

分类号

密级

UDC



北京林业大学

学术型硕士学位论文

川滇高山栎 *QaGDU3* 基因生态功能研究

Study on the ecological function of *QaGDU3* gene in an alpine oak
(*Quercus aquifolioides* Rehd. Et Wils.)

杨霖

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 二十日

独创性声明

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

签 名 栾霖

日 期： 2022.06.09

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： 栾霖 导师签名： 栾霖 日期： 2022.06.09

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

川滇高山栎 (*Quercus aquifolioides* Rehd. et Wils.) 为壳斗科 (Fagaceae) 栎属 (*Quercus* L.) 冬青栎组 (Section *Ilex*) 常绿树木, 广泛分布于中国西南地区横断山和喜马拉雅山的生物多样性热点地区。川滇高山栎地理分布广泛, 耐受极端环境, 是研究栎属植物物种形成和高山环境适应的理想材料。

基于 Illumina、Nanopore 和 Hi-C 数据获得了川滇高山栎高质量染色体水平的参考基因组, 组装得到的基因组大小约为 1.15Gb, 总共产生了 1 445 个 contigs, contig N50 约为 1.20Mb。96.02% (1 101.96Mb) 的组装基因组序列能够固定在 12 条染色体上, 是栎属植物中组装至染色体水平上较高质量的参考基因组。基因组注释显示川滇高山栎基因组重复序列占 58.04%, 其中长末端片段重复序列 (Long Terminal Repeats, LTRs) 是含量最丰富的重复类型, 占整个基因组的 40.86%。基因组中有 26 441 个蛋白质编码基因被预测, 其中 92.19% 的蛋白编码基因能够被功能注释。系统发育分析表明川滇高山栎与近缘物种欧洲栓皮栎 (*Quercus suber* L.)、加州白栎 (*Quercus lobata* Née) 和有柄栎 (*Quercus robur* L.) 聚为一枝。正选择检验和功能注释表明川滇高山栎基因组中 416 个正选择基因 (Positively Selected Genes, PSGs) 与“RNA 和受损 DNA 结合”、“光系统 II 氧复合物”和“核苷酸切除修复”等功能相关, 这些基因对川滇高山栎适应高山环境起着重要作用。

为了探索川滇高山栎对高山环境的适应机理, 通过正选择分析筛选出与“质子跨膜转运”功能相关的川滇高山栎正选择 *QaGDU3* 基因 (*Q. aquifolioides* *GLUTAMINE DUMPER 3*) 进行进一步分子功能研究。采用不依赖于连接反应的克隆 (Ligation-Independent Cloning, LIC) 和农杆菌花序侵染等方法获得 *QaGDU3* 转基因拟南芥 (*Arabidopsis thaliana* (L.) Heynh.)。利用生物信息学分析 *QaGDU3* 蛋白的结构特征和与其它物种 *GDU3* 蛋白的序列差异, 结合系统发育分析探究了 *GDU3* 在 *GDU* 家族中及 *QaGDU3* 与其它物种 *GDU3* 蛋白的进化关系。结果表明 *QaGDU3* 为酸性不稳定亲水跨膜蛋白, 无信号肽, 可能定位于细胞质膜上; 其二级结构主要由 α -螺旋和无规卷曲构成。*QaGDU3* 蛋白含有一个典型的“VIMAG”保守结构域, 属于阴离子渗透酶 *ArsB/NhaD* 超家族成员。系统进化分析显示 *GDU3* 基因只在种子植物中发现, 且 *GDU3* 在 *GDU* 家族中进化速率较快, *QaGDU3* 与 *QsGDU3* 亲缘关系最

近。通过实时荧光定量 PCR (quantificational Real-Time Polymerase Chain Reaction, qRT-PCR) 和图像分析等方法验证了转基因 *QaGDU3* 拟南芥 T2 代株系莲座大小随着 *QaGDU3* 基因表达量升高而变小, 且转基因植物 *QaGDU3* 基因的过表达能够激活水杨酸 (Salicylic Acid, SA) 途径相关基因的表达。*QaGDU3* 转基因拟南芥对茄病镰刀菌 (*Fusarium solani*) 具有抗性。

本研究通过构建川滇高山栎全基因组图谱, 为理解川滇高山栎的物种形成和适应性进化提供了遗传基础。进一步对正选择 *QaGDU3* 基因进行挖掘及功能鉴定, 探究了 *QaGDU3* 基因的分子功能, 也为探究川滇高山栎正选择基因功能在适应性进化中的作用提供了新信息。

关键词: 川滇高山栎, 适应性进化, *GDU3*, 基因克隆, 遗传转化

Study on the ecological function of *QaGDU3* gene in an alpine oak
(*Quercus aquifolioides* Rehd. Et Wils.)

Master Candidate: Lin Yang

(Ecology)

Directed by Prof. Dr. Fang K. Du

Abstract

The alpine oak (*Quercus aquifolioides* Rehd. et Wils.), belongs to section *Ilex* in Fagaceae, is an evergreen oak species, widely distributed across the Hengduan Mountains biodiversity hotspot in the Eastern Himalayas. The wide geographical distribution and ecological amplitude make it an ideal model for speciation and adaptation study in oak species.

This thesis presented a high-quality chromosome-level genome sequence of *Q. aquifolioides*, based on Illumina, Nanopore, and Hi-C data. De novo assembly of these long reads produced a contiguous assembly of 1 445 contigs with an N50 of ~1.20Mb and a cumulative size of ~1.15Gb. Hi-C scaffolding anchored 1 101.96Mb contigs (96.02% of draft assembly) onto 12 pseudo-chromosomes. Genome annotation revealed that 58.04% of the *Q. aquifolioides* genome composed of repeated sequences, Long Terminal Repeats (LTRs) constituted the predominant type of repeat elements in the genome and accounted for approximately 40.86%. 26 441 protein-coding genes were predicted and 92.19% of them were annotated. Phylogenetic analysis showed that four oaks species (*Q. aquifolioides*, *Q. suber* L., *Q. lobata* Née and *Q. robur* L.) are clustered together. Positive selection analysis suggest 416 genes were Positively Selected The functional annotations shows above Positively Selected Genes (PSGs) are potentially contributed to miscellaneous traits of adaptive importance, such as RNA and damaged DNA binding, photosystem II oxygen evolving complex and nucleotide-excision repair.

In order to explore the adaptation mechanism of *Q. aquifolioides* to the alpine environment, we screened the “proton transmembrane transport” positive selection *QaGDU3* (*Q. aquifolioides* *GLUTAMINE DUMPER 3*) gene for further molecular function. We used Ligation-Independent Cloning (LIC) and floral dipping transformation to clone it and obtain transgenic *Arabidopsis thaliana* (L.) Heynh. Then we investigated its structure, sequence difference and evolutionary relationship with other related

species. Results showed that the QaGDU3 is an acid and unstable hydrophilic transmembrane protein without signal peptide, predicted to be located in the cytoplasmic membrane. α -helix and random coil are the main structural elements. QaGDU3 has a typical "VIMAG" conserved domain, belongs to anion permease ArsB/NhaD superfamily. Phylogenetic analysis showed that *GDU3* is probably present only in seed plants. *GDU3* evolves faster in the GDU family, and the QaGDU3 is closely related to QsGDU3. The quantificational Real-Time Polymerase Chain Reaction (qRT-PCR) and image analysis showed that the rosette size of transgenic *A. thaliana* decreased with the increase of *QaGDU3* gene expression and the overexpression of the *QaGDU3* gene in transgenic plants activated the expression of genes involved in the Salicylic Acid (SA) pathway. *QaGDU3* transgenic *A. thaliana* exhibited resistance to *Fusarium solani*.

