the bar colored depending on the ancestry proportions estimated assuming K=2, K=3 or K=4

ABC 分析结果表明 8 个情景模型中情景 1 的后验概率最高，表明它是最合适两种栎树进化历史的情景：情景 1 的后验预测误差最低(直接估计：0.167，逻辑回归：0.213)（附录 C, D）。情景 1 表明川滇高山栎是这两种栎树的祖先物种，其首先扩散到青藏高原地区(*Q. aqu-West* lineage)，然后扩散到横断山脉地区(*Q. aqu-East* lineage)。刺叶高山栎在 t2 时起源于川滇高山栎，随后也发生了分化，发生在横断山脉地区(*Q. spi-West* lineage)的分化远早于秦岭地区(*Q. spi-East* lineage)。

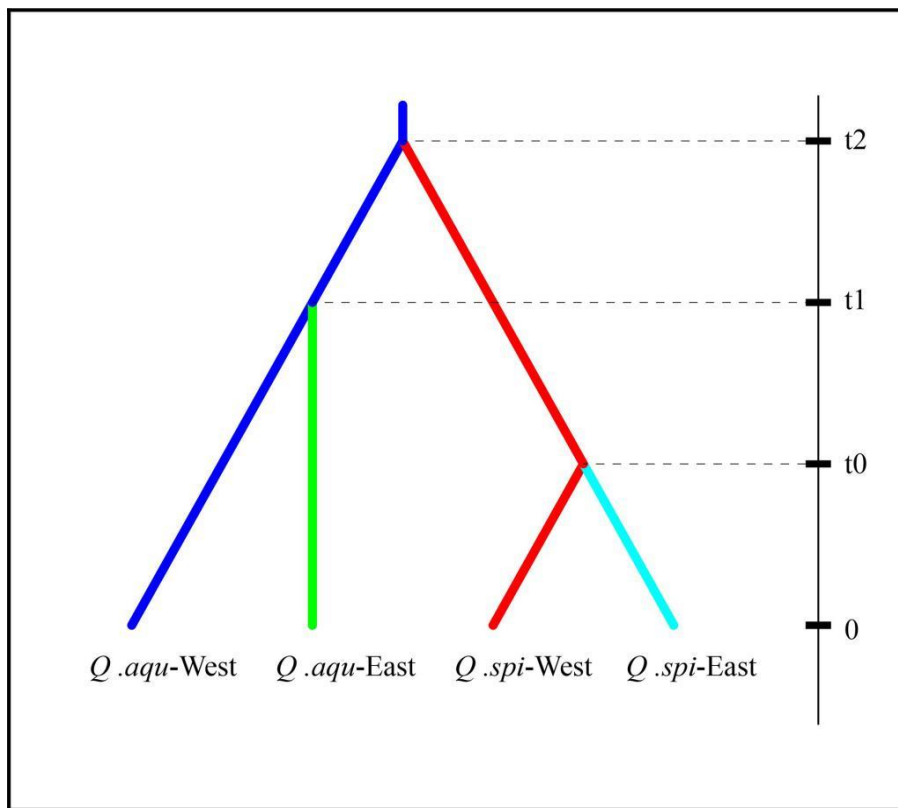


图 3.6 基于 DIYABC 模型推断的川滇高山栎和刺叶高山栎的种群进化历史图
Figure 3.6 The inferred demographic history of *Q. aquifolioides* and *Q. spinosa* with DIYABC

对于川滇高山栎和刺叶高山栎所有种群来说，分层分子方差分析(AMOVA)显示川滇高山栎和刺叶高山栎之间存在高水平的遗传分化，绝大部分的遗传变异发生在种群内($F_{ST} = 0.16$, 83.98%)。刺叶高山栎的遗传分化水平($F_{ST} = 0.21$)要高于川滇高山栎($F_{ST} = 0.07$)。川滇高山栎和刺叶高山栎西谱系的遗传分化程度要比东谱系的高(0.05 vs. 0.03; 0.14 vs. 0.01)，并且不论川滇高山栎还是刺叶高山栎以及各自的种内谱系，遗传变异都主要发生在种群内(表 3.3)。

表 3.3 基于 SNP 数据集的川滇高山栎和刺叶高山栎的分层分子方差分析(AMOVA)结果

Table 3.3 Hierarchical analyses of molecular variance (AMOVA) of *Q. aquifolioides* and *Q. spinosa* populations based on SNPs datasets

变异来源 Source of variation	自由度 Degrees of freedom	平方和 Sum of squares	方差 Variance component	变异百分比 (%) Percentage of variation (%)	固定指数 Fixation indices
All samples					
Among <i>Q. aquifolioides</i> and <i>Q. spinosa</i>	1	5596.04	81.35	5.89	$F_{CT} = 0.16$
Among populations within species	34	52266.42	140.02	10.13	$F_{SC} = 0.11$
Within populations	62	71933.45	1160.22	83.98	$F_{ST} = 0.16$
<i>Q. aquifolioides</i>					
Among lineages	1	3304.79	84.95	5.57	$F_{CT} = 0.06$
Among populations within lineages	16	23783.54	28.6	1.88	$F_{SC} = 0.02$
Within populations	30	42341.83	1411.39	92.55	$F_{ST} = 0.07$
<i>Q. aqu</i>-West lineage					
Among populations	6	13421.25	54.43	2.36	$F_{ST} = 0.05$
Within populations	9	21234	2359.33	97.64	
<i>Q. aqu</i>-East lineage					
Among populations	10	19285.77	59.71	3.28	$F_{ST} = 0.03$
Within populations	21	36926.67	1758.41	96.72	
<i>Q. spinosa</i>					

研究结果

Among lineages	1	6711.07	205.43	13.34	$F_{CT} = 0.13$
Among populations within lineages	16	24380.76	110.16	7.15	$F_{SC} = 0.08$
Within populations	32	39186.08	1224.57	79.51	$F_{ST} = 0.21$
<i>Q. spi</i> -West lineages					
Among populations	8	19367.29	269.39	13.93	$F_{ST} = 0.14$
Within populations	17	28297.17	1664.54	86.07	
<i>Q. spi</i> -East lineages					
Among populations	8	15320.79	25.29	1.29	$F_{ST} = 0.01$
Within populations	15	29722.25	1981.48	98.71	

3.4 遗传变异对环境梯度的响应模式

3.4.1 Mantel 检验

Mantel 检验的结果 (图 3.7, 表 3.4) 表明在川滇高山栎的所有种群(Mantel's $r = 0.36$, $P=0.002$)和刺叶高山栎的所有种群(Mantel's $r = 0.39$, $P < 0.001$)中均检测到了显著的距离隔离的空间遗传模式, 在两种栎树的東西谱系中并没有发现遗传距离和地理距离存在显著的相关性。在刺叶高山栎的所有种群(Mantel's $r = 0.33$, $P = 0.001$)和 *Q. spi-East* 谱系(Mantel's $r = 0.44$, $P = 0.02$)中发现了显著的环境隔离的空间遗传模式, 在 *Q. spi-West* 谱系(Mantel's $r = 0.17$, $P = 0.23$)中并没有发现显著的环境隔离的空间遗传模式。同时在川滇高山栎的所有种群以及种内谱系间均没有发现遗传距离和地理距离存在显著的相关性。

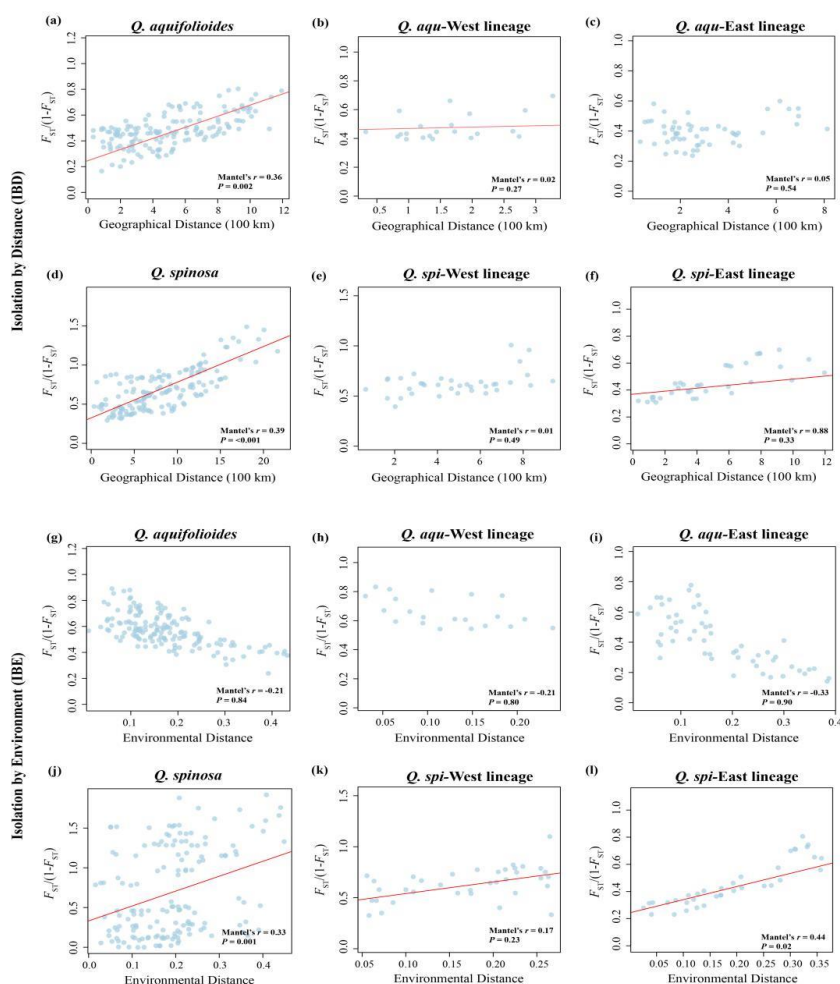


图 3.7 川滇高山栎和刺叶高山栎的遗传距离与地理和环境距离的 Mantel 检验结果
Figure 3.7 Mantel tests of genetic distance against geographic and environmental distances
Q. aquifolioides and *Q. spinosa*

表 3.4 川滇高山栎和刺叶高山栎 Mantel 和偏 Mantel 检验结果
 Table 3.4 Results of Mantel test and partial Mantel test in *Q. aquifolioides* and *Q. spinosa*

	<i>Q. aquifolioides</i>		<i>Q. aqu-West</i> lineage		<i>Q. aqu-East</i> lineage		<i>Q. spinosa</i>		<i>Q. spi-West</i> lineage		<i>Q. spi-East</i> lineage	
	Mantel's <i>r</i>	<i>P</i>	Mantel's <i>r</i>	<i>P</i>	Mantel's <i>r</i>	<i>P</i>	Mantel's <i>r</i>	<i>P</i>	Mantel's <i>r</i>	<i>P</i>	Mantel's <i>r</i>	<i>P</i>
<i>Mantel test</i>												
Isolation by Distance (IBD)	0.36	0.002	0.02	0.27	0.05	0.54	0.39	<0.001	0.01	0.49	0.88	0.33
Isolation by Environment (IBE)	-0.21	0.84	-0.21	0.80	-0.33	0.9	0.33	0.001	0.17	0.23	0.44	0.02
<i>partial Mantel test</i>												
IBD conditioned with environmental distance	0.44	0.004	0.18	0.479	0.14	0.398	0.26	<0.001	-0.02	0.547	0.31	0.03
IBE conditioned with geographical distance	-0.21	0.841	-0.27	0.852	0.35	0.02	0.17	0.011	0.17	0.007	0.079	0.05

注 a: 粗体数字表示数据显著

我们进一步采用控制环境距离或地理距离的影响来检测地理隔离和环境隔离(表 3.4)。偏 Mantel 检验结果显示当控制环境因子的影响时, 在川滇高山栎所有种群(Mantel's $r = 0.44, P = 0.004$)、刺叶高山栎所有种群(Mantel's $r = 0.26, P < 0.001$)和 *Q. spi-West* 谱系(Mantel's $r = 0.31, P = 0.03$)中均发现了显著的地理隔离模式, 在 *Q. spi-West* 谱系、*Q. aqu-West* 和 *Q. aqu-East* 中并没有发现显著的地理隔离模式。在控制地理因素的影响时, 我们在 *Q. aqu-East* 谱系(Mantel's $r = 0.35, P = 0.02$)、刺叶高山栎的所有种群(Mantel's $r = 0.17, P = 0.011$)、*Q. spi-West* 谱系(Mantel's $r = 0.17, P = 0.007$)和 *Q. spi-West* 谱系(Mantel's $r = 0.079, P = 0.05$)中发现了显著的环境隔离的遗传模式, 但是在川滇高山栎的所有种群(Mantel's $r = -0.21, P = 0.841$)以及 *Q. aqu-West* 谱系(Mantel's $r = -0.27, P = 0.852$)中并没有发现发现遗传距离和环境距离存在显著的相关性。考虑到偏 Mantel 检验结果揭示 *Q. spi-East* 谱系中地理隔离的模式要强于环境隔离(Mantel's $r = 0.31$ vs. 0.079), 因此我们只考虑 *Q. spi-East* 谱系中存在显著的地理隔离模式。

我们用 Mantel 和偏 Mantel 检验检测了遗传距离和四个气候因子之间的相关性(表 3.5)。Mantel 检验结果揭示了在川滇高山栎所有种群(Mantel's $r = 0.54, P = 0.001$)和 *Q. aqu-East* 谱系(Mantel's $r = 0.15, P = 0.009$)中遗传距离与年温度范围(bio07)存在显著的相关性; 在刺叶高山栎所有种群中 3 月份降水量(prec03)与遗传距离显著相关, 但是并没有揭示气候因子与 *Q. spi-West* 谱系的遗传距离有显著的相关性。偏 Mantel 检验也揭示了类似的结果: 在川滇高山栎所有种群(Mantel's $r = 0.44, P < 0.001$)、刺叶高山栎所有种群(Mantel's $r = 0.24, P = 0.006$)和 *Q. aqu-East* 谱系(Mantel's $r = 0.15, P = 0.03$)中 bio07 是影响遗传变异最重要的气候因子, 在 *Q. spi-West*(Mantel's $r = 0.29, P = 0.02$)和 *Q. spi-East* 谱系(Mantel's $r = 0.03, P = 0.04$)中 prec03 是影响遗传变异最重要的气候因子。

表 3.5 川滇高山栎和刺叶高山栎成对遗传距离($F_{ST} / (1-F_{ST})$)和四种环境距离之间的 Mantel 和偏 Mantel 检验结果

Table 3.5 Mantel test and partial Mantel test between pairwise genetic distance ($F_{ST} / (1-F_{ST})$) and four environmental distances in *Q. aquifolioides* and *Q. spinosa*

	<i>Q. aquifolioides</i>		<i>Q. aqu-West</i> lineage		<i>Q. aqu-East</i> lineage	
	Mantel's r	P	Mantel's r	P	Mantel's r	P
<i>Mantel test</i>						
bio07	0.54	0.001	0.77	0.934	0.15	0.009
bio08	0.05	0.375	0.01	0.682	0.02	0.486
prec03	-0.31	0.964	0.56	0.04	-0.38	0.971
prec09	-0.19	0.821	-0.1	0.773	-0.35	0.922
<i>partial Mantel test</i>						
bio07	0.44	<0.001	0.79	0.02	0.15	0.003

研究结果

	<i>Q. spinosa</i>		<i>Q. spi-West lineage</i>		<i>Q. spi-East lineage</i>	
	Mantel's <i>r</i>	<i>P</i>	Mantel's <i>r</i>	<i>P</i>	Mantel's <i>r</i>	<i>P</i>
bio08	0.05	0.411	0.01	0.565	0.01	0.497
perc03	-0.31	0.963	-0.64	0.993	-0.39	0.971
perc09	-0.19	0.809	-0.14	0.811	-0.37	0.933
<i>Mantel test</i>						
bio07	0.31	<0.001	0.2	0.429	0.34	0.068
bio08	0.18	0.174	0.13	0.494	0.165	0.228
perc03	0.13	0.187	0.09	0.303	0.06	0.448
perc09	0.07	0.378	0.19	0.258	0.23	0.263
<i>partial Mantel test</i>						
bio07	0.24	0.006	0.02	0.464	0.34	0.058
bio08	0.15	0.241	0.14	0.456	0.16	0.241
perc03	-0.01	0.55	0.29	0.02	0.03	0.04
perc09	-0.03	0.586	0.18	0.263	0.21	0.193

注 a: 粗体数字表示数据显著。

多元回归(MRM)结果(表 3.6)显示: 在川滇高山栎所有种群和 *Q. aqu-West* 谱系中最优模型均显著($P=0.018$ 和 $P=0.03$)。在川滇高山栎所有种群(Spearman's $\rho=0.52$, $P=0.009$)和 *Q. aqu-West* 谱系(Spearman's $\rho=0.58$, $P=0.038$)中遗传距离和地理距离显著相关, 同时 *Q. aqu-West* 谱系中遗传距离和 bio07(Spearman's $\rho=0.18$, $P=0.006$)以及 perc03(Spearman's $\rho=0.27$, $P=0.043$)显著相关。在刺叶高山栎所有种群($P=0.018$)和 *Q. spi-West* 谱系($P=0.006$)中最优模型均显著。遗传距离和地理距离在刺叶高山栎所有种群中存在显著相关性(Spearman's $\rho=0.40$, $P<0.001$), 在 *Q. spi-West* 谱系中没有发现显著相关性(Spearman's $\rho=-0.01$, $P=0.953$), 但在 *Q. spi-West* 谱系中遗传距离与 perc03 气候因子存在显著相关性(Spearman's $\rho=0.21$, $P=0.048$)。

表 3.6 MRM 估计的川滇高山栎和刺叶高山栎遗传距离与地理和环境距离的相关性结果
Table 3.6 Correlations of genetic distance with geographical and environmental distances in *Q. aquifolioides* and *Q. spinosa* estimated by MRM

	<i>Q. aquifolioides</i>		<i>Q. aqu-West lineage</i>		<i>Q. aqu-East lineage</i>	
	ρ	<i>P</i>	ρ	<i>P</i>	ρ	<i>P</i>
geographical distance	0.52	0.009	0.58	0.038	0.06	0.766
bio07	0.18	0.006	-0.4	0.113	0.41	0.024
bio08	0.04	0.7	-0.27	0.034	0.21	0.084
perc03	0.27	0.043	-0.62	0.068	0.49	0.058
perc09	-0.04	0.719	0.12	0.675	-0.14	0.521
ρ^2 and <i>P</i> of best model	0.14	0.018	0.62	0.03	0.31	0.111

	<i>Q. spinosa</i>		<i>Q. spi-West lineage</i>		<i>Q. spi-East lineage</i>	
	ρ	<i>P</i>	ρ	<i>P</i>	ρ	<i>P</i>
geographical distance	0.4	<0.001	-0.01	0.953	0.04	0.038
bio07	0.05	0.734	0.34	0.315	0.37	0.137
bio08	0.07	0.632	0.18	0.559	0.01	0.095
prec03	-0.03	0.793	0.21	0.048	0.07	0.006
prec09	0.18	0.054	0.19	0.441	0.17	0.618
ρ^2 and <i>P</i> of best model	0.25	0.018	0.13	0.006	0.19	0.538

注 a: 粗体数字表示数据显著。

综上所述，Mantel 和偏 Mantel 检验揭示了川滇高山栎和刺叶高山栎所有种群以及 *Q. spi-East* 谱系中发现了显著的地理隔离模式，而在 *Q. aqu-East* 谱系和 *Q. spi-West* 谱系中发现了显著的环境隔离模式。同时偏 Mantel 检验和 MRM 回归揭示了 bio07 是影响 *Q. aqu-East* 谱系遗传变异最主要的环境因子，而 prec03 是影响 *Q. spi-West* 谱系遗传变异最重要的环境因子。

3.4.2 冗余分析

冗余分析(RDA)结果（图 3.8，表 3.7）揭示了川滇高山栎所有种群、*Q. aqu-West* 和 *Q. aqu-East* 谱系的遗传变异都与气候因素显著相关，对川滇高山栎所有种群和 *Q. aqu-East* 谱系遗传变异最重要的环境因子是最潮湿季节的平均温度(bio08)，对 *Q. aqu-West* 谱系影响最大的气候因子 bio07。刺叶高山栎所有种群、*Q. spi-West* 和 *Q. spi-East* 谱系的遗传变异同样与气候因素显著相关，对刺叶高山栎所有种群和 *Q. spi-East* 谱系遗传变异影响最大的气候因子是 prec03，对 *Q. spi-West* 谱系影响最大的气候因子是九月份降水量(prec09)。偏 RDA 分析表明当控制地理因素的影响时川滇高山栎所有种群、*Q. aqu-West* 和 *Q. aqu-East* 谱系的遗传变异与气候因素显著相关，当控制气候因素的影响时并没有在 *Q. aqu-West* 和 *Q. aqu-East* 谱系发现遗传变异与地理因素显著相关。我们也在川滇高山栎所有种群中发现了地理因素的显著影响，但是地理因素(PVE=18.22, $P = 0.042$)的影响要小于气候变量(PVE=10.04, $P = 0.017$)。在刺叶高山栎所有种群、*Q. spi-West* 和 *Q. spi-East* 谱系的遗传变异同样与气候显著相关，但是刺叶高山栎所有中地理因素(PVE=17.9, $P = 0.001$)的影响要大于气候因素(PVE = 10.27, $P = 0.001$)。

结合 RDA 和 p RDA 的结果我们发现气候因素是影响川滇高山栎和刺叶高山栎遗传变异最重要的影响因素。同时我们发现对川滇高山栎和刺叶高山栎遗传变异产生影响的气候因子存在差异，温度(bio07 和 bio08)对川滇高山栎的遗传变异有显著影响，

而降水(prec03 和 prec09)对刺叶高山栎的遗传变异有显著影响。

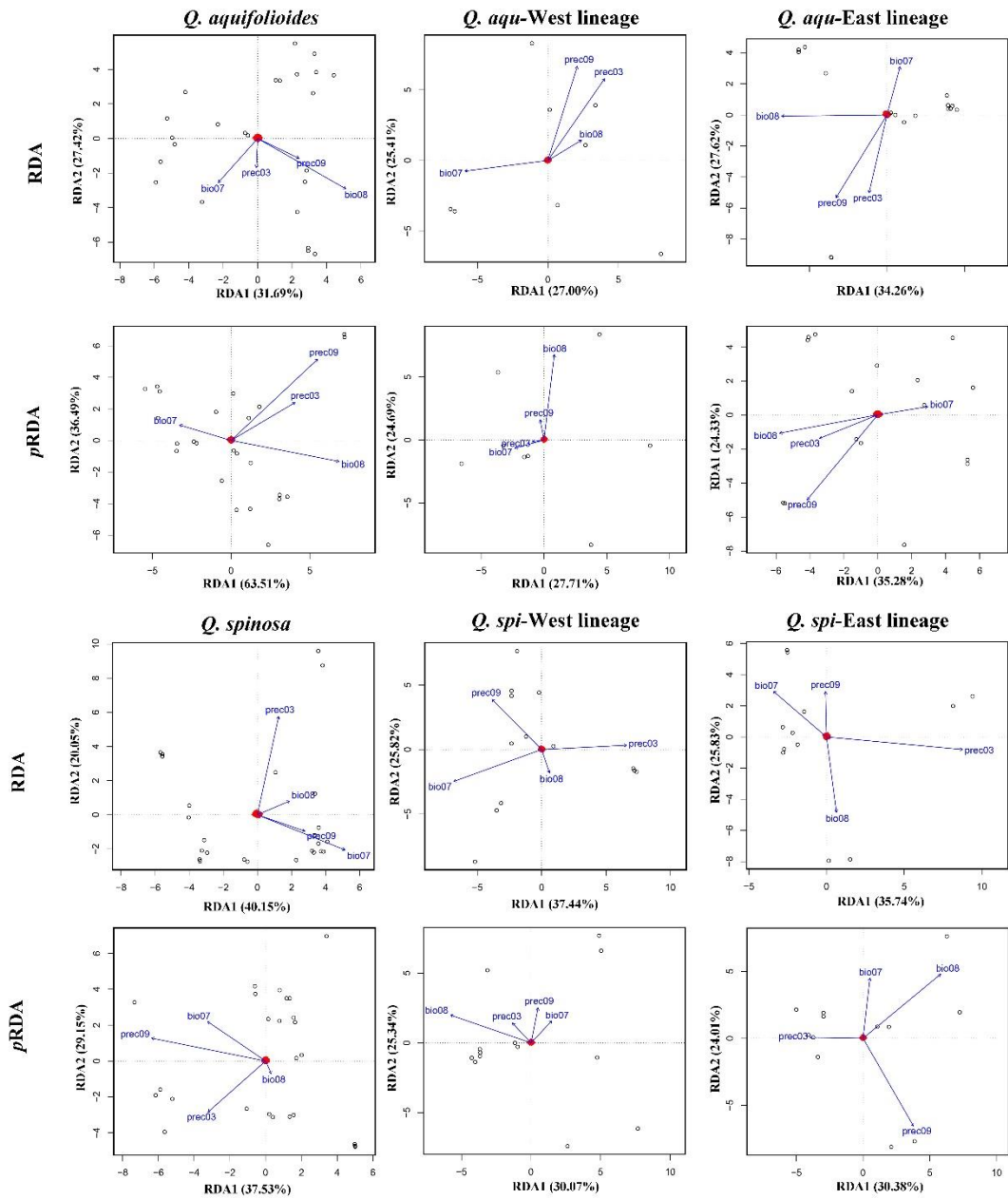


图 3.8 川滇高山栎和刺叶高山栎的冗余分析(RDA)和偏冗余分析(pRDA) 结果图

Figure 3.8 Triplots of redundancy analysis (RDA) and partial redundancy analysis (pRDA) results for

Q. aquifolioides and *Q. spinosa*

注 a: 环境变量被描述为蓝色向量 (箭头), 其中向量的长度表示该环境变量在解释方差中的贡献大小, 箭头之间的角度表示变量之间相关性的表示。

表 3.7 川滇高山栎和刺叶高山栎基于 RDA 和 pRDA 的与气候和地理变量相关的遗传变异结果

Table 3.7 Summary of the genetic variations associated with climate and geographic variables based on RDA and pRDA in *Q. aquifolioides* and *Q. spinosa*

	RDA			pRDA				RDA			pRDA		
	PVE	Eigenvalue	P	PVE	Eigenvalue	P		PVE	Eigenvalue	P	PVE	Eigenvalue	P
<i>Q. aquifolioides</i>							<i>Q. spinosa</i>						
climate	19.82	1.17	0.001	18.22	1.1	0.042	climate	17.9	1.21	0.001	17.9	1.21	0.001
geography				10.04	1.22	0.017	geography				10.27	1.39	0.001
bio07	21.96	1.03	0.325	23.49	1.04	0.367	bio07	22.86	1.1	0.192	22.86	1.1	0.19
bio08	28.51	1.34	0.001	26.85	1.19	0.038	bio08	24.37	1.18	0.071	24.37	1.18	0.07
prec03	25.9	1.22	0.008	25.1	1.11	0.164	prec03	26.54	1.28	0.006	26.54	1.28	0.01
prec09	23.64	1.11	0.082	24.56	1.08	0.221	prec09	26.24	1.27	0.006	26.24	1.27	0.01
<i>Q. aqu-West lineage</i>							<i>Q. spi-West lineage</i>						
climate	57.74	1.03	0.048	56.34	1.06	0.035	climate	37.73	1.21	0.001	33.89	1.13	0.04
geography				29.01	1.1	0.49	geography				17.11	1.14	0.28
bio07	25.93	1.06	0.047	24.69	1.05	0.043	bio07	26.86	1.3	0.005	25.05	1.13	0.01
bio08	23.69	0.97	0.671	24.24	1.03	0.465	bio08	23.44	1.14	0.112	25.05	1.13	0.31
prec03	25.1	1.03	0.276	24.76	1.05	0.434	prec03	22.29	1.08	0.168	22.21	1	0.51
prec09	25.29	1.04	0.24	26.31	1.12	0.045	prec09	27.41	1.33	0.001	27.7	1.25	0.04
<i>Q. aqu-East lineage</i>							<i>Q. spi-East lineage</i>						
climate	32.45	1.32	0.001	30.87	1.31	0.003	climate	42.09	1.27	0.016	34.44	1.04	0.05
geography				14.44	1.22	0.118	geography				16.51	1	0.51
bio07	21.43	1.13	0.103	24.45	1.28	0.056	bio07	23.59	1.2	0.125	26.7	1.11	0.35
bio08	33.71	1.78	0.001	29.59	1.55	0.001	bio08	22.19	1.13	0.111	24.52	1.02	0.51
prec03	22.61	1.19	0.066	22.18	1.16	0.168	prec03	32.43	1.65	0.006	27.19	1.13	0.04
prec09	22.24	1.18	0.079	23.78	1.24	0.084	prec09	21.79	1.11	0.224	21.59	0.9	0.71

注 a: PVE = percentage of explained variance; Eigenvalue-特征值;

注 b: 粗体数字表示数据显著。

3.4.3 梯度森林

梯度森林(GF)结果显示(图 3.9, 表 3.8): 对川滇高山栎所有种群和 *Q. aqu-West* 谱系来说, 地理因素(经度)是影响其遗传变异最主要的因素, 其次是气候因子。对 *Q. aqu-East* 谱系来说气候因子比地理因素在塑造其遗传变异方面的影响更大。Bio07 是对川滇高山栎所有种群和 *Q. aqu-East* 谱系影响最大的气候因子, 而 prec03 是对 *Q. aqu-West* 谱系影响最大的气候因子。地理因素(经度)同样是对刺叶高山栎所有种群和 *Q. spi-West* 谱系遗传变异影响最大的因素, bio07 是刺叶高山栎所有种群遗传变异最重要的气候因子, 而 prec03 是对 *Q. spi-West* 和 *Q. spi-East* 谱系影响最大的气候因子。

根据川滇高山栎和刺叶高山栎以及各自谱系内遗传变异沿环境梯度累计重要性的样条图我们发现(图 3.10): 在川滇高山栎所有种群以及 *Q. aqu-East* 谱系中 bio07 对遗传变异的影响较为平缓, 在 *Q. aqu-West* 谱系中 prec03 对遗传变异的影响在 10-15mm 范围内最为剧烈。另外在刺叶高山栎所有种群中 bio07 的影响样条曲线同样平滑, 而对于 *Q. spi-West* 谱系 prec03 对遗传变异的影响在 60-70mm 之间有明显的升高, 同样对于 *Q. spi-East* 谱系, prec03 对遗传变异的影响在 100-150mm 之间变化剧烈, 进一步说明了 prec03 在塑造 *Q. spi-West* 和 *Q. spi-East* 谱系遗传变异上的重要性。

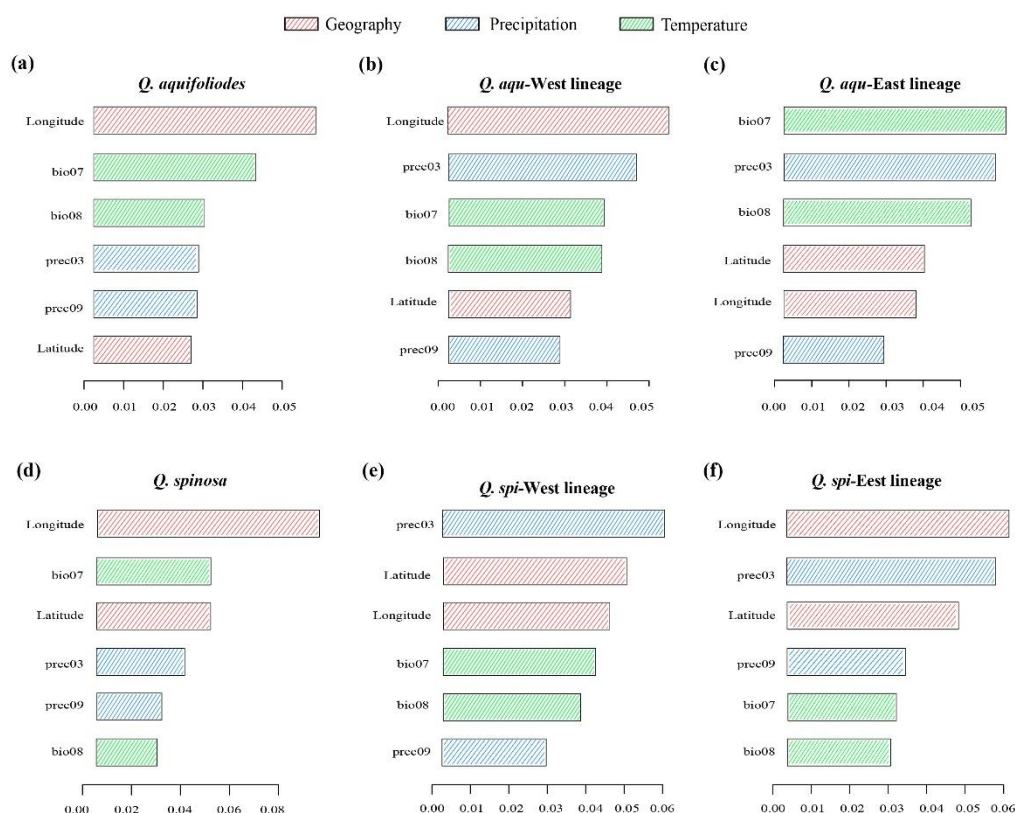


图 3.9 基于梯度森林(GF)分析的川滇高山栎和刺叶高山栎环境变量重要性排序图

Figure 3.9 Ranked importance of environmental variables based on analysis of a Gradient Forest (GF) for *Q. aquifolioides* and *Q. spinosa*

表 3.8 川滇高山栎和刺叶高山栎遗传变异和环境梯度之间的 GF 建模的每个环境变量的 R² 加权重要性(%)结果

Table 3.8 R²-weighted importance (%) of each environmental variables of GF modeling between genetic variants and environmental gradients in *Q. aquifolioides* and *Q. spinosa*

	<i>Q. aquifolioides</i>	<i>Q. aqu-West</i>	<i>Q. aqu-East</i>	<i>Q. spinosa</i>	<i>Q. spi-West</i>	<i>Q. spi-East</i>
Longitude	5.61	5.23	3.56	9.27	4.83	6.04
latiyude	2.16	2.89	3.79	4.75	5.31	4.76
bio07	4.09	3.67	5.96	4.78	4.41	3.09
bio08	2.8	3.64	5.04	2.53	3.98	3.01
prec03	2.65	4.44	5.67	3.69	6.44	5.54
prec09	2.61	2.64	2.71	2.71	3.02	3.19

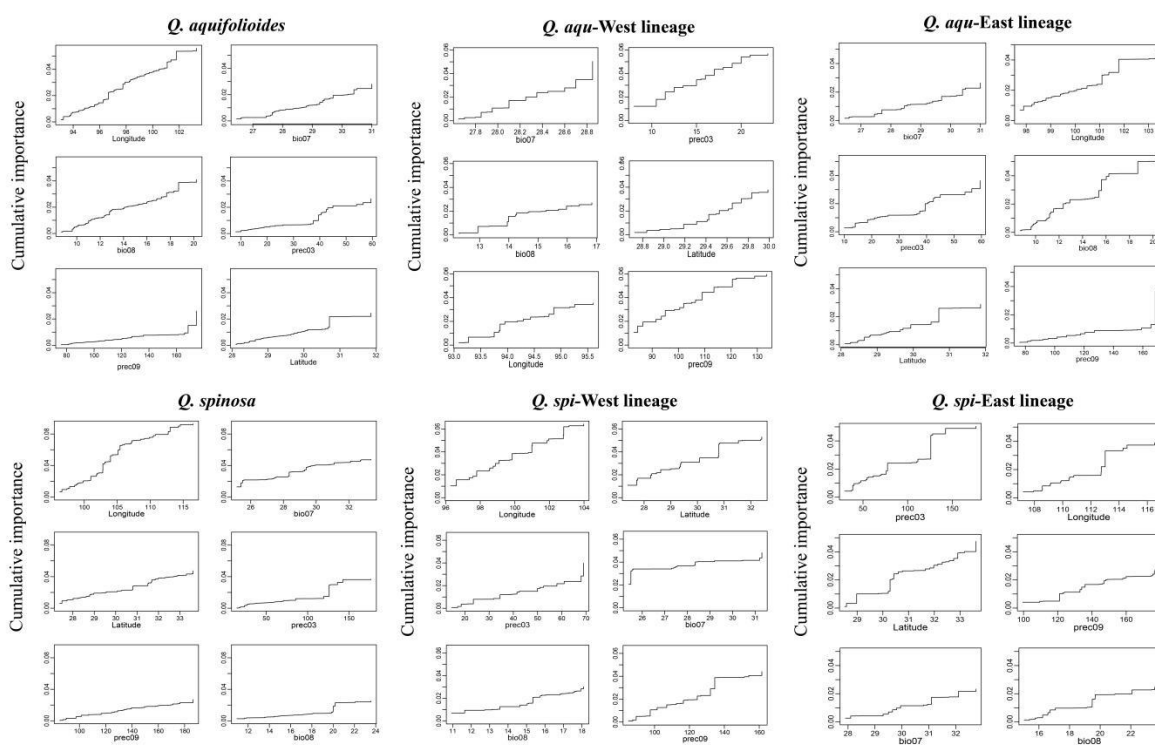


图 3.10 川滇高山栎和刺叶高山栎沿地理和环境梯度的遗传组成变化的 I-样条曲线图

Figure 3.10 I-splines showing changes in genetic distance along geographic and environmental distance gradients by Gradient Forest (GF) in *Q. aquifolioides* and *Q. spinosa*