This study provides a genetic basis for understanding the speciation and adaptive evolution of *Q. aquifolioides* at the genome-wide level. The mining and functional identification of the positive selection *QaGDU3* gene lays a foundation for exploring the molecular function of *QaGDU3* gene and also provides new information for exploring the function of positive selection genes of *Q. aquifolioides* in adaptive evolution.

Key words: *Quercus aquifolioides*, genome, adaptive evolution, *GDU3*, gene cloning, genetic transformation

目录

1 引言.....	1
1.1 栎属植物基因组研究进展.....	1
1.1.1 已公布的栎属植物基因组.....	1
1.1.2 栎属植物基因组重要功能.....	3
1.2 <i>GDU3</i> 基因研究进展.....	5
1.2.1 <i>GDU3</i> 研究现状.....	5
1.2.2 <i>GDU3</i> 与植物的抗性.....	5
1.3 镰刀菌对栎属植物的影响研究.....	7
1.4 川滇高山栎的研究概况.....	7
1.4.1 川滇高山栎物种介绍.....	7
1.4.2 川滇高山栎研究进展.....	8
1.5 技术路线.....	8
1.6 本文的研究内容、研究目的及意义.....	9
1.6.1 研究内容.....	9
1.6.2 研究目的和意义.....	9
2 高质量高山栎基因组组装与注释.....	11
2.1 基因组资源的获取.....	11
2.1.1 样品的采集与处理.....	11
2.1.2 实验方法.....	11
2.2 适应性进化及正选择分析.....	12
2.2.1 其它 13 个物种基因组数据获取.....	12
2.2.2 系统发育分析与分化时间估计.....	12
2.2.3 正选择基因鉴定及功能分析.....	13
2.3 研究结果.....	13
2.3.1 基因组结果.....	13
2.3.2 系统发育及分化时间估计.....	20
2.3.3 正选择基因及其功能.....	20
3 川滇高山栎 <i>QaGDU3</i> 基因克隆及其功能分析.....	23
3.1 试验材料.....	23
3.1.1 植物材料.....	23
3.1.2 试验试剂.....	23
3.1.3 实验试剂配制.....	24

目录

3.2 研究内容与方法.....	24
3.2.1 <i>QaGDU3</i> 基因的克隆及表达载体构建	24
3.2.2 <i>QaGDU3</i> 基因的生物信息学分析	28
3.2.3 <i>QaGDU3</i> 基因的系统发育分析	28
3.2.4 <i>QaGDU3</i> 基因遗传转化拟南芥与鉴定	29
3.2.5 荧光定量分析.....	30
3.2.6 <i>QaGDU3</i> 转基因拟南芥植株莲座大小测定	32
3.2.7 <i>QaGDU3</i> 转基因拟南芥对镰刀菌根腐病抗性研究	32
3.3 研究结果.....	33
3.3.1 <i>QaGDU3</i> 基因的克隆及序列分析结果	33
3.3.2 <i>QaGDU3</i> 基因的生物信息学分析结果	34
3.3.3 <i>QaGDU3</i> 基因家族的进化结果	37
3.3.4 <i>QaGDU3</i> 遗传转化拟南芥鉴定	41
3.3.5 <i>QaGDU3</i> 基因表达量与莲座大小关系	41
3.3.6 水杨酸途径中相关基因表达量测定	43
3.3.7 <i>QaGDU3</i> 基因抗镰刀菌根腐病研究结果	44
4. 讨论	48
4.1 川滇高山栎基因组特性.....	48
4.2 川滇高山栎对高山环境的适应性进化.....	49
4.3 <i>QaGDU3</i> 基因特征	49
4.4 <i>QaGDU3</i> 基因对川滇高山栎适应性影响	51
5 结论与展望	53
5.1 结论.....	53
5.2 展望.....	53
参考文献	55
附录	66
个人简介	68
导师简介	69
成果目录清单	70
致谢	71

1 引言

1.1 栎属植物基因组研究进展

DNA 是生物遗传信息的载体，其序列信息蕴含了生物体中各种复杂的生命过程，利用 DNA 测序技术可以对生物体的全基因组序列进行研究。采用第一代 Sanger 测序技术完成的第一个模式植物拟南芥 (*Arabidopsis thaliana* (L.) Heynh.) 的全基因组测序，开启了植物全基因组测序研究的时代 (Initiative 2000)。伴随高通量测序技术的发展，已测序的植物基因组的数量一直在稳步增加。木本植物是森林生态系统重要的组成部分，由于木本植物生命周期较长、基因组较大、杂合度较高，且多数发生过基因组的复制，这些瓶颈限制了林木全基因组测序研究的进程。杨属 (*Populus* L.) 植物是木本植物中基因组较小的物种，TusKana 等 (2006) 完成了第一个木本植物毛果杨 (*Populus trichocarpa* Torr. & Gray) 的全基因组测序，成为了木本植物基因组及遗传转化研究的模式物种。

栎属 (*Quercus* L.) 是壳斗科 (Fagaceae) 中数量最丰富的属，广泛分布在北美、欧洲和东亚地区 (Willis *et al.*, 1968; Manos *et al.*, 1999; Kremer *et al.*, 2012; Hubert *et al.*, 2014; Cavender, 2016; Denk *et al.*, 2018)。栎属植物对环境具有广泛的适应性 (Xu *et al.*, 2019)，能够为物种局地适应性进化机制的研究提供参考 (Petit *et al.*, 2013; Kremer, 2016)。高质量的参考基因组资源能够为探究栎属植物的起源及历史演化过程、生长发育背后的分子进化提供重要的遗传信息；同时对于鉴定和分析栎属植物适应性进化相关基因功能和调控机制提供了方向。

1.1.1 已公布的栎属植物基因组

高通量测序技术的发展促进了越来越多栎属植物的基因组得到测序。为了探索进化在栎属植物基因组水平上的印迹，法国科学家提出了“栎属植物基因组计划”并完成了第一个栎属植物—有柄栎 (*Quercus robur* L.) 的基因组测序 (Plomion *et al.*, 2018)，为栎属植物基因组研究提供了参考。除此之外，落叶栎属植物加州白栎 (*Quercus lobata* Née) (Sork *et al.*, 2016)、蒙古栎 (*Quercus mongolica* Fisch. ex Ledeb.) (Ai *et al.*, 2020) 及常绿栎属植物欧洲栓皮栎 (*Quercus suber* L.) (Ramos *et al.*, 2018) 的基因组也被测序并释放。基于当前已有数据，我们选择了有柄栎、加州白栎和欧洲栓皮栎的基因组与川滇高山栎的基因组进行了后续的比较分析。(表 1.1)。

(1) 有柄栎基因组研究进展

有柄栎 (*Quercus robur* L.) 是白栎组 (Sect. *Quercus*) 一种大型的落叶乔木，其

杂合度及重复较高，因此其基因组数据的获取具有巨大挑战，同时利用基因组数据探究有柄栎长寿的原因也具有重要意义。Lesur 等（2011）对一棵树龄约为 100 年的有柄栎进行了全基因组鸟枪法（Whole-Genome Shotgun, WGS）测序，初步预估其基因大小约为 740.00Mb/C。Plomion 等（2016）进一步通过 Illumina Hiseq 2000、Roche 454/GS-FLX、ABI-capillary 平台对有柄栎进行基因组测序，产生了 17910 个 scaffolds (>2kb)，总长度达 1.34Gb，Scaffold N50 (the scaffold size above which 50% of the total length of the sequence assembly can be found) 达 260.00kb。进一步组装得到了包含 1409 个 scaffolds，大小约为 814.28Mb，N50 约为 1.34Mb 的高质量单倍体基因组，其中 96.00% (716.60Mb) 的基因组序列被定位在 12 条染色体上。52.00% 的有柄栎基因组包含多种转座元件 (Transposable Elements, TEs)，其中 70.00% 由 I 类反转录转座子 (class I retrotransposons) 构成。在有柄栎基因组中预测得到 25808 个蛋白编码基因，较高的核苷酸多样性 (nucleotide diversity) 和较高比率的缺失突变 (deleterious mutations) 是有柄栎基因组重要的特征。

在有柄栎基因组中鉴定到 44 个可靠的体细胞突变 (somatic mutations) 位于 12 条染色体上，而且这种突变可从体细胞传递到生殖细胞并传递给后代。进一步研究发现，有柄栎从 21 条祖先型染色体 (ancestral eudicot karyotype) (Salse, 2016) 经过 5 次断裂 (fissions) 和 14 次融合 (fusions) 后形成了目前的 12 条染色体。且在有柄栎-桃树谱系分化后，有柄栎经历了局部基因复制 (占有柄栎基因组的 35.60%)。在有柄栎扩张的基因中 73.00% 属于串联重复基因 (Tandemly Duplicated Genes, TDGs)，而 TDGs 和谱系特异性选择的这种紧密关系对于抗病 (R) 基因尤为常见 (Hanada *et al.*, 2008; Zhang *et al.*, 2016)。从而验证了树木中可能存在一种趋同机制，使长寿物种中的这些基因家族得以扩张。

(2) 加州白栎基因组研究进展

加州白栎 (*Quercus lobata* Née) 是白栎组 (Sect. *Quercus*) 植物，在稳定土壤、维持水生和陆生动物栖息地和改善水质方面具有重要生态作用 (Dosskey *et al.*, 1997; Aitken and Whitlock, 2013)。加州白栎与有柄栎都属于白栎组，但它们有不同的进化史。有柄栎在更新世冰川期间 (Pleistocene glaciations) 经历了严重的瓶颈效应，大多数种群存活在欧洲南部避难所 (Petit *et al.*, 2002; Petit *et al.*, 2003)，而加州白栎至少在近 15 万年内没有经历严重的瓶颈效应，在最近几个间冰期内种群分布相对稳定 (Gugger *et al.*, 2013)。

Sork 等 (2016) 基于 Illumina HiSeq 2500 平台对加州白栎进行了基因组测序，基于 k-mer 估计显示其二倍体基因组大小约为 720.00~730.00Mb，杂合度约为 1.25%。最终组装得到加州白栎基因组 (v 1.0) 包含 94394 个 scaffolds，总长约为 1.17Gb，N50 约为 278.00kb。进一步基于 Pacific Biosciences 长序列结合 Hi-C 技术将基因组序

列组装至染色体水平 (v 3.0)，基因组大小约为 846.07Mb，N50 约为 66.42Mb (Sork *et al.*, 2021)。加州白栎与有柄栎和欧洲栓皮栎的核苷酸序列相似度分别为 96.00%和 93.00%。在加州白栎基因组中共有 61 773 个蛋白编码基因被注释，其中 39 373 个包含了与遗传和表型多样性相关的重复，这些重复包括通过古老的 γ 全基因组复制事件 (Whole-Genome Duplication events, WGD) 中保留的重复基因以及通过与抗性相关基因家族内串联复制而重复的基因，这些特征促进了加州白栎遗传和表型的变异，从而有助于其适应不断变化的环境。

(3) 欧洲栓皮栎基因组研究进展

欧洲栓皮栎 (*Quercus suber* L.) 是麻栎组 (Sect. *Cerris*) 的常绿乔木，其软木层具有可再生的罕见特性，对本地动植物物种的生存和防止脆弱地区的荒漠化具有重要的作用。但由于农业化、生物胁迫、火灾和气候变化等原因的影响，欧洲栓皮栎的种群数量逐渐呈现下降趋势。

Ramos 等 (2018) 通过筛选 28 棵自然生长的欧洲栓皮栎中纯合度最高的 HL8 个体进行 Illumina 基因测序。基于 k-mer 估计的基因组大小约为 1.06Gb，杂合度约为 1.62%。最终组装得到的欧洲栓皮栎基因组大小约为 953.30Mb (N50 约 465.16kb)，包含 23 344 个 scaffolds (>1kb)，基因组注释得到了 79 752 个基因。在所有类型的重复序列中反转录转座子含量最高，占欧洲栓皮栎基因组的 7.59%。

(4) 蒙古栎基因组研究进展

蒙古栎 (*Quercus mongolica* Fisch. ex Ledeb.) 是白栎组 (Sect. *Quercus*) 的落叶乔木，对害虫、疾病、寒冷和干旱具有高度抗性 (Hao *et al.*, 2017)。Ai 等 (2020) 对蒙古栎进行了基因组测序。基于 Illumina HiSeq X Ten 测序数据进行 k-mer 估计的基因组大小约为 842.80Mb，杂合度约为 1.09%。结合 PacBio Sequel II 测序平台的数据最终组装得到的蒙古栎基因组大小约为 809.83Mb (N50 约为 2.45Mb)。进一步利用 Hi-C 技术将 774.59Mb (95.65%) 基因组序列组装至 12 条假染色体上，其 scaffold N50 约为 66.74Mb。基因组注释得到 36 553 个蛋白编码基因，平均长约 6 084.00bp。蒙古栎基因组中重复序列含量占 68.46% (554.52Mb)，其中 44.07% 为长末端重复序列 (Long Terminal Repeats, LTRs)。比较基因组学分析表明蒙古栎与有柄栎亲缘关系最近，两者大约在 10.20 个百万年前 (Million years ago, Mya) 发生分化。蒙古栎基因组中经历了两次早于有柄栎的全基因组加倍事件，同时在蒙古栎基因组中鉴定得到 38 个正选择基因 (Positively Selected Genes, PSGs) 和 496 个与木材形成有关的基因。

1.1.2 栎属植物基因组重要功能

通过对有柄栎扩张的基因功能通过基因本体论 (GO, Gene Ontology) 数据库富

集分析发现这些基因主要与植物体内的生物交互作用 (biotic interactions) 相关。进一步对有柄栎特有的 126 个扩张基因功能进行 GO 富集分析发现 63.00% 的基因与植物免疫相关, 其中参与了对细菌或卵菌病原体的感知的相关基因可能通过增强有柄栎的抗性来影响其寿命。

加州白栎的串联重复基因具有许多可识别的 Pfam 功能域, 其中包括与植物抗性和防御相关的 NBS-ARC、LRR、TIR; 与信号转导相关的 Pkinase_Tyr; 葡萄糖糖基转移酶家族的 UDPGT; 与蛋白质-蛋白质相互作用、降解、信号转导和调节相关的 F-box 和与应激、耐受和解毒功能相关的谷胱甘肽 S-转移酶 GST_N。其中, 与植物抗病相关的基因 (R) 显著富集, 这些基因在植物防御机制中发挥重要的作用 (Gururani *et al.*, 2012)。