3.4.4 广义相异建模

广义相异建模(GDM)结果表明(图 3.11, 表 3.9): 地理因素是川滇高山栎所有种

群、*Q. aqu-West* 谱系、刺叶高山栎所有种群和 *Q. spi-East* 谱系中最重要环境变量，而气候因子在 *Q. aqu-East* 和 *Q. spi-West* 谱系的重要性高于地理因素。同时对川滇高山栎而言，bio07 是影响所有种群以及 *Q. aqu-East* 谱系种群遗传变异最主要的气候变量，而 prec03 在影响 *Q. aqu-West* 谱系种群遗传变异方面发挥重要作用。对刺叶高山栎而言，bio07 是所有种群遗传变异最重要的环境因子，而 prec03 对影响刺叶高山栎东西谱系种群的遗传变异中发挥最大作用。

广义相异建模的 I-样条曲线结果相较于梯度森林更为平滑，但是环境因子对遗传剧烈的区域也更易判断。I-样条曲线结果没有展示重要性为 0 的环境变量。在川滇高山栎所有种群中温度(bio07 和 bio08)在一定的范围内对遗传变异的影响有显著的上升，但是降水(prec03 和 prec09)的样条曲线没有波动，说明降水因素对川滇高山栎所有种群的遗传变异的影响不大。同样在 *Q. aqu-West* 谱系中，prec03 的样条曲线在 15-20mm 范围内变化明显，说明在此范围内 prec03 对其种群遗传变异的影响最大。在 *Q. spi-West* 和 *Q. spi-East* 谱系中，prec03 的样条曲线在一定范围内有显著上升，说明 prec03 在塑造其种群遗传变异中的重要性。

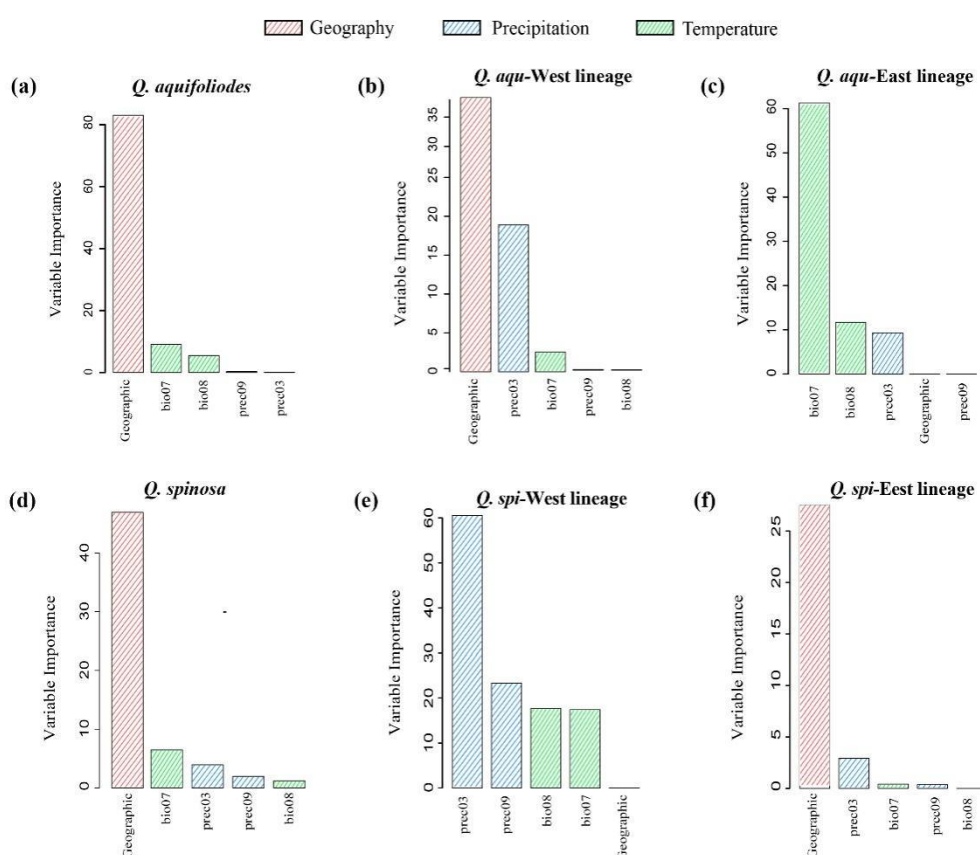


图 3.11 基于广义相异模型(GDM)分析的川滇高山栎和刺叶高山栎环境变量重要性排序图
 Figure 3.11 Ranked importance of environmental variables based on analysis of a generalized dissimilarity model (GDM) for *Q. aquifolioides* and *Q. spinosa*

表 3.9 基于广义相异建模的川滇高山栎和刺叶高山栎地理和环境变量的重要性(%)结果
Table 3.9 Variable importance (%) of each environmental variables of Generalized Dissimilarity Model correlated with geographic and environmental variables in *Q. aquifolioides* and *Q. spinosa*

	<i>Q. aquifolioides</i>	<i>Q. aqu-West</i>	<i>Q. aqu-East</i>	<i>Q. spinosa</i>	<i>Q. spi-West</i>	<i>Q. spi-East</i>
Geographic	83.00	0.00	35.26	46.93	0.02	0.00
bio07	9.00	61.3	2.55	6.47	17.38	2.94
bio08	5.46	11.68	0.06	1.21	17.68	0.37
prec03	0.03	9.32	18.89	3.94	60.55	27.5
prec09	0.25	0.00	0.11	1.99	23.29	0.42

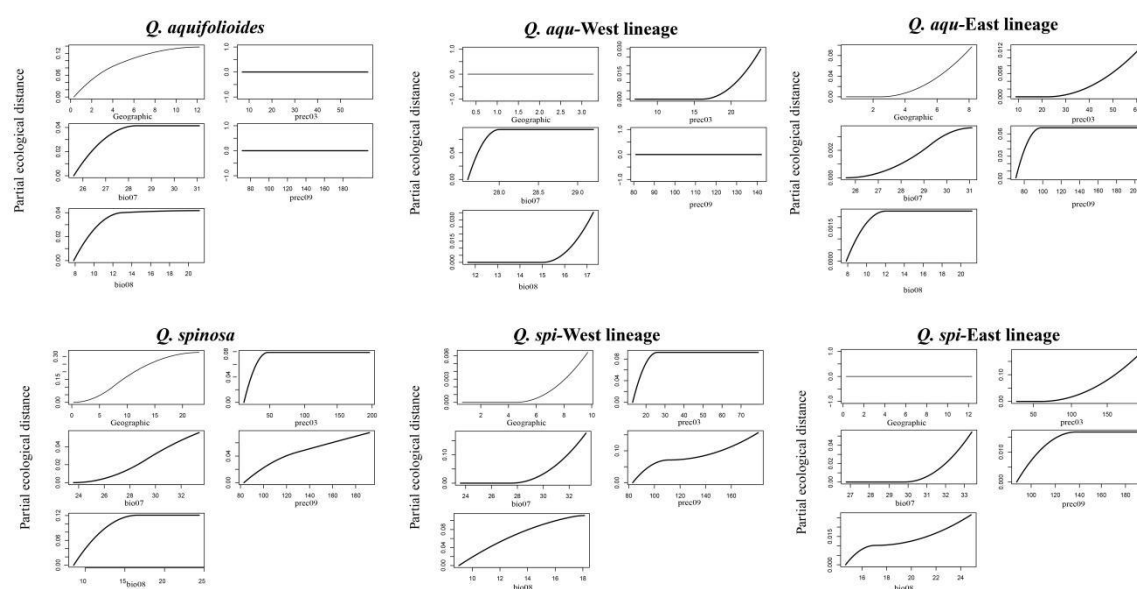


图 3.12 基于广义相异模型在川滇高山栎和刺叶高山栎中沿环境梯度的遗传变异的累积重要性图
Figure 3.12 Cumulative importance of genetic variation along environmental gradients by generalized dissimilarity model in *Q. aquifolioides* and *Q. spinosa*

3.5 异常位点检测结果

我们使用基于遗传分化异常位点挑选方法 BayeScan 方法在川滇高山栎所有种群中发现了 938 个异常 SNP 位点，在 *Q. aqu-East* 谱系中发现了 120 个与遗传分化相关的 SNP，但是我们并没有在 *Q. aqu-West* 谱系中发现与遗传分化相关的 SNP 位点。我们对川滇高山栎所有种群和 *Q. aqu-East* 谱系中挑选的异常 SNP 进行整合分析，发现有 26 个 SNP 在所有种群和 *Q. aqu-East* 谱系均被视为异常值，在 1 号、3 号和 12 号染色体上最多，分别有 4 个、7 个和 11 个。我们在刺叶高山栎所有种群中发现了 9624 个与遗传分化相关的异常 SNP，在 *Q. spi-West* 谱系中发现了 650 个异常 SNP。同时有 14 个异常 SNP 在所有种群和 *Q. spi-West* 谱系中均检测到，分布于 1 号和 9 号染

染色体上的异常 SNP 最多，分别有 3 个和 4 个(图 3.13)。

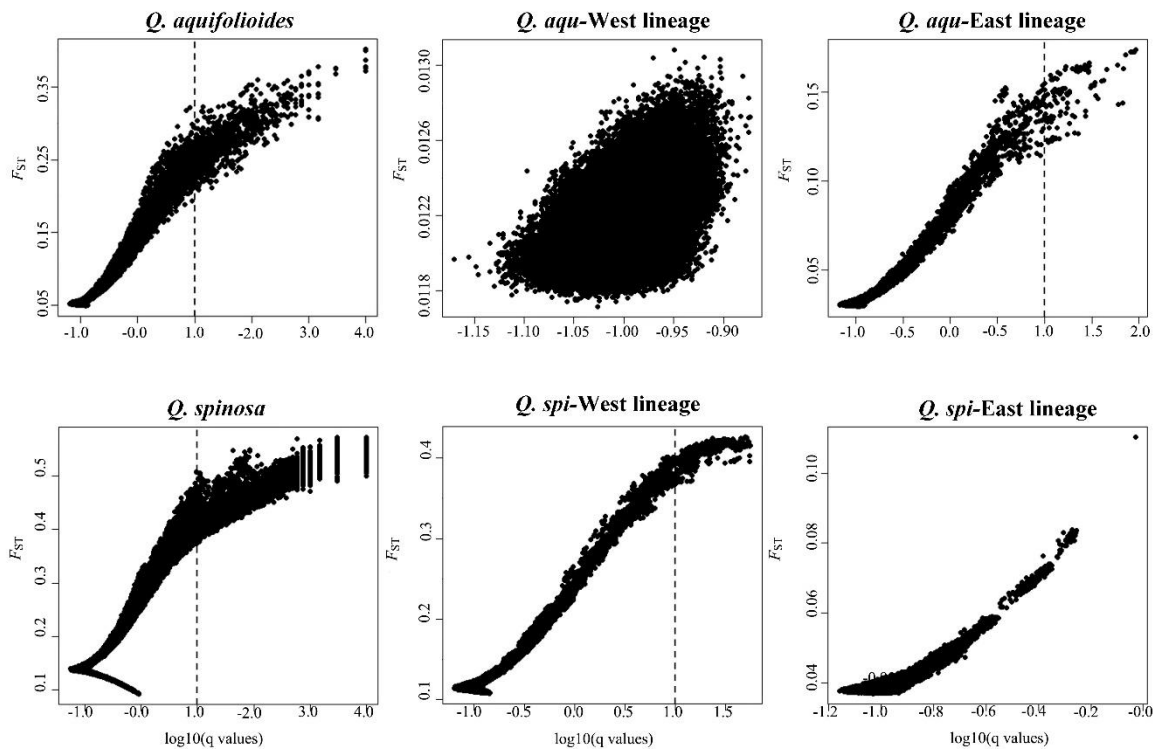


图 3.13 基于 BayeScan 方法挑选与遗传分化相关异常 SNP 位点结果图

Figure 3.13 Results of outlier SNP identification based on BayeScan

我们采用基因型-环境关联(GEAs)方法挑选与气候因子相关的异常 SNP 位点，我们采用上文研究的 4 种气候因子作为本分析的环境因子。我们在 *Q. aqu-West* 谱系发现了 2341、2670、2607、2946 个分别与 bio07、bio08、prec03 和 prec09 相关的异常 SNP 位点，同时这 4 个气候因子共同检测到 1 个异常位点，位于 1 号染色体；在 *Q. aqu-East* 谱系发现了 4931、7620、4464、6758 个分别与 bio07、bio08、prec03 和 prec09 相关的异常 SNP 位点，这 4 个气候因子共同检测到 4 个异常位点，位于 1、6、7、10 号染色体。在 *Q. spi-West* 谱系发现了 9620、2970、6205、6134 个分别与 bio07、bio08、prec03 和 prec09 相关的异常 SNP 位点，这 4 个气候因子共同检测到 8 个异常位点，位于 1、3、7、10 号染色体；在 *Q. spi-East* 谱系发现了 3838、4176、4718、6070 个分别与 bio07、bio08、prec03 和 prec09 相关的异常 SNP 位点，这 4 个气候因子共同检测到 1 个异常位点，位于 1 号染色体（图 3.14）。

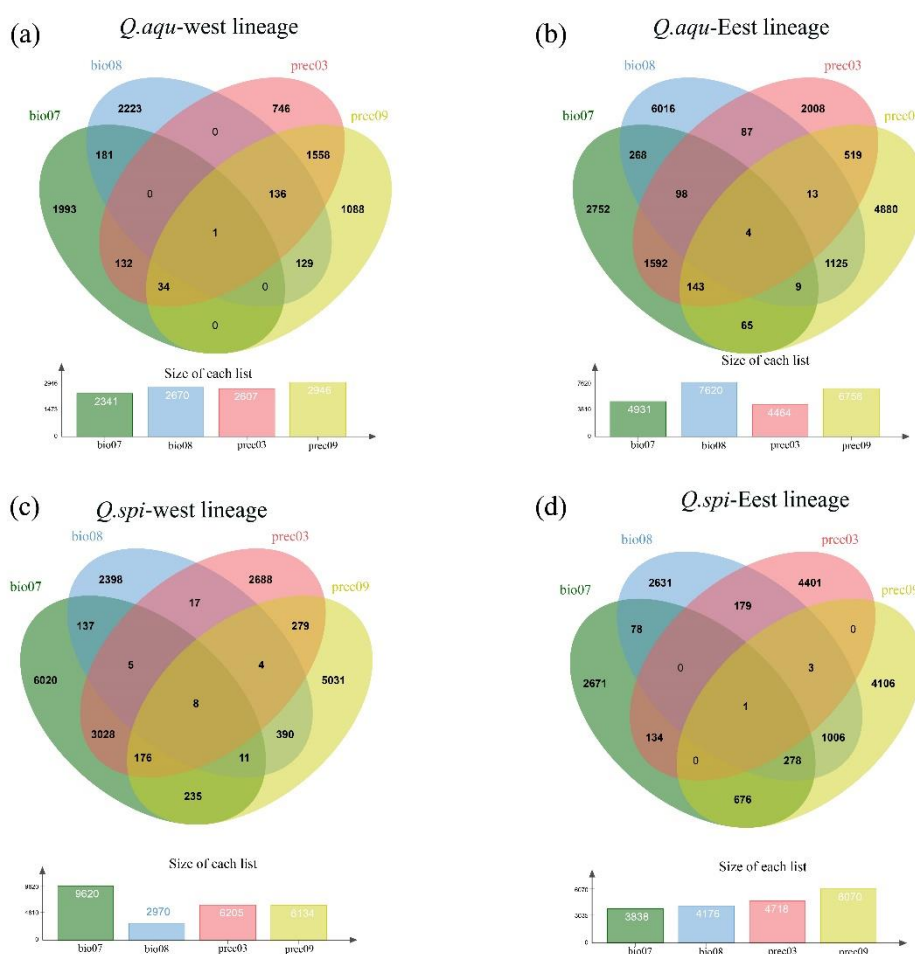


图 3.14 基于 LFMM 的与 4 个气候因子相关的异常位点数量分布的韦恩图

Figure 3.14 The Venn diagram showing the outliers of four climatic factors detected by LFMM

3.6 气候变化下的遗传偏移

3.6.1 梯度森林模型

对川滇高山栎而言(图 3.15 a), 在 2050 年和 2070 年 RCP26 碳排放情境下川滇高山栎的遗传偏移结果类似, 最大遗传偏移的地区(图中红色区域)出现在其预测分布范围的西部(西藏地区)和北部(四川北部), 代表着该地区的种群有着最大的不适应性风险。在 RCP85 碳排放情境下, 2050 年和 2070 年预测的红色区域(遗传偏移大的区域)进一步扩大, 尤其是西藏地区的种群面对气候变化时所受到威胁区域变大。相较于极端低的温室气体排放情景(RCP26), 极端高的温室气体排放情景(RCP85)对川滇高山栎未来命运的影响更大。整体来看, 云南省西北部地区的种群相较于其他地区有着最低的遗传偏移量, 说明该区域的种群有着最大的适应潜力。

对刺叶高山栎而言(图 3.15 b), 在 2050 年和 2070 年 RCP26 碳排放情境下川滇高山栎的遗传偏移结果类似, 没有明显高遗传偏移的地区。而在两年份 RCP85 碳排放

情景下红色区域面积明显扩大，尤其是在刺叶高山栎整个分布区域的北部地区，说明在此情境下该地区的种群有着最大的灭绝风险。在 2050 年和 2070 年两种碳排放情景下并没有在西藏地区发现高遗传偏移的地区，同样随着碳排放情景极端性的提升，也没有在整个预测分布范围的东部地区发现明显高的遗传偏移。