在蒙古栎 38 个正选择基因中, 有 34 个基因能够在 Swissprot 数据库进行功能注释。其中, 鉴定得到了编码生物多糖生物合成途径中关键酶的 *LpxB* 和 *LpxC5* 基因 (Li *et al.*, 2011), 它们可能参与调节植物对病原体的免疫反应 (Newman *et al.*, 1997; Dow *et al.*, 2000; Zeidler *et al.*, 2004; Shang-Guan *et al.*, 2018)。 *Pyridoxal reductase1* (*PLR1*) 基因是编码维生素 B6 代谢途径中的一种关键酶, 在抵抗渗透胁迫中起关键作用 (Herrero *et al.*, 2011); *Switch subunit 3* (*SWI3*) 为编码脱落酸 (ABA) 信号阳性调节因子的基因, 能够参与抵抗非生物胁迫 (Saez *et al.*, 2008)。这些正选择基因可能反映了蒙古栎对疾病、寒冷和干旱的抵抗力。

表 1.1 栎属植物基因组数据

Table 1.1 Oaks species genome data statistics

物种	基因组大小(Mb)	Scaffolds 数量	Scaffold N50 长度(base pair)	Contigs 数量	Contig N50 长度(base pair)	网址
有柄栎 <i>Quercus robur</i>	814.28	1 409	1 342 530	22 615	69 349	https://www.oakgenome.fr/
加州白栎 <i>Quercus lobata</i>	846.07	2 010	66 423 012	3 684	965 935	https://valleyoak.ucla.edu/
欧洲栓皮栎 <i>Quercus suber</i>	953.30	23 344	465 160	36 760	80 917	https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/002/906/115/GCF_002906115.1_CorkOak1.0/
蒙古栎 <i>Quercus mongolica</i>	809.83	321	66 735 633	735	2 446 788	https://www.binfire.authorea.com/doi/full/10.22541/au.160216757.73285629/v1
川滇高山栎 <i>Quercus aquifolioides</i>	1 147.54	-	-	1 445	1 202 933	http://www.oakofchina.org/

1.2 *GDU3* 基因研究进展

1.2.1 *GDU3* 研究现状

跨膜蛋白是一类结构独特，在植物细胞中广泛存在，并发挥重要生理功能的蛋白质。由于跨膜蛋白穿过膜的磷脂双层，因此其就成为了膜内外物质交换、能量和信号的传递的“桥梁”。其中参与氨基酸转运的跨膜蛋白是植物代谢过程中关键的调控因子（Tegeder *et al.*, 2000; Wipf *et al.*, 2002; Liu and Bush, 2006; Zhu *et al.*, 2011; Zhao *et al.*, 2012; Tegeder, 2012; Peng *et al.*, 2014），对植物的生长和发育起十分重要的作用。随着基因组测序技术的发展，越来越多的氨基酸转运跨膜蛋白基因在模式植物拟南芥中被分离、克隆及功能分析。其中 *GDU* 蛋白（Glutamine Dumper Protein, *GDU*）为谷氨酰胺（Gln）转运蛋白质，其包含一个功能未知的小膜蛋白，可能是参与调节氨基酸代谢的氨基酸转运蛋白的亚基。研究表明 *GDU* 基因家族为植物特有，其蛋白质结构存在两个保守结构域：一个假定的跨膜结构域和一个特有的高度保守的“VIMAG”疏水结构域，它可与位于膜上的蛋白质相互作用（Pratelli and Pilot, 2006），并参与氨基酸运输的各种途径（Malhotra *et al.*, 1993）。在“VIMAG”结构域中可以出现一种氨基酸由 G 到 R 的替换，这种替换不会改变“VIMAG”结构域的二级结构，表明该蛋白质的结构比较保守。

Pilot 等（2004）发现了一种拟南芥激活标签突变体 *gdu1-1D*，它可在水孔处积累结晶，化学分析表明该结晶主要含有谷氨酰胺。通过克隆 *AtGDU1* 基因，发现其编码一种新的包含一个假定跨膜结构的 17-kD 蛋白。*AtGDU1* 基因在植物的维管组织和水孔中表达，导致木质部汁液和叶片质外体中谷氨酰胺含量增加。同时观察发现转基因 *AtGDU1* 突变体植株比同时期生长的野生型植株小，与野生型植株相比，突变型植株叶片卷曲，呈深绿色。进一步对转基因 *AtGDU1* 突变体和野生型植株的大小（即莲座的直径）进行统计分析比较，结果显示转基因 *AtGDU1* 突变体植株叶片大小和 *AtGDU1* 蛋白积累之间存在明显的相关性，随着转基因植株 *AtGDU1* 蛋白积累增加，其叶片变小。

Pratelli 等（2010）进一步在拟南芥中鉴定到 7 个谷氨酰胺转运蛋白，编码这些蛋白的基因的过表达能够刺激氨基酸外向转运，且不同类型的 *AtGDU* 基因在根、维管组织和生殖器官中发挥不同的作用，例如 *AtGDU5* 在保卫细胞中增加编码 β -葡萄糖苷酸酶的 *Gus* 基因（ β -glucuronidase, *Gus*）的活性；*AtGDU3* 的表达在花药中被激活；*AtGDU4* 和 *AtGDU7* 在叶脉中发挥作用。

1.2.2 *GDU3* 与植物的抗性

致病机理相关的基因（Pathogenesis-Related genes, *PRs*）能够参与植物体过敏性

坏死反应 (Hypersensitive Response, HR), 使植物建立系统获得性抗性 (Systemic Acquired Resistance, SAR), 从而影响植物的防御反应。因此, 植物 SAR 的建立通常由 PRs 的诱导表达引起 (Ryals *et al.*, 1996; Edreva and Kostoff, 2004)。水杨酸 (Salicylic Acid, SA) 是一种重要的植物激素, 能够诱导 SAR, 当植物的某部位在遭受病原体侵染后, 植物体内与致病机理有关基因的表达增强, 导致植物体内 SA 的含量增加, 并参与植物防御反应 (Uknes *et al.*, 1992; Schenk *et al.*, 2000)。

植物体内至少有三种类型的 SA 调节因子, 分别为: I 型、II 型和 III 型调节因子 (Lu *et al.*, 2009)。I 型调节因子包括参与 SA 生物合成的酶, 例如: SA Induction-Deficient 2 (SID2) 通过异分支酸合酶 (Isochorismate Synthase, ICS) 使分支酸异构化为异分支酸, 再经过异分支酸丙酮酸裂解酶 (Isochorismate Pyruvate Lyase, IPL) 合成 SA (Wildermuth *et al.*, 2001); Agd2-Like Defense Response Protein 1 (ALD1) 是一种转氨酶, 能与体外多种底物相结合, 参与 SA 生物合成 (Song *et al.*, 2004); Wall-Associated Kinase 1 (WAK1) 是一种细胞质丝氨酸/苏氨酸激酶, 通过病原菌侵染或 SA 及其类似物 INA 诱导, 参与植物的防御反应 (He *et al.*, 1999)。II 型调节因子能够影响 SA 的积累, 但可能不是生物合成酶, 例如: Accelerated Cell Death 6 (ACD6) (Lu *et al.*, 2003; Dong, 2004; Lu *et al.*, 2005)、Enhanced Disease Susceptibility 1 (EDS1) (Falk *et al.*, 1999)、Phytoalexin Deficient 4 (PAD4) (Jirage *et al.*, 1999)、SID1/EDS5 (Nawrath *et al.*, 2002)、Flavin-Dependent Monooxygenase 1 (FMO1) (Mishina and Zeier, 2006)、Cam-Binding Protein 60-Like g (CBP60g) (Wan *et al.*, 2012) 和 SAR Deficient 1 (SARD1) (Sun *et al.*, 2018), 它们可能影响 SA 前体的可用性、SA 生物合成相关酶的活性和 SA 分解代谢。III 型调节因子在 SA 下游传递防御信号, 例如: Pathogen And Circadian Controlled 1 (PCC1) 可以通过 SA 依赖过程调节胁迫植物的开花 (Segarra *et al.*, 2010; Sauerbrunn and Schlaich, 2004); Nonexpressor of Pathogenesis-Related Genes 1 (NPR1) 对 SA 介导的防御信号至关重要, 同时也有独立的 NPR1 参与 SA 信号通路, 但尚未被分子鉴定 (He *et al.*, 1998)。这些调节因子能够独立或共同作用调控植物体内 SA 含量, 从而参与受病原菌侵染后植物体内的一系列防卫反应。

双生病毒 (Geminivirus) 是一组具有双生颗粒形态的单链环状植物 DNA 病毒 (Zerbini *et al.*, 2017), Chen 等 (2010) 通过对甜菜曲顶卷叶病毒 (Beet Severe Curly Top Virus, BSCTV) 这种双生病毒抗性突变体的遗传筛选, 发现 *AtGDU3* 基因能够参与双生病毒与宿主的相互作用, 它的过度表达能够激活 SA 途径中 *ACD6*、*PCC1*、*WAK1* 和 *PRI* 四个基因的表达, 从而减弱甜菜曲顶卷叶病毒 DNA 的复制, 降低植物对 BSCTV 的易感性。

目前关于 *GDU* 基因的研究还仅局限于模式植物拟南芥中。鉴于 *AtGDU3* 基因的超量表达能够增强转基因拟南芥的抗病性, 因此探究壳斗科栎属植物中 *GDU3* 基因

的作用和 *QaGDU3* 基因与川滇高山栎适应高山环境间的关系具有重要的意义。

1.3 镰刀菌对栎属植物的影响研究

疾病是导致栎属植物数量减少和死亡的重要原因之一 (Ragazzi *et al.*, 1995; Moreira and Martins, 2005; Linaldeddu *et al.*, 2011; Akilli *et al.*, 2013)。真菌作为导致疾病的重要因素, 能够使栎属植物树冠枯死或根系受损, 影响树木健康。镰刀菌 (*Fusarium spp.*) 是土壤中普遍存在的重要的真菌, 它不仅是引起苜蓿、大豆等豆科饲料作物根腐病病害的主要病原菌 (Chi *et al.*, 1964; Leath, 1978; 郭玉霞等, 2009; Samac *et al.*, 2016; 方香玲等, 2019;), 其产生的毒素对栎属植物的组织细胞也具有明显毒害作用。

Alidadi 等 (2019) 从伊朗健康和衰退的波斯栎 (*Quercus brantii* Lindl.) 幼苗的枝条中分离出部分真菌, 根据形态学和分子鉴定真菌种类, 首次发现了两种与波斯栎相关的病原菌: 黑斑病的致病菌 (*Didymella glomerata*) 和新暗色柱节孢菌 (*Neoscytalidium dimidiatum*); 此外, 首次发现枝顶孢霉属菌种 (*Acremonium sp.*)、子囊菌锥毛壳属菌种 (*Coniochaeta sp.*)、壳孢属菌种 (*Cytospora ribis*)、三线镰刀菌 (*Fusarium tricinctum*)、茄病镰刀菌 (*F. solani*) 和粉红单端孢菌 (*Trichothecium roseum*) 为波斯栎内生真菌。Vujanovic 等 (1999) 研究验证了茄病镰刀菌能够导致北美红栎 (*Quercus rubra* L.)、纳氏栎 (*Quercus nuttallii* E.J.Palmer)、柳叶栎 (*Quercus phellos* L.) 和黑栎 (*Quercus nigra* L.) 树叶的枯萎和树干树枝一年生溃疡病。Lynch 等 (2013) 进一步在加利福尼亚西海岸的加州槲栎 (*Quercus agrifolia* L.) 溃疡的树枝和树干样本的坏死组织中分离了部分真菌, 结合形态学和系统发育分析, 鉴定得到了茄病镰刀菌 (*Fusarium solani*)、褐顶孢菌 (*Phaeoacremonium mortonia*)、疣凸假蕉孢壳菌 (*Diatrypella verruciformis*) 等, 并对一年生的加州槲栎幼苗伤口进行接种, 发现每个真菌都导致了加州槲栎木质部不同程度的坏死, 其中茄病镰刀菌的致病性最强。

1.4 川滇高山栎的研究概况

1.4.1 川滇高山栎物种介绍

(1) 形态特征

川滇高山栎 (*Quercus aquifolioides* Rehd. et Wils.) 为壳斗科 (Fagaceae) 栎属 (*Quercus* L.) 冬青栎组 (Section *Ilex*) 常绿乔木或灌木, 生于干旱阳坡或山顶 (端木炳, 1994; 王国严和徐阿生, 2008; Denk *et al.*, 2018)。根据中国植物志记载: 川滇高山栎叶片整体呈椭圆形或倒卵形, 叶片全缘或有刺状锯齿, 幼叶两面有黄棕色腺

毛，老叶背面有黄棕色单毛或粉状鳞秕。叶柄长 2-5 毫米，有时近无柄；雄花序长 5-9 厘米，花序轴及花被均被疏毛；果序长 0.5-2.5 厘米，有花 1-4 朵。坚果基部被浅杯形壳斗包裹，壳斗内壁有密绒毛；坚果呈卵形或长卵形，无毛；花期 5-6 月，果期 9-10 月（陈焕镛和黄成就, 1998）。

(2) 生长习性 & 地理分布

川滇高山栎是硬叶常绿阔叶林的主要物种，广泛分布于中国西南地区横断山和喜马拉雅山的生物多样性热点地区，海拔分布范围在 1 700m-4 600m，为该地区的建群树种（Feng *et al.*, 2016; Meng *et al.*, 2017）。

(3) 主要价值

川滇高山栎根系发达，萌生能力强，因此对金沙江、雅鲁藏布江等江河源头及沿河地区的水土保持和水源涵养有十分重要的作用（端木炘, 1994; 黄亚洲, 2011）。川滇高山栎在遭受火灾、砍伐等干扰后有较强的再生能力，是川西、滇西地区高山栎灌丛的主要更新方式，对维持川、滇高海拔地区及横断山脉一些脆弱生态系统的稳定性具有十分重要的生态功能（杨玉坡等, 1992; 谢静等, 2014; 王超等, 2021）。同时，川滇高山栎树皮含单宁木质坚硬，被人们用作薪柴、机舱版、木钉等。川滇高山栎也是著名食用菌松茸的伴生树种，其坚果具有一定的食用、药用和较高的经济价值（王国严等, 2011; 程晓彬等, 2014）。

1.4.2 川滇高山栎研究进展

近年来，川滇高山栎的研究重点集中在系统演化、生物地理等方面（王玉垚, 2018; 王天瑞, 2020）。Meng 等（2017）使用叶绿体（chloroplast DNA, cpDNA）片段和核微卫星（nuclear Simple Sequence Repeats, nSSR）分子标记，通过种群遗传学分析表明为适应青藏高原的隆升过程中的环境变化，高山栎组植物形成了冷暖区分布的两个谱系，其中川滇高山栎分布于冷区域。Du 等（2017; 2020）进一步基于叶绿体片段、核微卫星和候选基因数据在川滇高山栎物种分布范围内将其分为了西藏（Tibet）和横断山脉-川西高原（Hengduan Mountains–Western Sichuan Plateau, HDM-WSP）两个谱系，且在喜马拉雅-横断山脉/川西高原谱系中发现位于两个基因中的三个单核苷酸多态性位点对旱季降水存在强的局部适应。Liu 等（2022）进一步基于核微卫星数据发现了川滇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。