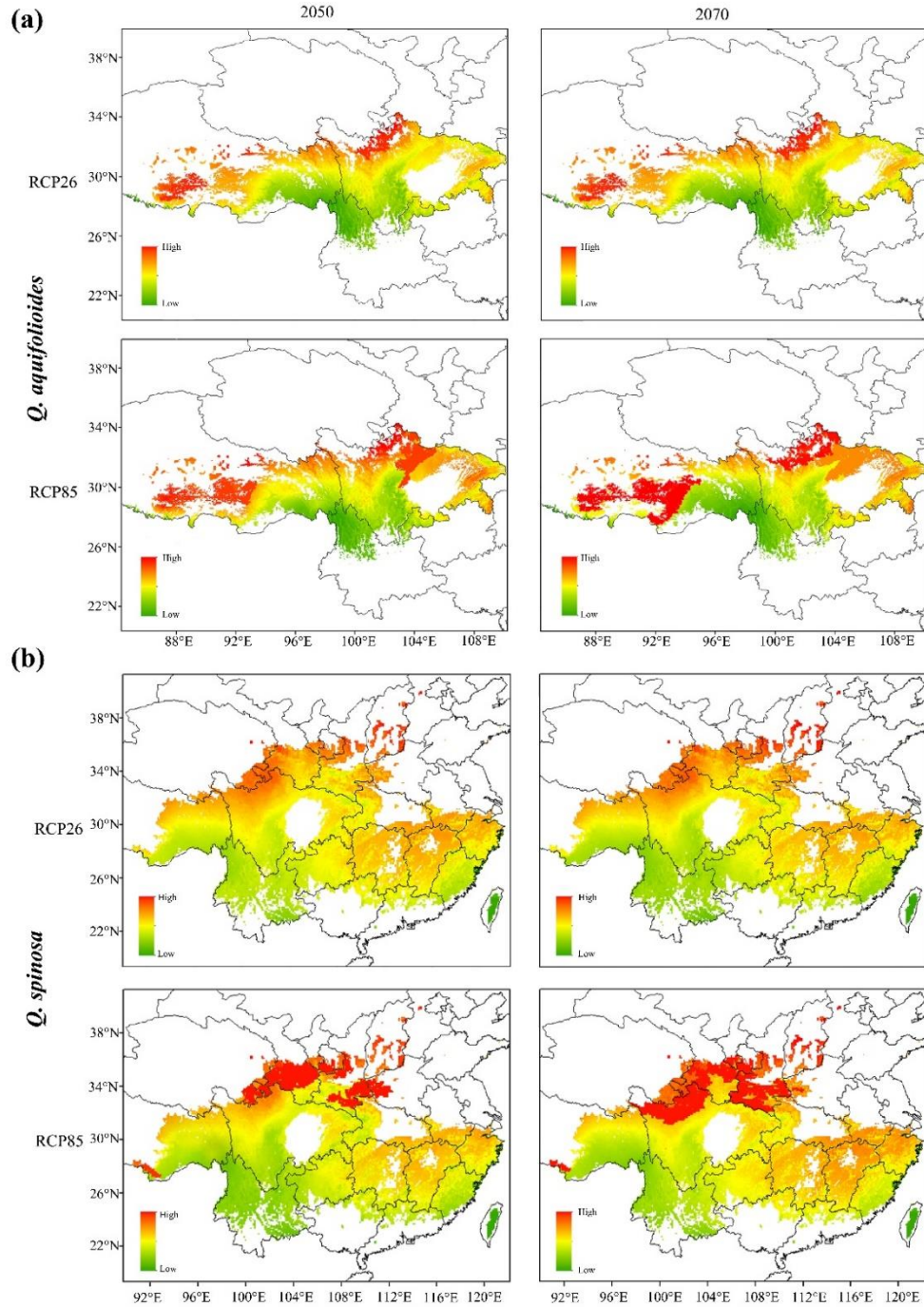


图 3.15 在 RCP26 和 RCP85 情景下，2050 年和 2070 年川滇高山栎和刺叶高山栎分布预测的遗传偏移结果

Figure 3.16 Predicted genetic offset across the distribution of *Q. aquifolioides* and *Q. spinosa* in the years 2050 and 2070 under scenario RCP26 and RCP85

3.6.1 非适应性风险分析

RONA 分析结果显示无论在川滇高山栎还是刺叶高山栎种群中, 最具代表性的环境因素是三月份降水量(prec03), 其次是九月份降水量(prec09)。对川滇高山栎而言, 在 2050 年的 RCP26 和 RCP85 两种碳排放情景下, 位于四川省东北部的平武县(PW)种群有着最大的 RONA 值, 说明该种群在 2050 年气候预测下有着最大的灭绝风险, 同时位于横断山脉其他地区的种群比位于西藏地区的种群的 RONA 值低, 说明横断山脉地区的种群, 尤其是位于云南省东北部有着更大的适应潜力(图 3.16 a, 附录 E)。2070 年在 RCP26 和 RCP85 两种碳排放情景下 RONA 分析的预测结果大致相同, 川滇高山栎所有种群的 RONA 值相差不大, 东西谱系的差异也不明显(图 3.16 a, 附录 F)。

对刺叶高山栎而言, 在 2050 年的 RCP26 和 RCP85 两种碳排放情景下预测结果大致相同, 位于西藏地区通麦县(TM)以及位于四川省和甘肃省交接处(JZG)的种群有着最大的 RONA 值(图 3.16 b, 附录 D); 这一推论在 2070 年 RCP26 和 RCP85 两种碳排放情景下的预测结果中得到了验证: TM 和 JZG 是刺叶高山栎所有种群中 RONA 值最高的, 即在未来气候变化下有着最大的灭绝风险(图 3.16 b, 附录 E)。同时结合 2050 年和 2070 年两种碳排放情景下的预测结果, 发现云南地区的种群(GZ、XGJQ)的 RONA 值相较于其他地区较低, 说明云南地区的种群在气候变化下有着更大的适应潜力(图 3.16 b)。

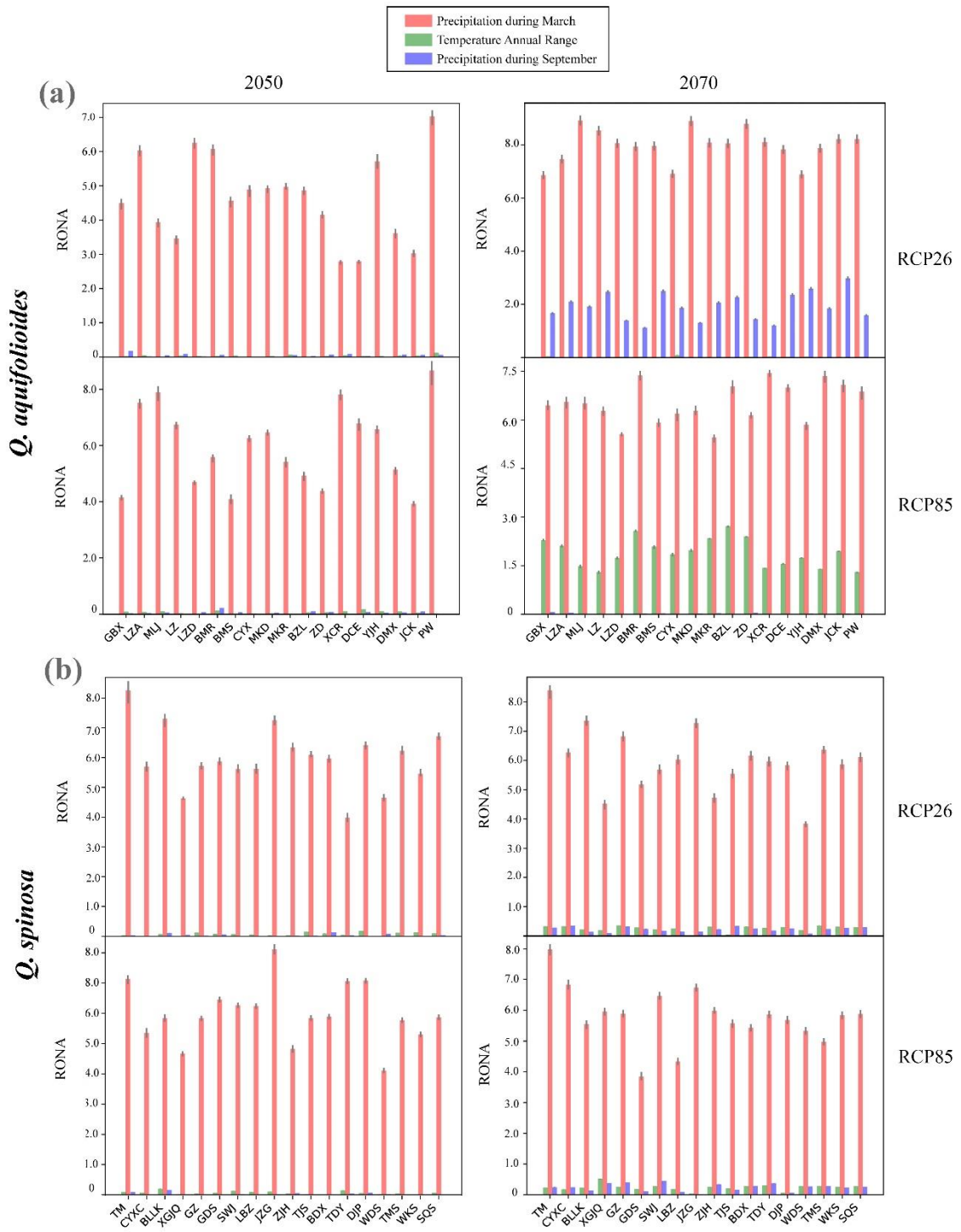


图 3.15 2050 年和 2070 年情景 RCP 26 和 RCP 85 预测模型下川滇高山栎和刺叶高山栎三个环境因素下的非适应风险(RONA)

Figure 3.15 Risk of nonadaptedness (RONA) for three environmental factors of *Q. aquifolioides* and *Q. spinosa* under scenario RCP26 and RCP85 prediction models in 2050 and 2070

4 讨论

4.1 遗传多样性和遗传结构

本研究分别使用 358 661 个和 346 281 个高质量的单核苷酸多态性位点(SNP)评估了川滇高山栎和刺叶高山栎的遗传多样性和遗传结构。研究表明川滇高山栎和刺叶高山栎在各自的分布范围内其种群结构模式可以分为东西两个谱系,分别为分布于青藏高原地区的 *Q. aqu-West*, 分布于横断山脉地区的 *Q. aqu-East* 和 *Q. spi-East* 谱系, 以及分布于秦岭山脉和中国东部温暖低地的 *Q. spi-West* 谱系(图 3.5)。两个常绿栎物种种内谱系具有相对不同的遗传成分,这可能是西部谱系和东部谱系之间基因流动的障碍所致,之前在中国亚热带其他物种中的研究也发现了这种强烈的遗传结构,比如银杏(*Ginkgo biloba*) (Gong *et al.*, 2008)、青冈(*Cyclobalanopsis glauca*) (Xu *et al.*, 2015)。基于 SSR 数据的 ABC 分析进一步揭示了两种栎树谱系间的祖先关系(图 3.6),川滇高山栎作为祖先种群,刺叶高山栎起源于川滇高山栎,随后两物种逐渐开始种内分化。这种强烈的种群结构可能与青藏高原和邻近地区的造山历史造成的长期地理隔离模式有关(Qiu *et al.*, 2011; Wen *et al.*, 2014), 青藏高原近期的造山运动以及气候波动催化了该地区物种的种内分化、多样化和适应。地形和气候的联合作用可能会导致许多具有不同小气候的小型破碎生境,这可能会影响自然选择的方向,并可能促进物种的种内高度分化(Sobel *et al.*, 2010)。青藏高原地区的复杂地质结构和气候带的垂直移动形成了高山植被的自然地理屏障和冰川避难所(Chen *et al.*, 2011), 复杂的山地地形是栎树种子或花粉传播的有效地理屏障(Li *et al.*, 2013; Liu *et al.*, 2013), 并为这两个栎树物种种内谱系的独立进化提供了潜在条件。青藏高原的隆升事件为该地区栎树的祖先种群不断向东扩散提供了机会,并逐渐引发和促进了栎树物种的形成和多样化(Zhou, 1992)。

遗传多样性的结果表明川滇高山栎的遗传多样性水平高于刺叶高山栎,这一结果可能与刺叶高山栎是从川滇高山栎分化出来的有关。另外在两物种种内谱系间西部地区的遗传多样性水平要明显高于东部谱系(表 3.2),可能与西部谱系的种群间的高遗传分化有关: AMOVA 分析中 *Q. aqu-West* ($F_{ST} = 0.05$)的遗传分化高于 *Q. aqu-East* 谱系($F_{ST} = 0.05$), 同样 *Q. spi-West* ($F_{ST} = 0.14$)的遗传分化高于 *Q. spi-East* 谱系($F_{ST} = 0.01$)。西部谱系的高遗传分化导致种群间的基因交流困难,因此会导致种群间的遗传多样性水平变低。另外横断山脉在第四纪冰期时可以作为物种避难所,因此会维持相当高的物种多样性(López-Pujol *et al.*, 2011; Favre *et al.*, 2015)。基于川滇高山栎和刺叶高山栎的 15 对 SSR 引物数据表明两物种之间存在不对称的历史基因流,历史基因流的方向是从川滇高山栎到刺叶高山栎,而谱系间的历史基因流都是从西谱系到东谱系

(Liu *et al.*, 2022)。谱系间的不对称基因交流使得两物种的西谱系得到了更多的私有等位基因, 使其遗传多样性逐渐变高。

4.2 基于基因型-环境关联的遗传变异的响应模式

基因型-环境关联(GEA)分析, 包括 Mantel 检验、冗余分析、梯度森林模型和广义相异性建模, 通常将环境变量和空间遗传结构整合到分析框架中以检测适应性变异(Feng and Du, 2022)。GEA 分析用于量化遗传变异和气候条件之间的相互作用模式, 对于了解对环境的进化反应的机制至关重要(Hansen *et al.*, 2012)。我们在川滇高山栎所有种群、刺叶高山栎所有种群以及 *Q. spi-East* 谱系中发现了显著的 IBD 模式(图 3.7, 表 3.4), 表明 *Q. spi-East* 谱系的遗传变异主要是由选择性中性进化过程驱动的, 而不是来自强大的环境驱动力。复杂的山脉地质结构可能形成种子或花粉传播的天然地理屏障, 为植物种内谱系的形成和独立进化提供潜在条件(Li *et al.*, 2013; Liu *et al.*, 2013)。在川滇高山栎的种内谱系间没有发现显著的 IBD 模式, 说明该谱系的种群的遗传距离与地理结构没有显著相关性。相反我们在 *Q. aqu-East* 和 *Q. spi-West* 谱系中发现了显著的 IBE 模式(表 3.4), 表明高原上极端的环境条件可能被视为重要的气候屏障, 而不是地理障碍。极端的气候条件可能影响着 *Q. aqu-East* 和 *Q. spi-West* 谱系种群的生长发育, 比如低温、缺氧、强紫外照射条件影响着高山栎植物的生长发育和物候。地理隔离可能导致种间和种内分化; 对当地环境因素的适应强化了这种差异性, 并逐渐形成了这种显著的 IBE 模式(Gao *et al.*, 2021)。偏 Mantel 检验和 MRM 进一步表明 bio07 是 *Q. aqu-East* 谱系遗传变异最重要的环境因子, prec03 是驱动 *Q. spi-West* 谱系遗传变异最重要的环境因子。温度可能是 *Q. aqu-East* 谱系遗传变异的主要驱动因素, 它可能通过影响代谢过程来影响微生境中植物的生长(Wahid *et al.*, 2007)。虽然栎树具有很强的耐旱、贫瘠的能力, 但降水可能对植物物候和生长有很大的影响, 进而影响 *Q. spi-West* 谱系适应气候变化的能力。考虑到该地区的异质微环境条件, 这些种群可能已经适应了其栖息地的微环境条件。

GF 和 GDM 结果表明对 *Q. aqu-East* 谱系来说气候因子比地理因素在塑造其遗传变异方面的影响更大, 但对 *Q. spi-West* 谱系遗传变异影响最大的因素是地理因素(经度), 非线性模型揭示 *Q. spi-West* 谱系中地理因素的重要性要高于环境因素, 可能与非线性模型中地理因素的量化模式有关。prec03 是对 *Q. spi-West* 谱系的遗传变异影响最大的环境因子, 与偏 Mantel 检验的结果是一致的。GF 模型的 I-样条曲线结果显示 prec03 对 *Q. spi-West* 谱系遗传变异的影响在 100-150mm 之间变化剧烈(图 3.10), 也进一步表明了降水量达到每一个阈值会对栎树生长发育或者适应产生影响。同样在 GDM 模型的 I-样条曲线可以看出(图 3.12), 在 *Q. spi-West* 谱系中 prec03 的样条曲线在一定范围内有显著上升, 也进一步说明了 prec03 在塑造遗传变异中的重要性。

结合线性模型和非线性模型分析的结果, 在 *Q. aqu-East* 和 *Q. spi-West* 谱系中适应性遗传变异部分由环境因素构成, 而在 *Q. aqu-East* 谱系中环境因素可能主要由温度梯度驱动, 而在 *Q. spi-West* 谱系环境因素可能主要由降水梯度驱动。

4.3 与生态适应相关的基因组特征

生态适应会在基因组区域留下选择的遗传印记(Joshi *et al.*, 2001; Mosca *et al.*, 2012)。关于气候变化的生态适应的研究有助于理解当地种群维持和适应快速气候变化的能力(Fournier-Level *et al.*, 2011), 同时了解当地种群基因库中的局部适应遗传变异的来源, 对于在未来环境中选择具有增产潜力的种群或个体非常重要(Redden, 2013)。挑选与选择相关的异常位点是一种有效的识别遗传印记的方法, 所有方法都有各自的局限性和优势(Hoban *et al.*, 2016; Sork, 2018)。在本研究中, 我们结合了基于 F_{ST} 和基因型-环境关联(GEA)的方法, 提高了异常位点检测的可靠性。我们基于 BayeScan 方法发现了一组可能与生态适应相关的位点, 发现有 26 个 SNP 在川滇高山栎所有种群和 *Q. aqu-East* 谱系均被视为异常值, 分布在 1 号、3 号和 12 号染色体上最多; 同时有 14 个异常 SNP 在所有种群和 *Q. spi-West* 谱系中均检测到, 分布于 1 号和 9 号染色体上的异常 SNP 最多。然而, 由于对环境异质性的忽视, 这种方法的假阳性结果比率很高(Eveno *et al.*, 2008; Hoban *et al.*, 2016)。LFMM 通过检测多态性和环境变量之间的高度相关性, 提供了一种研究局部适应特征的方法。该方法已被证明是一种可靠的方法, 即使 IBD 等种群进化历史因素倾向于使该分析产生偏差(Stucki *et al.*, 2017)。我们使用 LFMM 在全基因组范围内发现了大量与四种气候因子相关的异常位点, 通过整合我们在川滇高山栎和刺叶高山栎种内谱系间发现有 14 个异常 SNP 由四个气候因子共同检测到, 表明它们可能是在该谱系的不同环境的选择压力下进化而来。这也表明了环境因素作为重要的选择驱动力, 可能会驱动遗传变异, 并促进栎树适应其当地环境条件。Du 等(2020)年基于候选基因的方法研究川滇高山栎的生态适应性, 基因型-环境关联分析研究发现在 Tibet 和 HDM-WSP 谱系中确定了 12 个与气候变化相关的基因, 说明了该物种在一定程度上适应了干旱环境。近期发表的全基因组序列数据和测序技术的发展为揭示物种生态适应的分子机制奠定了基础, 可以在基因组水平寻找生态适应的分子印记。

4.4 物种优先保护区的确定

4.4.1 遗传多样性的空间分布

更好地了解遗传多样性的空间分布对于制定有效和高效的保护策略是必要的(Petit *et al.*, 1998)。应优先保护保留最高等位基因丰富度(AR)和局部共有等位基因(LCA)的种群, 因为在遗传多样性最高的种群中找到高适应能力的育种材料的可能性

更高, 这表明存在适应于特定环境的基因型(Frankel *et al.*, 1995)。基于空间遗传多样性结果发现, 川滇高山栎和刺叶高山栎的优先保护区域不同(图 3.4)。对于川滇高山栎而言, 优先保护区域应该是位于横断山脉, 因为该地区的种群包含最高的 AR 和 LCA。此外, 位于 HDM 的最东端平武, 其 AR 和 LCA 低于其他地区, 可能意味着有着最高的灭绝风险。因此在保护活动中也应考虑在分布范围边缘隔离的种群, 以防止该地区物种的灭绝。对于刺叶高山栎, 优先保护区域应该位于具有最高遗传多样性的秦岭山脉。第二个优先保护区应位于四川云南的交界地带, 该地区的高空间多样性可能与高物种丰富度有关。*Q. aqu-East* 和 *Q. spi-East* 谱系中的 AR 和 LCA 高于 *Q. aqu-West* 和 *Q. spi-West* 谱系, 表明这些种群接受了更多的迁移, 并在进化和多样化中发挥了重要作用(López-Pujol *et al.*, 2011)。*Q. aqu-West* 和 *Q. spi-West* 谱系中较低的遗传多样性可能反映了该地区的种群收缩或灭绝-重新定殖动态(Qiu *et al.*, 2011; Du *et al.*, 2017; Meng *et al.*, 2017)。

4.4.2 未来气候条件下的基因组脆弱性

树木通过迁移和适应新的环境条件来应对不断变化的环境, 与适应的新突变的产生相比, 持续的遗传变异有助于更快的适应性进化(Aitken *et al.*, 2008)。虽然未来环境变化对种群适合度的精确关系不容易确定, 为了更好模拟未来气候变化情境下基因型-环境关系之间的差异, Fitzpatrick 和 Keller (2015)基于梯度森林模型(GF)预测了香脂杨(*Populus balsamifera*)种群未来气候下遗传高度不匹配的地理区域, Rellstab 等(2016)开发了一个简化的模型-非适应性风向分析(RONA), 并预测了栎属(*Quercus spp.*)种群对未来当地气候的适应能力。目前这两种方法在树木未来命运预测中得到了广泛的应用, 如金合欢(*Acacia koa*, Gugger *et al.*, 2018)、侧柏(*Platycladus orientalis*, Jia *et al.*, 2020)和矮桦树(*Betula nana*, Borrell *et al.*, 2020)等, 表明了这两种模型在未来气候变化下预测物种命运的可靠性。本研究使用上述两种基于线性和非线性预测模型预测了未来气候情景下基因组脆弱性, 使用这种方法我们推断川滇高山栎和刺叶高山栎在未来气候变化下存在的潜在风险。我们的梯度森林模型(GF)分析表明, 川滇高山栎在横断山脉地区的西南部受到的影响最小, 而在整个分布区的北部和西藏地区西部边缘预测到了相对较高的遗传偏移(图 3.16 a)。同样我们也在刺叶高山栎分布区的北部边缘地区预测到了相对较高的遗传偏移(图 3.16 b)。这些高风险地区需要快速适应气候变化, 否则这些地区的种群可能会出现种群收缩现象。在预测高遗传偏移的领域中, 我们应该考虑使用一种混合原种, 将原生地与选定的非本地苗种混合以增加多样性和恢复力。栎树的生命周期长、世代间隔较长, 通过有利突变在种群中出现和传播新的适应性等位基因的速度可能太慢, 无法应对未来快速的气候变化。

我们使用 RONA 测试和比较了两个年份(2050 年和 2070 年)两种不同的碳排放场景(RCP26 和 RCP85)的未来气候相匹配的等位基因频率所需的预期变化, 假设通过等

位基因频率向当前经历预测气候的种群中出现的频率的线性变化进行适应, 并利用这些信息推断原位适应的潜力, 以及对基因流动或管理干预(如辅助迁移)的需要。我们的 RONA 结果表明三月份降水量(prec03)是最具代表性的环境因素(图 3.15), 说明随着未来气候变化, 降水量的变化对川滇高山栎和刺叶高山栎种群的适应性有着较大影响, 气候因子需要较大变化以确保适应预测的未来环境下变化。Martins 等(2018)基于 GF 预测表明了 *Q. rugosa* 种群未来预测中降水季节性是最具代表性的环境因素, Du 等(2020)年基于 RONA 分析发现了川滇高山栎两个谱系间最具代表性的环境因素是旱季降水量(prec01), 表明降水可能会对整个景观中这两个物种的等位基因频率的转换有很大影响。我们还发现位于四川省东北部的平武县(PW)川滇高山栎种群, 以及位于西藏地区通麦县(TM)以及位于四川省和甘肃省交接处(JZG)的刺叶高山栎种群有着最大的 RONA 值, 这些种群需要通过基因流引入“预适应”的等位基因, 以促进种群应对未来气候变化的能力。我们还发现云南地区的川滇高山栎和刺叶高山栎种群在气候变化下都有着很大的适应期潜力, 该地区的种群所经历的相对温暖和潮湿的气候条件可能有助于生态适应, 说明这些种群可能包含更多的潜在适应性多态性。

5 结论与展望

5.1 结论

本研究基于全基因组重测序技术在 49 个川滇高山栎和刺叶高山栎样品中一共产生了 1851.78G 的原始下机数据, 平均测序深度超过了 30X, 碱基正确识别率在 99% 以上(Q20)的碱基数目百分比平均值超过了 97%, 经过变异检测和 SNP 过滤后分别在川滇高山栎和刺叶高山栎中发现了 358 661 个和 346 281 个单核苷酸多态性位点(SNP)。本研究中得到的下机数据和 SNP 位点是到目前为止获得的川滇高山栎和刺叶高山栎最饱和的遗传标记数据集。以重测序数据集为基础加上本实验室的 nSSR 数据集研究了川滇高山栎和刺叶高山栎的遗传多样性和分化、遗传变异对气候的响应模式、与生态适应相关的基因组特征以及气候变化的未来命运, 主要得到以下结论:

(1) 川滇高山栎和刺叶高山栎都可以被划分成两个不同的遗传类群, 根据地理位置的不同命名为东西谱系, 分别为于青藏高原地区的 *Q. aqu-West*, 分布于横断山脉地区的 *Q. aqu-East* 和 *Q. spi-West* 谱系, 以及分布于秦岭山脉和中国东部温暖低地的 *Q. spi-East* 谱系。ABC 种群统计分析表明川滇高山栎的分化时间早于刺叶高山栎, 然后两物种逐渐发生了种内遗传分化, 且川滇高山栎的分化水平高于刺叶高山栎。两栎树物种的种内分化可能与青藏高原造山运动相关, 复杂的山脉地形是栎树种子和花粉传播的有效地理屏障, 为这两物种种内谱系的独立进化提供了潜在条件。川滇高山栎的遗传多样性水平高于刺叶高山栎, 且两物种种内谱系间西部地区的遗传多样性水平要明显高于东部谱系, 可能与西部谱系种群间高遗传分化水平相关。

(2) 基因型环境关联分析在川滇高山栎和刺叶高山栎的谱系间揭示了不同的遗传变异的响应模式。在 *Q. aqu-East* 谱系中发现了显著的 IBE 模式, 但是在 *Q. aqu-West* 谱系中并没有发现遗传距离和地理或气候距离有显著的相关性。在 *Q. spi-East* 谱系发现了显著的 IBD 模式, 相反在 *Q. spi-West* 谱系发现了显著的 IBE 模式, 揭示了 *Q. aqu-East* 和 *Q. spi-West* 谱系的遗传变异主要来自环境的强大选择压力。同时我们还发现 bio07 是 *Q. aqu-East* 谱系遗传变异最重要的气候因子, prec03 是驱动 *Q. spi-West* 谱系遗传变异最重要的环境因子。表明温度和降水分别在塑造 *Q. aqu-East* 和 *Q. spi-West* 谱系种群遗传变异中起着重要作用。

(3) 我们使用异常位点检测方法在川滇高山栎和刺叶高山栎的基因组中发现了自然选择的遗传印记。基于 BayeScan 方法发现有 26 个 SNP 在川滇高山栎所有种群和 *Q. aqu-East* 谱系均被视为异常值, 在刺叶高山栎所有种群和 *Q. spi-West* 谱系中均检测到 14 个异常 SNP。基于 LFMM 的方法在川滇高山栎和刺叶高山栎的谱系内均发现与四个气候因子生态适应相关的异常 SNP 位点, 表明了环境因素作为重要的选

择压力的作用，可能会驱动遗传变异。

(4) 基于空间遗传多样性结果我们发现川滇高山栎和刺叶高山栎的优先保护区不同，这一结果也被基因组脆弱性区域的预测结果所验证。川滇高山栎在整个分布区的北部和西藏地区西部边缘预测到了相对较高的遗传偏移，在刺叶高山栎分布区的北部边缘地区预测到了相对较高的遗传偏移，这些地区的种群必须转移更多的潜在适应性等位基因频率，以适应未来气候变化。RONA 分析揭示了位于分布区边缘的种群有着最大的灭绝风险，同时三月份降水量(prec03)是最具代表性的环境因素。

5.2 展望

(1) 种群对气候变化的实际进化反应将比预测的更加复杂，因为适应是多种进化过程的结果，比如迁移、突变、重组和进化历史动态也会影响物种的生态适应能力。因此在未来研究中要考虑种群动态对物种遗传变异的影响，比如基因流的大小和方向以及有效种群大小的变化。

(2) 基因组扫描可能对具有广泛连锁不平衡、搭便车效应或适应性隔离(isolation-by-adaptation)的物种更有效。具有不同降水和温度的地形复杂区域，例如我们的研究区域，可能比同质的区域更容易推动选择。基因组扫描可以在染色体水平识别影响局部适应模式的与环境梯度相关的异常位点，与适应性基因组区域相关的或通过适应表现出分离的中性位点可能具有很高的信息量。另外基因组扫描可以与田间试验（如同质园实验）相结合，以验证假定适应性遗传变异的功能相关性。

(3) 我们只确定了可能参与生态适应过程的单核苷酸多态性位点，并没有深入研究这些位点的具体功能。未来研究中可以将可能参与生态适应的位点进行功能富集分析，从基因功能方面揭示物种的生态适应的遗传基础。另外还可以将遗传变异信息和表型变异信息相结合，可能有助于确定假定的候选基因。

(4) 我们对川滇高山栎和刺叶高山栎未来命运的预测是基于对基因组信息的模拟，而没有考虑其他潜在因素的影响，例如表型可塑性和种群平衡阶段对局部适应的影响。未来研究中需要进一步研究整个基因组的特征，以便在选择压力下获得更精确的基因功能，并阐明潜在的多基因适应模式。此外，可以将预测模型与生态位模型结合以及阻力分析相结合，更全面地研究物种当前和未来分布模式。这种非线性、多变量的环境关联方法可能有助于通过确定合适的种子来源来指导种子选择，从而对未来气候做出更充分的反应

(5) 目前用于评估和揭示树种适应性变异的景观基因组学研究只关注单一物种，而不是群落水平。通过分析同一景观中的多个物种，可以更好的评估物种和景观之间的适应性共性，从而更全面的描述物种的适应性遗传机制。

参考文献

- 陈焕镛, 黄成就. 中国植物志: 第二十二卷[M]. 北京: 科学出版社, 1998.
- 端木炘. 我国栎属资源的综合利用[J]. 河北林果研究, 1994, 9 (2): 177-181.
- 王天瑞. 基于微卫星分子标记的川滇和刺叶高山栎的种群和景观遗传学研究[D]. 北京林业大学, 2020.
- 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. 中国科学: 生命科学, 2021, 51: 167-178.
- 王玉垚. 基于干旱及缺氧候选基因的生态适应研究: 以川滇高山栎为例[D]. 北京林业大学, 2018.
- Adopted I. Climate change 2014 synthesis report[J]. IPCC: Geneva, Switzerland, 2014.
- Aitken S N, Yeaman S, Holliday J A, Wang T, Curtis-McLane S. Adaptation, migration or extirpation: climate change outcomes for tree populations[J]. Evolutionary Applications, 2008, 1(1): 95-111.
- Aldrich P R, Cavender-Bares J. *Quercus*[M]. Wild crop relatives: Genomic and breeding resources. Springer, Berlin, Heidelberg, 2011: 89-129.
- Alexander D H, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals[J]. Genome Research, 2009, 19: 1655-1664.
- Bay R A, Harrigan R J, Le U V, Gibbs H L, Smith T B, Ruegg K. Genomic signals of selection predict climate-driven population declines in a migratory bird[J]. Science, 2018, 359(6371): 83-86.
- Beaumont M A, Balding D J. Identifying adaptive genetic divergence among populations from genome scans[J]. Molecular Ecology, 2010, 13(4): 969-980.
- Beaumont M A, Nichols R A. Evaluating loci for use in the genetic analysis of population structure[J]. Proceedings of the Royal Society of London, 1996, 263(1377): 1619-1626.
- Borrell J S, Zohren J, Nichols R A, Buggs R. Genomic assessment of local adaptation in dwarf birch to inform assisted gene flow[J]. Evolutionary Applications, 2020, 13(1): 161-175.
- Bragg J G, Supple M A, Andrew R L, Borevitz J O. Genomic variation across landscapes: insights and applications[J]. New Phytologist, 2015, 207: 953-967.
- Breed M F, Harrison P. A., Blyth C, Prowse T A A, Steane D A, Mohr J J. The potential of genomics for restoring ecosystems and biodiversity[J]. Nature Reviews Genetics, 2019, 20(10): 615-628.
- Breiman L. Statistical modeling: The two cultures (with comments and a rejoinder by the author) [J]. Statistical Science, 2001, 16(3): 199-231.
- Browne L, Wright J W, Fitz-Gibbon S, Gugge P F, Sork V L. Adaptational lag to temperature in valley oak (*Quercus lobata*) can be mitigated by genome-informed assisted gene flow[J]. Proceedings of the National Academy of Sciences of the United States of America, 2019, 116(50): 25179-25185.
- Burger W C. The species concept in *Quercus*[J]. Taxon, 1975, 24, 45-50.

- Catchen J, Hohenlohe P A, Bassham S, Amores A, Cresko W A. Stacks: An analysis tool set for population genomics[J]. *Molecular Ecology*, 2013, 22(11): 3124-3140.
- 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: 669-692.
- Chen D M, Kang H Z, Liu C J. An overview on the potential quaternary glacial refugia of plants in China Mainland[J]. *Bulletin of Botanical Research*, 2011, 31(5): 623-632.
- Chen Y, Chen Y, Shi C, Huang Z, Zhang Y, Li S. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data[J]. *Gigascience*, 2018, 7(1):1-6.
- Coop G, Witonsky D, Rienzo A D, Pritchard J K. Using environmental correlations to identify loci underlying local adaptation[J]. *Genetics*, 2010, 185(4):1411-1423.
- Cornuet J M, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, Marin J M, Estoup A. DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data[J]. *Bioinformatics*, 2014, 30: 1187-1189.
- Danecek P, Auton A, Abecasis G, Albers C A, Anks E B, Depristo M A. The variant call format and VCFtools[J]. *Bioinformatics*, 2011, 27(15): 2156-8.
- De Beaulieu A L H, Lamant T. Guide illustré des chênes[M]. Edilens, 2010.
- Denk T, Grimm G W, Manos P S, Deng M, Hipp A. An updated infrageneric classification of the oaks: review of previous taxonomic schemes and synthesis of evolutionary patterns[J]. *bioRxiv*, 2017, 13-38.
- Diffenbaugh N S, Singh D, Mankin J S. Unprecedented climate events: Historical changes, aspirational targets, and national commitments[J]. *Science Advances*, 2018, 4(2): eaao3354.
- 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.
- Du F K, Wang T R, Wang Y Y, Ueno S, de Lafontaine, G. Contrasted patterns of local adaptation to climate change across the range of an evergreen oak, *Quercus aquifolioides*[J]. *Evolutionary Applications*, 2020, 00: 1-15.
- Ellis N, Smith S J, Pitcher C R. Gradient forests: calculating importance gradients on physical predictors[J]. *Ecology*, 2012, 93(1): 156-168.
- Etterson J R, Shaw R G. Constraint to adaptive evolution in response to global warming[J]. *Science*, 2001, 294: 151-154.
- Eveno E, Collada C, Guevara M A, Leger V, Soto A, Diaz L, Garnier-Gere P H. Contrasting patterns of

- selection at *Pinus pinaster* Ait. Drought stress candidate genes as revealed by genetic differentiation analyses[J]. *Molecular Biology and Evolution*, 2008, 25: 417-437.
- Excoffier L, Hofer T, Foll M. Detecting loci under selection in a hierarchically structured population[J]. *Heredity*, 2009, 103(4): 285-298.
- Excoffier L, Lischer H E L. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows[J]. *Molecular Ecology Resources*, 2010, 10(3): 564-567.
- Favre A, Päckert M, Pauls S U, Jähnig S C, Uhl D, Michalak I. The role of the uplift of the Qinghai-Tibetan Plateau for the evolution of Tibetan biotas[J]. *Biological reviews of the Cambridge Philosophical Society*, 2015, 90: 236-253.
- Ferguson D K. Oaks of Asia[J]. *Systematic Botany*, 2006, 31(4): 881-881.
- Felsenstein J. The theoretical population genetics of variable selection and migration[J]. *Annual Review of Genetics*, 1976, 10: 253-280.
- Feng L, Du F K. Landscape genomics in tree conservation under a changing environment[J]. *Front. Plant Sci*, 2022, 13: 822217.
- Feng L, Zheng Q J, Qian Z Q, et al. Genetic structure and evolutionary history of three Alpine Sclerophyllous oaks in East Himalaya-Hengduan Mountains and adjacent regions[J]. *Frontiers in plant science*, 2016, 7: 1688.
- Ferrier S, Manion G, Elith J, Richardson K. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment[J]. *Diversity and Distributions*, 2007, 13(3): 252-264.
- 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.
- Fisher R A. The Genetical theory of natural selection[J]. *Biological Reviews*, 1930, 1(4): 1419
- 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.
- Fitzpatrick M C, Sanders N J, Normand S, Svenning J C, Ferrier S, Gove A D, Dunn R R. Environmental and historical imprints on beta diversity: Insights from variation in rates of species turnover along gradients[J]. *Proceedings of the Royal Society B Biological Sciences*, 2013, 280(1768): 20131201.
- Foll M. BayeScan v2. 1 User manual[J]. *Ecology*, 2012, 20: 1450-1462.
- Forester B R, Lasky J R, Wagner H H, Urban D L. Comparing methods for detecting multilocus adaptation with multivariate genotype-environment associations[J]. *Molecular Ecology*, 2018, 27(9): 2215-2233.