1.5 技术路线

本研究首先利用生物信息分析方法对川滇高山栎的基因组进行测序、组装和注释，构建了高质量的高山栎参考基因组图谱。通过系统发育及正选择基因鉴定，初

步探究了川滇高山栎对高山环境适应性进化的基因及功能。进一步筛选出正选择 *QaGDU3* 基因，利用分子克隆的手段，将其转化至异源表达载体拟南芥中，同时结合生物信息分析方法对其基因功能进行了验证。结合荧光定量分析和镰刀菌根腐病的抗性研究，揭示了 *QaGDU3* 基因与植物抗性之间的关系，推测 *QaGDU3* 基因对川滇高山栎适应性的影响。

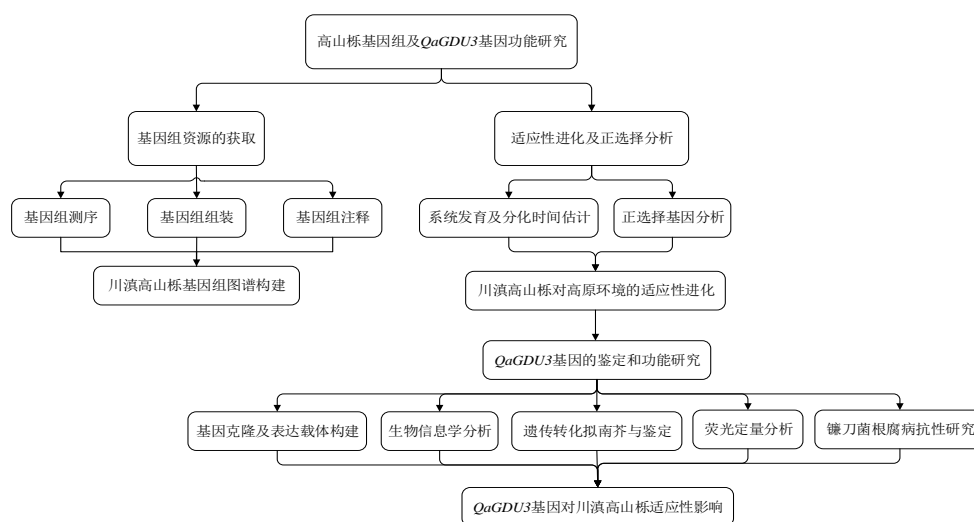


图 1.1 研究技术路线图

Figure 1.1 Research technology route

1.6 本文的研究内容、研究目的及意义

1.6.1 研究内容

本研究结合基因组学、生物信息学和分子克隆等手段，以川滇高山栎为研究对象，构建川滇高山栎基因组图谱，探究川滇高山栎的进化地位和分类起源，分析鉴定了川滇高山栎基因组中的正选择基因及功能，克隆得到了正选择 *QaGDU3* 基因并对其分子功能进行了探究。

1.6.2 研究目的和意义

本研究拟在解决以下几个方面的问题：

- (1) 结合二代 Illumina 和三代 Nanopore 测序数据构建染色体水平高质量的川滇高山栎参考基因组图谱，通过基因组注释信息全面解析川滇高山栎基因组序列特征，丰富栎属植物的基因资源。
- (2) 构建系统进化树，从全基因组水平探究川滇高山栎的进化地位，并鉴定川滇高山栎基因组中正选择基因，探究川滇高山栎适应性进化的基因功能。
- (3) 从川滇高山栎基因组中克隆 *QaGDU3* 基因并进行生物信息学分析，构建异源植

物表达载体，将其转化至拟南芥中。结合表型观察与统计分析，探究转基因 *QaGDU3* 拟南芥大小与 *QaGDU3* 基因表达量之间的关系。

- (4) 进一步通过实时荧光定量 PCR (quantificational Real-Time Polymerase Chain Reaction, qRT-PCR) 分析不同转基因拟南芥中 *QaGDU3* 与水杨酸 (Salicylic Acid, SA) 途径相关基因的表达关系。结合 *QaGDU3* 转基因拟南芥对镰刀菌的抗性结果，探究 *QaGDU3* 基因与植物抗性间的相关作用，推测 *QaGDU3* 基因对川滇高山栎适应性影响。

本研究获得的川滇高山栎基因组资源，填补了栎属冬青栎组基因组资源的空白，其基因组数据为进一步研究该物种的种群遗传及保护提供了参考，为其它栎属物种的基因组研究提供宝贵的资源。正选择 *QaGDU3* 基因的挖掘为进一步解析该基因的功能、探究其在植物体内的作用机制提供了参考，同时对于鉴定和分析栎属植物适应性进化相关基因功能和调控机制提供了方向。

2 高质量高山栎基因组组装与注释

2.1 基因组资源的获取

2.1.1 样品的采集与处理

本研究所需幼嫩的叶片样品，以及用于 RNA-seq 的嫩芽组织均采集于中国云南省大理白族自治州剑川县滇南镇水井村 214 国道旁的一棵川滇高山栎（北纬 26°22'45", 东经 99°58'21"）（图 2.1）。将采集得到的新鲜叶片立刻放置于 -80°C 液氮中，后续带回实验室进行 DNA/RNA 的提取（Doyle, 1987）。同时为了保证测序样品质量和测序结果准确性，我们直接对采取得到的新鲜的川滇高山栎植物样本进行 Hi-C 测序相关实验操作（Dekker *et al.*, 2002）。

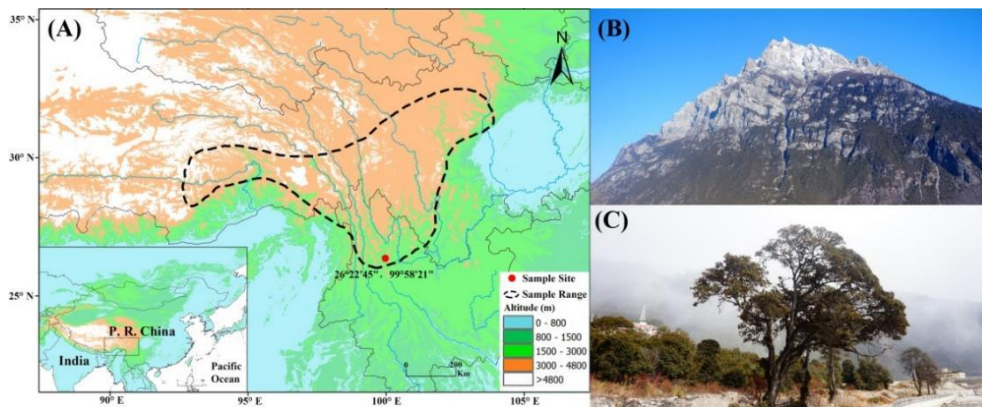


图 2.1 川滇高山栎分布区域和用于基因组测序的个体位置 (A)、生境 (B) 及个体样本 (C)

Figure 2.1 *Quercus aquifolioides* distribution range (dashed line), and location of the sequenced tree (red dot) (A), local habitat (B) and photo of the tree sampled for genome sequencing (C)