- Fournier-Level A, Korte A, Cooper M D, Nordborg M, Schmitt J, Wilczek A M. A map of local adaptation in *Arabidopsis thaliana*[J]. *Science*, 2011, 334: 86-89.
- Frankel O H, Brown A H D, Burdon J. The genetic diversity of wild plants[J]. The conservation of plant biodiversity, Cambridge University Press, UK. 1995, 10-38.
- 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.
- Frichot E, Schoville S D, Bouchard G, François O. Testing for associations between loci and environmental gradients using latent factor mixed models[J]. *Society for Molecular Biology and Evolution*, 2013, 30: 1687-1699.
- Gao Y, Yin S, Chu H, Zhang Y, Wang H, Chen H, Liu C, Dai D, Tang L. Genome-Wide SNPs provide insights on the cryptic genetic structure and signatures of climate adaption in *Amorphophallus albus* Germplasms[J]. *Frontiers in Plant Science*, 2021, 12: 1554.
- García-Ramos G, Kirkpatrick M. Genetic models of adaptation and gene flow in peripheral populations[J]. *Evolution*, 1997, 51: 21-28.
- Gong W, Chen C, Dobeš C, Fu C X, Koch M A. Phylogeography of a living fossil: Pleistocene glaciations forced *Ginkgo biloba* L. (Ginkgoaceae) into two refuge areas in China with limited subsequent postglacial expansion[J]. *Molecular Phylogenetics and Evolution*, 2008, 48: 1094-1105.
- Goslee S C, Urban D L. The ecodist package for dissimilarity-based analysis of ecological data[J]. *Journal of Statistical Software*, 2007, 22(7): 1-19.
- Gugger P F, Liang C T, Sork V L, Hodgskiss P, Wright J W. Applying landscape genomic tools to forest management and restoration of Hawaiian koa (*Acacia koa*) in a changing environment[J]. *Evolutionary Applications*, 2018, 11: 231-242.
- Guillot G, Rousset F. Dismantling the Mantel tests[J]. *Methods in Ecology and Evolution*, 2013, 4(4).
- Haldane J B S, The Causes of Evolution[M]. New York: Harper, 1932.
- Hamrick J L, Godt M J W, Sherman-Broyles S L. Factors influencing levels of genetic diversity in woody plant species[M]. *Population genetics of forest trees*. Springer, Dordrecht, 1992, 95-124.
- Hansen M M, Olivieri I, Waller D M, Nielsen E E. Monitoring adaptive genetic responses to environmental change[J]. *Molecular Ecology*, 2012, 21: 1311-1329.
- Hardy G H. Mendelian proportions in a mixed population[J]. *Science*, 1908, 28(706): 49-50.
- Harte J, Saleska S R, Levy C. Convergent ecosystem responses to 23-year ambient and manipulated warming link advancing snowmelt and shrub encroachment to transient and long-term climate–soil carbon feedback[J]. *Global Chang Biology*, 2015, 21(6): 2349-2356.
- Hijmans R J. Geosphere: spherical trigonometry. R package geosphere version 1.5-7[J]. 2014, 1549-1553.
- Hipp A L, Manos P S, González-Rodríguez A, Hahn M, Kaproth M, McVay J D , Avalos S V, Cavender-

- Bares J. Sympatric parallel diversification of major oak clades in the Americas and the origins of Mexican species diversity[J]. *New Phytologist*, 2018, 217: 439-452.
- Hoban S, Kelley J L, Lotterhos K E, Antolin M F, Bradburd G, Lowry D B. Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions[J]. *American Naturalist*, 2016, 188: 379.
- Hohenlohe P A, Phillips P C, Cresko W A. Using population genomics to detect selection in natural populations: key concepts and methodological considerations[J]. *International Journal of Plant Sciences*, 2010, 171(9): 1059-1071.
- Holderegger R, Kamm U, Gugerli F. Adaptive vs. neutral genetic diversity: implications for landscape genetics[J]. *Landscape Ecology*, 2006, 21(6): 797-807.
- Holliday J A, Aitken S N, Cooke J E, Staton M, Whetten R W, Plomion C. Advances in ecological genomics in forest trees and applications to genetic resources conservation and breeding[J]. *Molecular Ecology*. 2017, 26, 706-717.
- Hughes A R, Inouye B D, Johnson M T, Underwood N, Vellend M. Ecological consequences of genetic diversity[J]. *Ecology Letters*, 2008, 11(6): 609-623.
- Isabel N, Holliday J A, Aitken S N. Forest genomics: advancing climate adaptation, forest health, productivity, and conservation. *Evolutionary Applications*, 2020, 13(1): 3-10.
- Jia K H, Zhao W, Maier P A, Wang T, Wang X R, Mao J F. Landscape genomics predicts climate change-related genetic offset for the widespread *Platycladus orientalis* (Cupressaceae) [J]. *Evolutionary Applications*, 2020, 13(4): 665-676.
- Joshi J, Schmid B, Caldeira M C, Dimitrakopoulos P G, Good J, Harris R, Lawton J H. Local adaptation enhances performance of common plant species[J]. *Ecology Letters*, 2001, 4: 536-544.
- Ju M M, Feng L, Yang J, Yang Y C, Chen X D, Zhao G F. Evaluating population genetic structure and demographic history of *Quercus spinosa* (Fagaceae) based on specific length amplified fragment sequencing[J]. *Frontiers in genetics*, 2019: 965.
- Kalinowski S T. HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness[J]. *Molecular Ecology Notes*, 2005, 5: 187-189.
- Kawecki T J, Ebert D. Conceptual issues in local adaptation[J]. *Ecology Letters*, 2004, 7(12).
- Kremer A, Hipp A L. Oaks: an evolutionary success story[J]. *New Phytologist*, 2020, 226(4): 987-1011.
- Le Corre V, Fabrice R, Xavier R. DNA polymorphism at the *FRIGIDA* gene in *Arabidopsis thaliana*: extensive nonsynonymous variation is consistent with local selection for flowering time[J]. *Molecular Biology and Evolution*, 2002(8): 1261-1271.
- Legendre P, Anderson M J. Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments[J]. *Ecological Monographs*, 1998, 69: 1-24.

- Legendre P, Legendre L. Complex ecological data sets[M]. Elsevier, 2012, 24: 1-57.
- Lenormand T. Gene flow and the limits to natural selection[J]. Trends in Ecology and Evolution, 2002, 17: 183-189.
- Leroy T, Louvet J M, Lalanne C, Le Provost G, Labadie K, Aury J M, Delzon S, Plomion C, Kremer A. Adaptive introgression as a driver of local adaptation to climate in European white oaks [J]. New Phytologist, 2020, 226(4):1171-1182.
- Li H, Durbin R. Fast and accurate long-read alignment with burrows-wheeler transform[J]. Bioinformatics, 2010, 26: 589-595.
- Li L, Abbott R J, Liu B B, Sun Y S, Li L L, Zou J. Pliocene intraspecific divergence and Plio-Pleistocene range expansions within *Picea likiangensis* (Lijiang spruce), a dominant forest tree of the Qinghai-Tibet Plateau[J]. Molecular Ecology, 2013, 22(20): 5237-5255.
- Li Y, Zhang X X, Mao R L, Yang J, Miao C Y, Li Z, Qiu Y X. Ten Years of Landscape Genomics: Challenges and Opportunities[J]. Frontiers in Plant Science, 2017, 8: 2136.
- Li Y, Zhang Y, Liao P C, Wang T, Wang X, Ueno, S., & Du, F. K. 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), 1-18.
- Liao Z, Nobis M P, Xiong Q, et al. Potential distributions of seven sympatric sclerophyllous oak species in Southwest China depend on climatic, non-climatic, and independent spatial drivers[J]. Annals of Forest Science, 2021, 78(1): 1-22.
- Liancourt P, Spence L A, Song D S, Lkhagva A, Sharkhuu A, Boldgiv B. Plant response to climate change varies with topography, interactions with neighbors, and ecotype[J]. Ecology, 2013, 94: 444-453.
- Liang Y Y, Shi Y, Yuan S, Zhou B F, Chen X Y, An Q Q, Ingvarsson P K, Plomion C, Wang B S. Linked selection shapes the landscape of genomic variation in three oak species[J]. New Phytologist, 2022, 233(1): 555-568.
- Lichstein J W. Multiple regression on distance matrices: a multivariate spatial analysis tool[J]. Plant Ecology, 2007, 188: 117-131.
- Lischer H E L, Excoffier L. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs[J]. Bioinformatics, 2012, 28(2): 298-299.
- Liu K, Qi M. and Du F K. Population and Landscape Genetics provide insights into species conservation of two evergreen oaks in Qinghai-Tibet plateau and adjacent regions[J]. Frontiers in. plant Science. 2022, 13:858526.
- Liu J, Möller M, Provan J, Gao L M, Poudel R C, Li D Z. Geological and ecological factors drive cryptic speciation of yews in a biodiversity hotspot[J]. New Phytologist, 2013, 199: 1093-1108.
- Liu Y, Li Y, Song J, Song J L, Zhang R P, Yan Y, Wang Y Y, Du F K. Geometric morphometric analyses

- of leaf shapes in two sympatric Chinese oaks: *Quercus dentata* Thunberg and *Quercus aliena* Blume (Fagaceae)[J]. *Annals of Forest Science*, 2018, 75(4): 1-12.
- López-Pujol J, Zhang F M, Sun H Q, Ying T S, Ge S. Centres of plant endemism in China: places for survival or for speciation? [J]. *Journal of Biogeography*, 2011, 38(7): 1267-1280.
- Manel S, Joost S, Epperson B K, Holderegger R, Storfer A, Rosenberg M S, Scribner K T, Bonin A, Fortin M J. Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field[J]. *Molecular Ecology*, 2010, 19(17): 3760-3772.
- Manion G, Lisk M, Ferrier S. Package 'gdm', A toolkit with functions to fit, plot, and summarize Generalized Dissimilarity Models: CRAN Repository, R[J], 2018.
- Manthey J D, Moyle R G. Isolation by environment in White-breasted Nuthatches (*Sitta carolinensis*) of the Madrean Archipelago sky islands: a landscape genomics approach[J]. *Molecular Ecology*, 2015, 24(14): 3628-3638.
- Martins K, Gugger P F, Llanderal-Mendoza J, Gonzalez-Rodriguez A, Fitz-Gibbon S T, Zhao J L, Rodriguez-Correa H, Oyama K, Sork V L. Landscape genomics provides evidence of climate-associated genetic variation in Mexican populations of *Quercus rugosa*[J]. *Evolutionary Applications*, 2018, 11: 1842-1858.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A. The Genome Analysis Toolkit: a MapReduce framework for analyzing next generation DNA sequencing data[J]. *Genome Research*, 2010, 20(9): 1297-1303.
- McKown A D, Klápště J, Guy R D, El-Kassaby Y A, Mansfield, S D. Ecological genomics of variation in bud-break phenology and mechanisms of response to climate warming in *Populus trichocarpa*[J]. *New Phytologist*, 2018, 220(1): 300-316.
- 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
- Moeller D A, Tiffin P. Geographic variation in adaptation at the molecular level: a case study of plant immunity genes[J]. *Evolution*, 2008, 62: 3069-3081.
- Mosca E, Eckert A J, Pierro E A, Rocchini D, Porta N, Belletti P, Neale D B. The geographical and environmental determinants of genetic diversity for four alpine conifers of the European Alps[J]. *Molecular Ecology*, 2012, 21: 5530-5545.
- Naimi B, Hamm N A S, Groen T A, Skidmore A K, Toxopeus A G. Where is positional uncertainty a problem for species distribution modelling?[J]. *Ecography*, 2014, 37(2): 191-203.
- Neale D B, Kremer A. Forest tree genomics: growing resources and applications[J]. *Nature Reviews Genetics*, 2011, 12(2): 111-122.
- Oksanen J, Blanchet F G, Kindt R, Legendre P, Minchin P R, O Hara R B, Simpson G L, Solymos P,

- Stevens M H H, Wagner H. vegan: Community ecology package[J]. R package version, 2017, 14(6), 927-930
- Oldfield S, Eastwood A. Red list of oaks[M]. Fauna and Flora International, 2007.
- Peakall R, Smouse P E. GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research[J]. Molecular Ecology Notes, 2006, 6: 288-295.
- Petit RJ, Bodénès C, Ducouso A, Roussel G, Kremer A. Hybridization as a mechanism of invasion in oaks [J]. New Phytologist, 2004, 161(1):151-164
- Petit R J, Hampe A. Some evolutionary consequences of being a tree[J]. Annual Review of Ecology, Evolution, and Systematics, 2006, 37: 187-214.
- Petit R J, Mousadik A, Pons O. Identifying populations for conservation on the basis of genetic markers[J]. Biological Conservation, 1998, 12: 844-855.
- Plomion C, Aury J M, Amselem J, Quesneville H, Kremer A, Salse J. Oak genome reveals facets of long lifespan[J]. Nature Plants, 2018, 4(7): 440-452.
- Prugnolle F, Manica A, Charpentier M, Guégan J F, Guernier V, Balloux F. Pathogen-driven selection and worldwide HLA class I diversity[J]. Current biology, 2005, 15(11): 1022-1027.
- Purcell S, Neale B, Todd-Brown K, Thomas, L, Ferreira M A R, Bender D. 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.
- 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: 225-244.
- Redden R. New approaches for crop genetic adaptation to the abiotic stresses predicted with climate change[J]. Agronomy, 2013, 3: 419-432.
- Rellstab C, Gugerli F, Eckert A J, Hancock A M, Holderegger R. A practical guide to environmental association analysis in landscape genomics[J]. Molecular Ecology, 2015, 24: 4348-4370.
- Rellstab C, Zoller S, Walthert L, Lesur I, Pluess A R, Graf R, Gugerli F. Signatures of local adaptation in candidate genes of oaks (*Quercus* spp.) with respect to present and future climatic conditions[J]. Molecular Ecology, 2016, 25: 5907-5924.
- Richards E, Reichardt M, Rogers S. Preparation of genomic DNA from plant tissue[J]. Current Protocols in Molecular Biology, 1994, 27(1): 231-237.
- Ruegg K, Bay R A, Anderson E C, Saracco J F, Harrigan R J, Whitfield M, Paxton E H, Smith T B. Ecological genomics predicts climate vulnerability in an endangered Southwestern songbird[J]. Ecology Letters, 2018, 21(7): 1085-1096.
- Rushton B S. Natural hybridization within the genus *Quercus* L[C]. Annales des sciences forestières. EDP

- Sciences, 1993, 50(Supplement): 73s-90s.
- Savolainen O, Lascoux M, Merilä J. Ecological genomics of local adaptation[J]. Nature Reviews Genetics, 2013, 14: 807-820.
- Shaw R G, Etterson J R. Rapid climate change and the rate of adaptation: Insight from experimental quantitative genetics[J]. New Phytologist, 2012, 195(4): 752-765.
- Skelton R P, Anderegg LDL, Diaz J, Klinga M M, Pappera P, Lamarqued L J, Delzone S, Dawson T E, Ackerly D D. Evolutionary relationships between drought-related traits and climate shape large hydraulic safety margins in western North American oaks[J]. Proceedings of the National Academy of Science, 2021, 118.
- Slatkin M. Isolation by distance in equilibrium and non-equilibrium populations[J]. Evolution, 1993, 47(1): 264-279.
- Smouse P E, Long J C, Sokal R R. Multiple regression and correlation extensions of the Mantel test of matrix correspondence[J]. Systematic Biology, 1986, 35: 627-632.
- Sobel J M, Chen G F, Watt L R, Schemske D W. The biology of speciation[J]. Evolution, 2010, 64(4): 295-315.
- Sork V L, Aitken S N, Dyer R J, Eckert A J, Legendre P, Neale D B. Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate[J]. Tree Genetics and Genomes, 2013, 9: 901-911.
- Sork V L, Squire K, Gugger P F, Steele S E, Levy E D, Eckert A J. Landscape genomic analysis of candidate genes for climate adaptation in a California endemic oak, *Quercus lobata*[J]. American Journal of Botany, 2016, 103(1): 33-46.
- Sork V L, Waits L. Contributions of landscape genetics-approaches, insights, and future potential[J]. Molecular Ecology, 2010, 19: 3489-3495.
- Sork V L. Genomic studies of local adaptation in natural plant populations[J]. Journal of Heredity, 2018, 109: 3-15.
- Spitze K. Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation[J]. Genetics, 1993, 135: 367-374.
- Storz J F, Sabatino S J, Hoffmann F G, Gering E J, Moriyama H, Ferrand N. The molecular basis of high-altitude adaptation in deer mice[J]. PLoS Genetics, 2007, 3: 448-459.
- Stucki S, Orozco-Terwengel P, Forester B R, Duruz S, Colli L, Masembe C. High performance computation of landscape genomic models including local indicators of spatial association[J]. Molecular Ecology Resources, 2017, 17(5): 1072-1089.
- Van Zonneveld M, Scheldeman X, Escribano P, Viruel M A, Van D P, Garcia W, Tapia C, Romero J, Siqueñas M, Hormaza J I. Mapping genetic diversity of cherimoya (*Annona cherimola* Mill.):

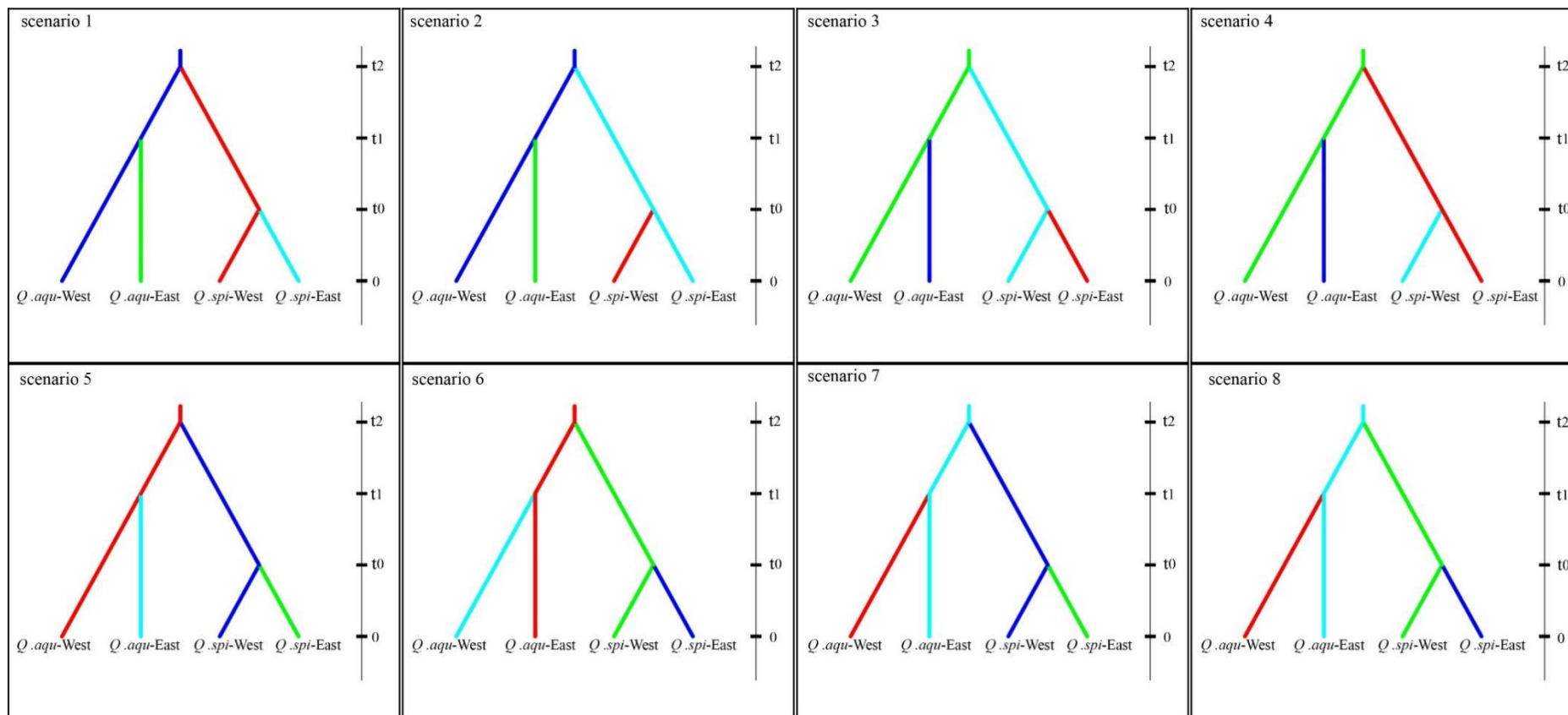
- application of spatial analysis for conservation and use of plant genetic resources[J]. PLoS One, 2012, 7(1).
- Van Strien M, Holderegger R, Van Heck H. Isolation-by-distance in landscapes: Considerations for landscape genetics[J]. Heredity, 2015, 114(1): 27-37.
- Vitalis R, Dawson K, Boursot P. Interpretation of variation across marker loci as evidence of selection[J]. Genetics, 2001, 158: 1811-1823.
- Wahid A, Gelani S, Ashraf M, Foolad M R. Heat tolerance in plants: An overview[J]. Environmental & Experimental Botany, 2007, 61(3): 199-223.
- Wang I J, Bradburd G S. Isolation by environment[J]. Molecular Ecology, 2014, 23: 5649-5662.
- Wang I J, Summers K. Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog[J]. Molecular Ecology, 2010, 19: 447-458.
- Wang J, Street N R, Scofield D G, Ingvarsson P K. Variation in linked selection and recombination drive genomic divergence during allopatric speciation of European and American aspens[J]. Molecular Biology and Evolution, 2016, 33(7): 1754-1767.
- Wang T, O'Neill G A, Aitken S N. Integrating environmental and genetic effects to predict responses of tree populations to climate[J]. Ecological Applications, 2010, 20: 153-163.
- Weinberg W. On the demonstration of heredity in man[J]. Papers on human genetics, 1908.
- Wen J, Zhang J Q, Nie Z L, Zhong Y, Sun H. Evolutionary diversifications of plants on the Qinghai-Tibetan Plateau[J]. Frontiers in Genetics, 2014, 5: 4.
- Wernberg T, Coleman M A, Bennett S, Thomsen M S, Tuya F, Kelaher B P. Genetic diversity and kelp forest vulnerability to climatic stress[J]. Scientific Reports, 2018, 8(1): 1851.
- Whitlock M C. Evolutionary inference from Q_{ST} [J]. Molecular Ecology, 2008, 17: 1885-1896.
- Whitlock M C. Modern approaches to local adaptation[J]. The American Naturalist, 2015, 186(S1): S1-S4.
- Wilczek A M, Cooper M D, Korves T M, Schmitt J. Lagging adaptation to warming climate in *Arabidopsis thaliana*[J]. Proceedings of the National Academy of Sciences, 2014, 111: 7906-7913.
- Willis C, Ruhfel B, Primack R, Miller-Rushing A, Davis C. Phylogenetic patterns of species loss in Thoreau's woods are driven by climate change[J]. Proceedings of the National Academy of Sciences, 2008, 105: 17029-17033.
- Wright S I, Gaut B S. Molecular population genetics and the search for adaptive evolution in plants[J]. Society for Molecular Biology and Evolution, 2004, 22: 506-519.
- Wright S. Evolution in Mendelian populations[J]. Genetics, 1931, 16: 97-159.
- Wright S. The genetical structure of populations[J]. Annals of Eugenics, 1949, 15(1): 323-354.

- Xu J, Deng M, Jiang X L, Westwood M, Song Y G, Turkington R. Phylogeography of *Quercus glauca* (Fagaceae), a dominant tree of East Asian subtropical evergreen forests, based on three chloroplast DNA interspace sequences. *Tree Genetics and Genomes*, 2015, 11(1):1-17.
- Yeaman S, Whitlock M C. The genetic architecture of adaptation under migration–selection balance[J]. *Evolution*, 2011, 65: 1897-1911.
- Zhao Y P, Fan G, Yin P P, Sun S, Ge S. Resequencing 545 ginkgo genomes across the world reveals the evolutionary history of the living fossil[J]. *Nature Communications*, 2019, 10(1): 1-10.
- Zhou B F, Shi Y, Chen X Y, Yuan X Y, Liang Y Y, Wang B S. Linked selection, ancient polymorphism, and ecological adaptation shape the genomic landscape of divergence in *Quercus dentata*[J]. *Journal of Systematics and Evolution*. 2022, 00: 1-14.
- Zhou Z K. Origin, phylogeny and dispersal of *Quercus* from China[J]. *Acta Botanica Yunnanica*, 1992, 14(3): 227-236.

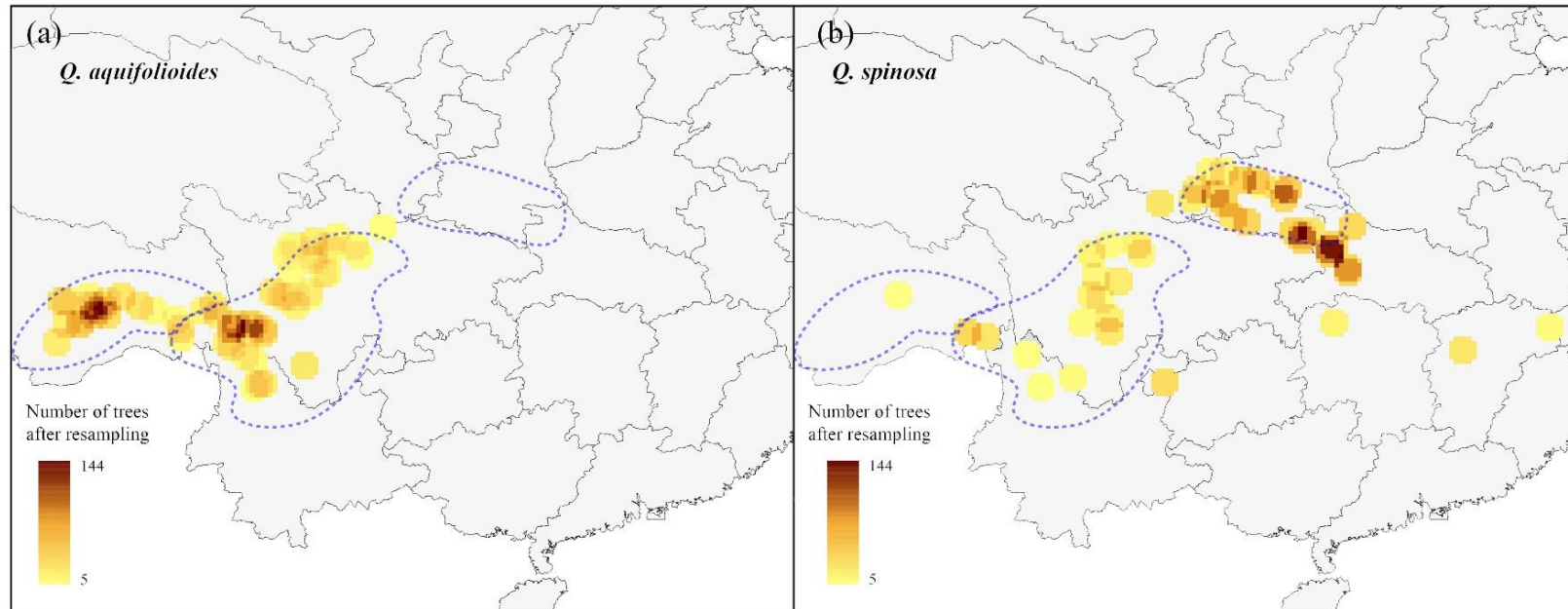
附录

附录 A 川滇高山栎和刺叶高山栎 DIY-ABC 测试的 8 种进化场景

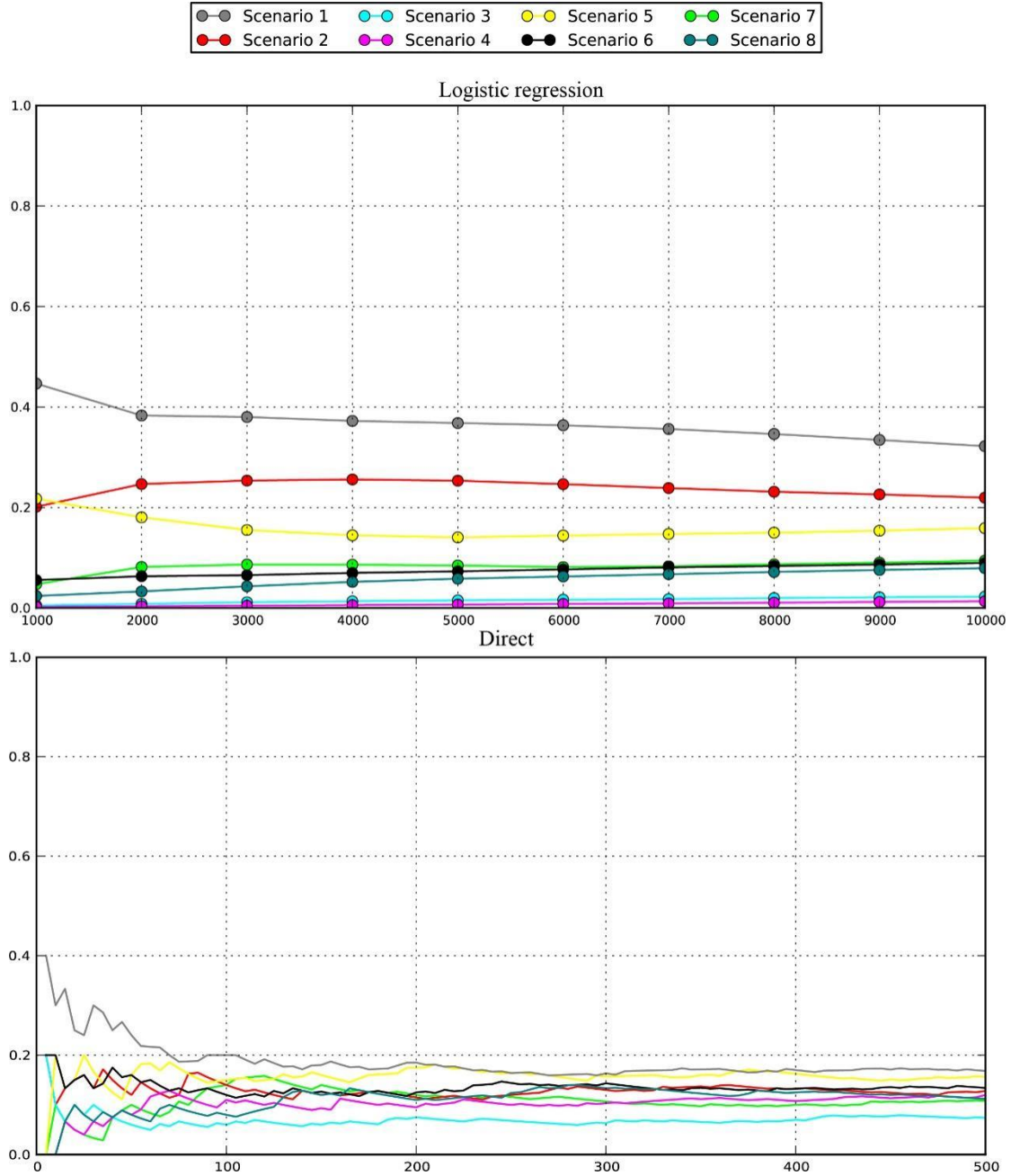
Appendix A. Assuming colonization scenarios tested of *Q. aquifolioides* and *Q. spinosa* using DIY-ABC



附录 B 川滇高山栎和刺叶高山栎重新采样后每个网格的树分布图地
Appendix B. The map of Number of trees per grid after resampling of *Q. aquifolioides* and *Q. spinosa*



附录 C ABC 模型两种线性回归方法（直接估计和逻辑回归）在不同场景的结果
Appendix C. The results of two linear regression methods (direct estimation and logistic regression) in the ABC model in different scenarios



附录 D 川滇高山栎和刺叶高山栎 DIY-ABC 分析 8 个场景的概率值及其 95% 置信区间结果
 Appendix D. The probability values with their 95% confidence intervals of *Q. aquifolioides* and *Q. spinosa* with DIY-ABC

n	scenario1	scenario2	scenario3	scenario4	scenario5	scenario6	scenario7	scenario8
<i>Direct estimate</i>								
50	0.1[0.0000,0.36 30]	0.12[0.0000,0.4 048]	0.06[0.0000,0.2 682]	0.08[0.0000,0.3 178]	0.16[0.0000,0.4 813]	0.16[0.0000,0.4 813]	0.24[0.0000,0.6 144]	0.08[0.0000,0.3 178]
100	0.14[0.0000,0.4 441]	0.14[0.0000,0.4 441]	0.06[0.0000,0.2 682]	0.11[0.0000,0.3 843]	0.15[0.0000,0.4 630]	0.12[0.0000,0.4 048]	0.2[0.0000,0.55 06]	0.08[0.0000,0.3 178]
150	0.14[0.0000,0.4 441]	0.12[0.0000,0.4 048]	0.06[0.0000,0.2 682]	0.0933[0.0000,0 .3483]	0.16[0.0000,0.4 813]	0.1267[0.0000,0 .4182]	0.18[0.0000,0.5 168]	0.12[0.0000,0.4 048]
200	0.12[0.0000,0.4 048]	0.115[0.0000,0. 3946]	0.075[0.0000,0. 3059]	0.095[0.0000,0. 3520]	0.175[0.0000,0. 5081]	0.125[0.0000,0. 4149]	0.185[0.0000,0. 5254]	0.11[0.0000,0.3 843]
250	0.116[0.0000,0. 3967]	0.12[0.0000,0.4 048]	0.068[0.0000,0. 2887]	0.1[0.0000,0.36 30]	0.164[0.0000,0. 4886]	0.144[0.0000,0. 4517]	0.164[0.0000,0. 4886]	0.124[0.0000,0. 4129]
300	0.1067[0.0000,0 .3772]	0.13[0.0000,0.4 248]	0.0633[0.0000,0 .2768]	0.1033[0.0000,0 .3701]	0.1567[0.0000,0 .4753]	0.1433[0.0000,0 .4505]	0.1633[0.0000,0 .4874]	0.1333[0.0000,0 .4313]
350	0.0971[0.0000,0 .3567]	0.1371[0.0000,0 .4387]	0.0657[0.0000,0 .2829]	0.1114[0.0000,0 .3872]	0.16[0.0000,0.4 813]	0.1343[0.0000,0 .4331]	0.1714[0.0000,0 .5018]	0.1229[0.0000,0 .4106]
400	0.1[0.0000,0.36 30]	0.1325[0.0000,0 .4297]	0.07[0.0000,0.2 936]	0.1075[0.0000,0 .3790]	0.1625[0.0000,0 .4859]	0.1325[0.0000,0 .4297]	0.17[0.0000,0.4 993]	0.125[0.0000,0. 4149]
450	0.1067[0.0000,0 .3772]	0.1244[0.0000,0 .4138]	0.0778[0.0000,0 .3125]	0.1133[0.0000,0 .3912]	0.1511[0.0000,0 .4650]	0.1356[0.0000,0 .4356]	0.1711[0.0000,0 .5012]	0.12[0.0000,0.4 048]
500	0.108[0.0000,0. 3801]	0.128[0.0000,0. 4208]	0.074[0.0000,0. 3035]	0.12[0.0000,0.4 048]	0.156[0.0000,0. 4741]	0.134[0.0000,0. 4326]	0.168[0.0000,0. 4957]	0.112[0.0000,0. 3884]
<i>logistic regression</i>								