2.1.2 实验方法

本研究采用商品化的试剂盒对川滇高山栎样本材料的 DNA 和 RNA 进行提取。使用 Ultra Long 植物提取试剂盒（希望组，中国）从川滇高山栎叶片组织中提取 DNA，采用 RNA prep Pure 多糖多酚植物总 RNA 提取试剂盒（Cat #DP441, QIAGEN, Hilden, Germany）提取川滇高山栎 RNA。川滇高山栎全基因组测序、组装和注释委托北京希望组生物科技有限公司进行。

2.2 适应性进化及正选择分析

2.2.1 其它 13 个物种基因组数据获取

本研究选取 14 个不同进化位置中具有代表性且基因组组装质量较好的双子叶植物进行比较基因组及系统发育分析。从 NCBI (<http://www.ncbi.nlm.nih.gov/>) 下载得到除川滇高山栎外其它 13 个物种的基因组序列, 其中包括 2 个常绿木本植物: 欧洲栓皮栎 (*Quercus suber*; PRJNA433227) (Ramos *et al.*, 2018) 和川滇高山栎 (*Quercus aquifolioides*); 10 个落叶木本植物: 加州白栎 (*Quercus lobata* v1.0; PRJNA308314) (Sork *et al.*, 2016)、有柄栎 (*Quercus robur*; PRJEB19898) (Plomion *et al.*, 2018)、桦树 (*Betula pendula*; PRJEB14544) (Salojärvi *et al.*, 2017)、中国板栗 (*Castanea mollissima*; PRJNA527178) (Xing *et al.*, 2019)、欧洲白蜡 (*Fraxinus excelsior*; PRJEB4958) (Sollars *et al.*, 2017)、苹果 (*Malus domestica*; PRJNA534520) (Daccord *et al.*, 2017)、桃树 (*Prunus persica*; PRJNA31227) (Verde *et al.*, 2013)、葡萄 (*Vitis vinifera*; PRJEA18785, PRJDB5761) (Jaillon *et al.*, 2007; Canaguier *et al.*, 2017)、毛果杨 (*Populus trichocarpa*; PRJNA17973) (TusKana *et al.*, 2006) 和巨桉 (*Eucalyptus grandis*; PRJNA49047) (Myburg *et al.*, 2014) 和以及 2 个草本植物: 草莓 (*Fragaria vesca*; PRJNA66853) (Shulaev *et al.*, 2011) 和拟南芥 (*Arabidopsis thaliana*; PRJNA311266) (Zapata *et al.*, 2016)。然后使用 OrthoMCL v2.0.9 (Li *et al.*, 2003) 对以上所选择的 14 个物种的基因家族进行鉴定: 首先对 14 个物种基因组中所有基因的可变剪切进行过滤, 保留每个基因最长的转录本用于分析。利用 BLASTp (Jacob *et al.*, 2008), 参数为 $E\text{-value} \leq 1 \times 10^{-5}$, 将 14 个物种的所有蛋白质序列分别进行两两之间的 All-v-All 比对, 并基于 OrthoMCL v2.0.9 软件中的马尔可夫聚类算法模型对比对结果进行聚类分析, 得到所有物种的基因家族聚类结果。同时将 BLASTp 比对最好的结果作为直系同源基因, 最后保留至少包含 10 个物种的直系同源基因家族。

2.2.2 系统发育分析与分化时间估计

利用 OrthoMCL v2.0.9 软件提取 14 个用于比较基因组分析的物种中的所有单拷贝直系同源基因。将鉴定得到的单拷贝基因家族的蛋白序列通过 Mafft v7.313 (Katoh and Standley, 2013) 进行比对, 比对较差的区域通过 Gblocks v0.91b (Castresana, 2000) 过滤掉, 然后将保守位点提取并连接起来, 最终得到每个物种特有的比对后的序列。在此基础上, 使用 RAxML v8.2.10 (Stamatakis, 2006) 软件, 采取 GTRGAMMA 预测最佳替代结构模型, 以葡萄为外类群, bootstrap 重复值设为 100, 构建 14 个物种的最大似然系统发育树。

基于化石记录物种进化时间，分别以 Time Tree (<http://www.timetree.org>) (Hedges *et al.*, 2006) 数据库中的葡萄和拟南芥 (111.99-131.13Mya)，拟南芥和巨桉 (97.83-115.81Mya) 以及巨桉和中国板栗 (5.62-16.18Mya) 的分化时间进行校正。使用 PAML v4.9e (Yang, 2007) 软件包中基于马尔可夫链中的 MCMCTREE 软件估计分化时间，MCMCTREE 软件运行的主要参数为 clock=2, RootAge=<1.17, model=7, BDparas=110, kappa_gamma=62, alpha_gamma=11, rgene_gamma=23.18, sigma2_gamma=14.5。

2.2.3 正选择基因鉴定及功能分析

使用 PAML 软件包中的 CodemL 工具，选择枝位点模型 (branch-site model) 检测川滇高山栎基因组中的正选择基因 (Yang, 2007)。以川滇高山栎为前景枝 (foreground branches)，加州白栎、有柄栎、欧洲栓皮栎和毛果杨分枝定义为背景枝 (background branches)。重要参数为：替代模型 (alternative model) (Model A) : model = 2, NSsites = 2, fix_omega = 0; 零假设模型 (null model) (Model A1) : model = 2, NSsites = 2, fix_omega = 1, omega = 1。对零假设模型和替代模型得出的最大似然值进行似然比检测 (the likelihood ratio, LRT) (Zhang *et al.*, 2005)，并利用卡方检验 (自由度为 1) 计算 p 值，然后结合错误发现率 (False Discovery Rate, FDR) 对 p 值进行校正，以 FDR<0.05 作为正选择的阈值。通过经验贝叶斯法 (Bayes Empirical Bayes, BEB) 筛选出后验概率值 ≥ 0.95 的正选择基因，而其它基因被鉴定为假阳性并排除 (Yang *et al.*, 2005)。

使用 clusterProfiler (Yu *et al.*, 2012) 富集软件对川滇高山栎基因组中正选择基因功能在基因本体论 (GO, Gene Ontology) 和京都基因与基因组百科 (KEGG, Kyoto Encyclopedia of Genes and Genomes) 两个数据库中进行富集分析，并筛选 P-value<0.05 为显著富集的结果。

2.3 研究结果

2.3.1 基因组结果

(1) DNA 检测及质控

提取的川滇高山栎基因组 DNA 经琼脂糖凝胶电泳结果显示 DNA 条带清晰明亮，无拖尾 (图 2.2)。核酸浓度检测结果显示 OD260/OD280 的比值为 1.83，OD260/OD230 比值为 2.39，说明提取的川滇高山栎 DNA 完整度较好，纯度较高，总量满足建库要求，可以用于建库及测序。