基于全基因组重测序的川滇高山栎和刺叶高山栎生态适应研究

100	0.0473[0.0000,0	0.2014[0.0065,0	0.0052[0.0000,0	0.0022[0.0000,0	0.2176[0.0477,0	0.0556[0.0000,0	0.4467[0.2151,0	0.024[0.0000,0.
0	.1214]	.3963]	.0931]	.0922]	.3874]	.1367]	.6782]	1070]
200	0.0817[0.0181,0	0.2467[0.0808,0	0.0082[0.0000,0	0.0036[0.0000,0	0.1805[0.0735,0	0.0632[0.0000,0	0.3831[0.2205,0	0.033[0.0000,0.
0	.1454]	.4127]	.0819]	.0800]	.2874]	.1346]	.5456]	1045]
300	0.0865[0.0338,0	0.2536[0.1137,0	0.0114[0.0000,0	0.0046[0.0000,0	0.1552[0.0753,0	0.0654[0.0061,0	0.3801[0.2466,0	0.0432[0.0000,0
0	.1392]	.3935]	.0703]	.0667]	.2351]	.1248]	.5136]	.1015]
400	0.0861[0.0410,0	0.2558[0.1343,0	0.0134[0.0000,0	0.0058[0.0000,0	0.1447[0.0786,0	0.0698[0.0180,0	0.3723[0.2578,0	0.052[0.0016,0.
0	.1312]	.3773]	.0636]	.0590]	.2109]	.1216]	.4868]	1024]
500	0.0846[0.0452,0	0.2534[0.1460,0	0.0152[0.0000,0	0.0069[0.0000,0	0.1406[0.0831,0	0.0729[0.0273,0	0.3681[0.2667,0	0.0583[0.0141,0
0	.1240]	.3608]	.0585]	.0531]	.1981]	.1186]	.4695]	.1025]
600	0.0814[0.0468,0	0.2465[0.1510,0	0.0165[0.0000,0	0.0081[0.0000,0	0.1442[0.0919,0	0.0768[0.0362,0	0.3637[0.2725,0	0.0628[0.0239,0
0	.1161]	.3419]	.0541]	.0483]	.1964]	.1173]	.4549]	.1018]
700	0.0831[0.0513,0	0.2387[0.1533,0	0.0176[0.0000,0	0.0092[0.0000,0	0.1473[0.0992,0	0.0808[0.0442,0	0.3563[0.2736,0	0.0672[0.0323,0
0	.1149]	.3240]	.0507]	.0446]	.1953]	.1174]	.4390]	.1021]
800	0.0868[0.0567,0	0.2314[0.1545,0	0.0197[0.0000,0	0.0106[0.0000,0	0.1499[0.1051,0	0.0835[0.0500,0	0.3463[0.2707,0	0.0717[0.0398,0
0	.1169]	.3083]	.0493]	.0424]	.1948]	.1170]	.4219]	.1035]
900	0.0902[0.0613,0	0.226[0.1557,0.	0.0213[0.0000,0	0.0121[0.0000,0	0.1538[0.1113,0	0.0865[0.0551,0	0.3345[0.2650,0	0.0757[0.0459,0
0	.1191]	2963]	.0485]	.0412]	.1964]	.1178]	.4040]	.1054]
100	0.0943[0.0662,0	0.2196[0.1553,0	0.0226[0.0000,0	0.0135[0.0000,0	0.159[0.1181,0.	0.0896[0.0601,0	0.3221[0.2580,0	0.0794[0.0514,0
00	.1224]	.2838]	.0478]	.0405]	1999]	.1192]	.3862]	.1074]

附录 E 基于 2050 年未来气候预测的川滇高山栎和刺叶高山栎种群中 RCP26 和 RCP85 计算的非适应风险结果
 Appendix E. Summary of Risk of non-Adaptedness (RONA) calculated for RCP26 and RCP85 in *Q. aquifolioides* and *Q. spinosa* populations based “Future” climatic predictions for 2050

	RCP26			RCP85		
	bio07	prec03	prec09	bio07	prec03	prec09
<i>Q. aquifolioides</i>						
GBX	0.0175794	4.490766233	0.18004691	0.085826	4.14825091	0.029621636
LZA	0.0433363	6.034157688	0.0158414	0.081697	7.516359396	0.036741334
MLJ	0.0044455	3.944192459	0.0471758	0.099471	7.894464132	0.058430708
LZ	0.0386237	3.459793989	0.08539373	0.028461	6.747190316	0.00762477
LZD	0.0261191	6.255546386	0.0156066	0.004786	4.69905771	0.065950848
BMR	0.0302585	6.073986536	0.06017037	0.132177	5.590134956	0.217521426
BMS	0.0387662	4.564180602	0.00796432	0	4.0879594	0.064664636
CYX	0.008841	4.897831793	0.00796042	0.018818	6.262211312	0
MKD	0.026436	4.932786499	0.0079172	0.028236	6.464963469	0.04370677
MKR	0.0667983	4.968057259	0.05545447	0.004771	5.420128331	0
BZL	0.0088621	4.863357975	0.03164917	0.059539	4.930574864	0.105966639
ZD	0.0220083	4.158860498	0.07600019	0.063866	4.381736507	0.080224622
XCR	0.0437299	2.784212452	0.0873378	0.107486	7.811442378	0.007603005
DCE	0.0220707	2.798765241	0.02384172	0.166381	6.783768928	0.072485024
YJH	0.021986	5.717241878	0	0.109948	6.574173512	0.050974417
DMX	0.0394726	3.610306045	0.06819519	0.093748	5.141295501	0.059196992
JCK	0.0389295	3.027263392	0.06115336	0.051063	3.928929571	0.092112501
PW	0.1224252	7.027388595	0.06122769	0.023729	8.672480565	0.007597251
#SNPs	8393	10957	7963	8393	10957	7963
Min R ²	1.28E-09	3.68E-10	2.79E-09	2.04E-09	9.42E-09	0
Max R ²	0.5707446	0.613266802	0.56163038	0.77439	0.671688363	0.617897559

基于全基因组重测序的川滇高山栎和刺叶高山栎生态适应研究

Average R ²	0.1743	0.0722	0.2048	0.059147	0.09571964	0.049793444
<i>Q. spinosa</i>						
TM	0.035369	8.249626026	0.02518726	0.091579	7.134887037	0.085588362
CYXC	0.0126254	5.693862792	0.00291741	0.077311	5.354273128	0.011630504
BLLK	0.070125	7.301969532	0.10944107	0.201292	5.847088179	0.158181295
XGJQ	0.0038409	4.63355789	0.0421407	0.022675	4.662886722	0.011127864
GZ	0.1199249	5.720823782	0.01468495	0.044393	5.832847153	0.011488182
GDS	0.0749623	5.875653159	0.06114629	0.063784	6.452919323	0.014738916
SWJ	0.0668684	5.612681273	0.01091495	0.130061	6.265217136	0.016845502
LBZ	0.0484105	5.625682	0.00266549	0.085055	6.240277606	0.013631686
JZG	0.0232444	7.250080419	0.00516236	0.106179	8.120728337	0.022843923
ZJH	0.0335883	6.336413096	0.01571156	0.040661	4.825195136	0.055048007
TJS	0.1444078	6.097416778	0.01503784	0.016706	5.845205407	0.003539699
BDX	0.0920556	5.961272524	0.13224873	0.007762	5.889220797	0.00338621
TDY	0.040579	3.985030073	0.02013971	0.150466	7.05936851	0.039015506
DJP	0.1770478	6.412658964	0.00627997	0.05645	7.079088759	0.065951296
WDS	0.0038109	4.652400824	0.07187599	0.015271	4.108482138	0.025546502
TMS	0.1141888	6.230632191	0.01230006	0.042448	5.775794425	0.014903339
WKS	0.1286686	5.458473187	0.01056102	0.002153	5.304108251	0.011169461
SQS	0.0953711	6.711676082	0.03861404	0.058315	5.869714691	0.01030504
#SNPs	17693	14870	8780	17693	14870	8780
Min R ²	0	3.09E-10	1.69E-09	0	5.29E-16	1.41E-34
Max R ²	0.5456589	0.464672051	0.52274316	0.632056	0.582342516	0.589717677
Average R ²	0.0346205	0.034932569	0.0537328	0.041219	0.029393473	0.073493132

附录 F 基于 2070 年未来气候预测的川滇高山栎和刺叶高山栎种群中 RCP26 和 RCP85 计算的非适应风险结果
 Appendix F. Summary of Risk of non-Adaptedness (RONA) calculated for RCP26 and RCP85 in *Q. aquifolioides* and *Q. spinosa* populations based “Future” climatic predictions for 2070

	RCP26			RCP85		
	bio07	prec03	prec09	bio07	prec03	prec09
<i>Q. aquifolioides</i>						
GBX	0.041196	6.863696	1.661431	2.288823	6.449195	0.068063
LZA	0.002922	7.462761	2.099445	2.114445	6.549222	0.047033
MLJ	0	8.920484	1.90904	1.491171	6.518649	0.004238
LZ	0.00875	8.538684	2.467355	1.310415	6.286628	0.004233
LZD	0.023639	8.061599	1.385364	1.746179	5.563925	0.001425
BMR	0.026543	7.941116	1.123125	2.570904	7.378349	0.016471
BMS	0.011647	7.958477	2.495658	2.088579	5.918471	0.028921
CYX	0.090577	6.914189	1.867204	1.848479	6.196375	0.012448
MKD	0.008162	8.896092	1.304225	1.976009	6.282327	0.019474
MKR	0.008635	8.084873	2.062369	2.344172	5.457235	0.033154
BZL	0.01278	8.058924	2.266214	2.723583	7.032386	0.014014
ZD	0.024023	8.789273	1.436755	2.400043	6.152994	0.041267
XCR	0.034075	8.103953	1.204744	1.425753	7.517532	0.00423
DCE	0.020793	7.831577	2.35457	1.554352	7.00099	0.007086
YJH	0.031187	6.891679	2.58902	1.744597	5.855506	0.016723
DMX	0.008514	7.874218	1.843762	1.399129	7.359804	0.012597
JCK	0.005811	8.220438	2.975418	1.949996	7.080524	0.017661
PW	0.018824	8.219535	1.589916	1.298965	6.878407	0.005658
#SNPs	8393	10957	7963	8393	10957	7963
Min R ²	8.83E-07	6.93E-08	8.22E-07	3.51E-09	2.01E-15	8.35E-10
Max R ²	0.328787	0.388585	0.411748	0.576206	0.679609	0.66335

基于全基因组重测序的川滇高山栎和刺叶高山栎生态适应研究

Average R ²	0.074295	0.042871	0.048126	0.055199	0.081581	0.084063
<i>Q. spinosa</i>						
TM	0.329931	8.405032	0.281952	0.238799	7.984965	0.247856
CYXC	0.329955	6.277156	0.342023	0.177352	6.835743	0.244344
BLLK	0.215085	7.36561	0.135858	0.228695	5.550434	0.142739
XGJQ	0.188805	4.528838	0.096034	0.516361	5.954101	0.381899
GZ	0.355977	6.822387	0.314492	0.267983	5.892665	0.407823
GDS	0.285232	5.182611	0.235116	0.181679	3.845858	0.115027
SWJ	0.223344	5.693228	0.174526	0.278741	6.472115	0.449605
LBZ	0.253823	6.029782	0.144555	0.192311	4.336613	0.095064
JZG	0.02412	7.273609	0.140157	0.046584	6.73802	0.026824
ZJH	0.311579	4.722798	0.230367	0.26176	5.974958	0.343788
TJS	0.038408	5.544801	0.336403	0.209939	5.581339	0.166679
BDX	0.313857	6.162258	0.252028	0.279515	5.423217	0.290531
TDY	0.262997	5.966117	0.178665	0.30685	5.858537	0.373289
DJP	0.297476	5.832397	0.255035	0.059273	5.683883	0.06728
WDS	0.196593	3.833818	0.079229	0.287561	5.327621	0.270215
TMS	0.35267	6.365221	0.23325	0.290477	4.970988	0.275742
WKS	0.309846	5.869582	0.263436	0.258901	5.836382	0.24249
SQS	0.297791	6.112138	0.288702	0.282451	5.877399	0.26577
#SNPs	17693	14870	8780	17693	14870	8780
Min R ²	0	8.37E-10	0	0	2.82E-10	1.54E-09
Max R ²	0.617154	0.547192	0.586997	0.665527	0.570983	0.782673
Average R ²	0.058851	0.041855	0.046738	0.063997	0.032066	0.07674

个人简介

刘可可，男，汉族，群众，山东菏泽曹县人。本科毕业于青岛农业大学园林与林学院林学专业，2019年9月-2022年7月就读于北京林业大学生态与自然保护学院，指导教师为杜芳教授，研究方向为修复生态学，

研究生期间选修了分子生态学理论方法、高级森林生态学、生态学研究方法、生态学进展等10门课程，共23学分，主要从事川滇高山栎和刺叶高山栎生态适应研究。在读期间以第一作者发表SCI论文1篇。

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和云南大学生态与环境学院客座教授。

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三项，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23篇，发明专利1项，论文总计被引用865（谷歌学术统计）。2014年在北京林业大学创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；2018年起担任国际栎树协会专业委员会成员和北京生态修复协会理事；2019年当选中国植物学会女科学家协会委员；2021年开始担任SCI期刊Ecology & Evolution的Associate Editor。

获得成果目录清单

[1] Liu K, Qi M and Du FK. Population and Landscape Genetics Provide Insights into Species Conservation of Two Evergreen Oaks in Qinghai-Tibet Plateau and Adjacent Regions[J]. *Frontiers in Plant Science*. 2022, 13:858526.

致谢

转眼间即将与三年的研究生生涯挥手告别，三年的时光，一千多个日日夜夜就这样悄然划过。我从刚入学的欢喜到毕业的不舍，在这所美丽的校园里有惆怅彷徨，也有砥砺前行，在这里学到的必将使我受益终生。海纳百川，取则行远，这一路上我得到过许多人的帮助，离别之际，我想诚挚地感谢所有帮助过我的人们。

很感谢这三年间遇到的每一位老师，特别要感谢的是我的导师杜芳教授，感谢杜老师给我这样一个难得的学习机会，感谢杜老师三年间的敦敦教诲和悉心指导。杜老师是一个务实的人，老师在学习上给予了我很多实质性帮助，在入学之前杜老师就给我强调阅读文献的重要性，并一直要求我们阅读文献的同时做好记录，这种好习惯让我在论文写作的过程中少了很多麻烦。杜老师是一个负责的人，研一做实验时我总是毛手毛脚、粗心大意，各种不规范的操作频频发生，杜老师在一次跑胶实验中及时纠正我的不规范操作，并强调了分子实验安全的重要性，为了保证我实验进度又教我点样，在此感谢杜老师的指导和帮助。杜老师是一个用心的人，老师对我的小论文非常上心，我从 21 年 9 月份才开始小论文的撰写，由于自己英语写作基础差、逻辑不清晰给自己带来不少麻烦，但杜老师从不厌其烦地逐步引导我解决各种问题，帮助我扩展思路，精心点拨、热忱鼓励；从选题、定题开始，一直到最后论文的反复修改、润色，杜老师总会以专业标准严格要求我，使我的论文更加条理化、规范化，最终在 22 年 5 月份顺利发表。小论文发表过程是艰辛的，但庆幸有杜老师的指导和帮助，再次感谢杜老师的付出和包容。杜老师是一个热心的人，老师在工作上也给了我很大帮助，给我推荐了几个非常好的与生态相关的工作，并时常询问和关注我面试的情况，我最终能顺利拿到心仪的 offer 都要感谢杜老师的帮助。杜老师严谨求实、一丝不苟的治学态度和勤勉的工作态度深深感染了我，给了我巨大的启迪和鼓舞，这将成为我一生的财富让我受益终生。

感谢西安交通大学的冯力老师和西北民族大学的王文婷老师在景观阻力分析中的指导和帮助；感谢四川大学的王婧老师在 fastsimocal 分析上的帮助；感谢北京林业大学的尹康权老师在空间遗传多样性分析上的指导和帮助；感谢云南大学赵建立老师、褚雪师姐在梯度森林分析上的指导和帮助。感谢以上所有老师和同学的指导和帮助，谢谢你们！

感谢北京林业大学 402 分子实验室的每一位同学，感谢祁敏师姐在 DNA 提取、PCR 扩增实验上的帮助，感谢师姐在小论文写作上的帮助，帮我逐字逐句修改文章。感谢罗艳君帮我修改毕业论文，对我的毕业论文进行了实质性的帮助；感谢杨霖、康艺馨帮我修改文章；感谢师妹张艺在实验上的帮助；感谢师弟秦伟在数据分析、处理上的帮助；感谢师妹原明星学习和生活上的帮助。感谢已经毕业的师兄

王天瑞、许阳以及师姐李月娟、苏蔚，感谢师兄王天瑞在学习上的指导和帮助，感谢师兄在研一时带我做实验、一步步教我做分析，不厌其烦的解决我的各种问题。感谢许阳师兄在外业和学习上的帮助，师兄在外业时教会我很多辨别栎树的知识，教会了我如何制作外业计划。再次感谢以上提到的各位同学，感谢你们的付出。

感谢我的父母和姐姐，你们是我不断前进的动力，你们是我砥砺前行的保障。感谢你们让我坚定了读研的想法，感谢你们的理解和支持，感谢你们在我做错事时的包容和教导，感谢你们二十多年的付出与陪伴。感谢我的女朋友，感谢你在我心情低落时的耐心开导，感谢你一直以来的默默付出，感谢你七年时光的支持与陪伴，感谢有你！

感谢关心、支持、帮助过我的所有人！行文至此，落笔为终，愿未来以人为镜，不断追求，不断进取。