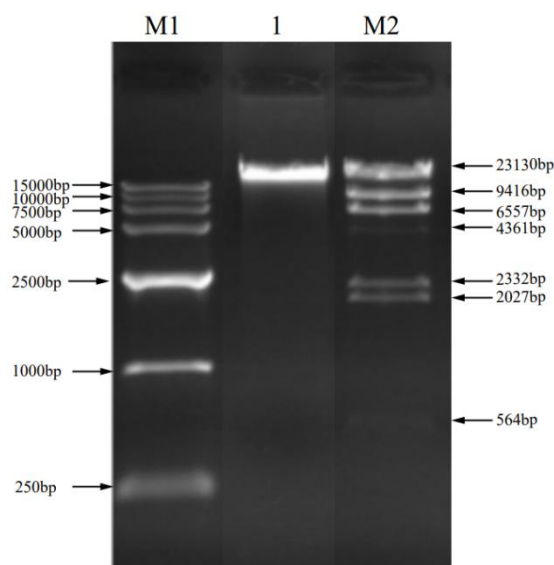


图 2.2 川滇高山栎 DNA 凝胶电泳图

Figure 2.2 The DNA electrophorogram of *Quercus aquifolioides*

注 a: M1: 15kb DNA Marker; M2: λ DNA/HindIII; 1: 川滇高山栎 DNA。

(2) 测序数据统计

去除二代测序下机数据中低质量、含接头以及重复的序列后得到了 124.53Gb 高质量的 clean reads，三代测序数据通过 GridION X5 平台的 21 个 MinION flow cells 生成了 7 213 768 个 reads，包含 99 231 406 295 个碱基，N50 为 19 894.86bp，通过 5 个额外的 flow cells 产生超长 Ultra-long reads，平均每个 cell 数据 10.39Gb。此外，通过 Illumina NovaSeq 平台进行双端测序，还构建了 Hi-C 文库，得到 223.66Gb 的数据。总共有 484.94Gb 的数据用于后续的基因组分析（表 2.1）。

以 500bp 无重叠区域为窗口计算基因组 GC 含量并统计其测序深度，根据 GC depth 分布图显示，发现 GC 含量集中在 30%-40%，测序深度集中在 50X-150X（图 2.3）。

表 2.1 川滇高山栎基因组测序信息

Table 2.1 Sequencing data statistics of *Quercus aquifolioides*

测序类型	测序平台	文库大小 (bp)	过滤后数据量 (Gb)
Genome short reads	Illumina Hiseq X ten	350	124.53
Nanopore and ultra-long reads	Nanopore platform	20 000	124.73
Hi-C reads	Illumina NovaSeq	300-700	223.66
Transcriptome short reads	Illumina NovaSeq 6000	150-500	12.02
Total	-	-	484.94

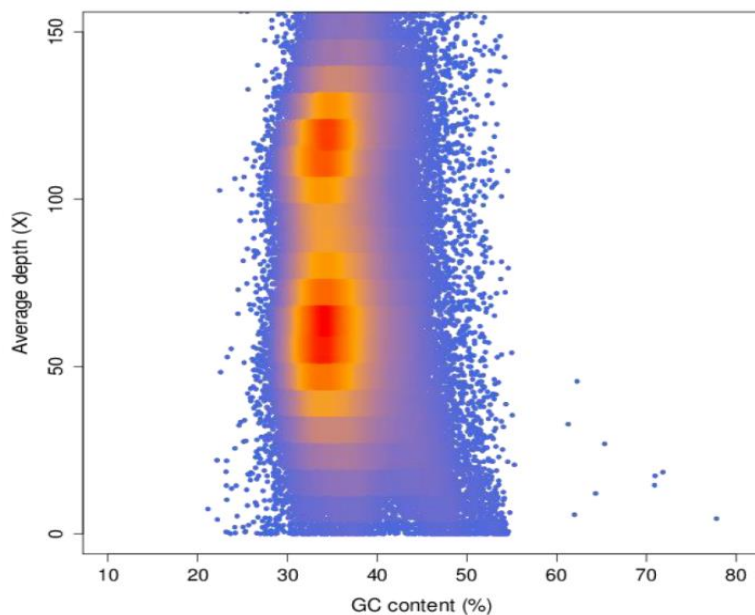


图 2.3 川滇高山栎基因组 GC 含量与测序深度关系

Figure 2.3 The relationship between GC content and sequencing depth of *Quercus aquifolioides* genome

(3) 基因组组装

结合 124.73Gb (109X) 三代 Nanopore 数据和 124.53Gb (109X) 二代 Illumina 数据对川滇高山栎基因组进行初步组装，纠错后最终组装得到川滇高山栎基因组大小约为 1.15Gb，contig N50 数量为 304 个，contig N50 长度约 1.20Mb (表 2.2)。

表 2.2 川滇高山栎基因组组装信息

Table 2.2 Assembly statistics of the *Quercus aquifolioides* genome

统计类型	Contig 长度 (bp)	Contig 数量
N50	1 202 933	304
N60	983 692	410
N70	804 782	538
N80	613 413	700
N90	416 514	925
Longest	6 700 662	1
Total	1 147 540 021	1 445
Length>=10kb	1 147 540 021	1 445

(4) Hi-C 辅助组装

利用 LACHESIS 软件基于 Hi-C 文库获得的 223.66Gb (195X) 的 clean data，将川滇高山栎基因组进行染色体水平的组装。最终将 1 301 条 scaffolds 组装到 12 条假

染色体上，覆盖基因组的96.02%（1 101.96Mb）（表 2.3）。结果可以明显的区分 12 个染色体组，且在每个染色体组内相邻的序列间（对角线）交互作用强度高，而不相邻的序列间（非对角线）的交互作用强度弱（图 2.4），证明基因组组装效果较好。

表 2.3 川滇高山栎 scaffolds 聚类结果

Table 2.3 The result of *Quercus aquifolioides* clustered scaffolds

染色体	大小 (bp)	Scaffold 数量
cluster01	135 503 775	151
cluster02	133 707 883	179
cluster03	132 465 767	155
cluster04	97 117 131	126
cluster05	91 241 144	103
cluster06	74 660 713	98
cluster07	80 804 455	108
cluster08	84 628 911	96
cluster09	75 188 798	70
cluster10	68 973 880	95
cluster11	69 943 310	77
cluster12	57 727 289	74
Total	1 101 963 056	1 301

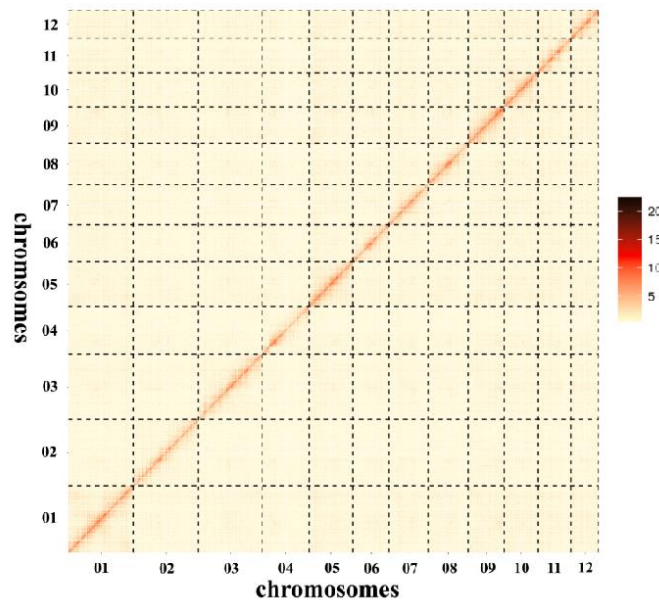


图 2.4 川滇高山栎基因组 Hi-C 交互热图

Figure 2.4 The Hi-C heatmap of *Quercus aquifolioides* genome showing genome-wide all-by-all interactions

(5) 组装质量评估

为了评估组装序列的完整性，本研究从 BUSCO 和 LTR Assembly Index (LAI) 两方面来评价川滇高山栎基因组组装质量。BUSCO 结果显示基因组完整性为 95.40% (BUSCO groups 总数为 1 375)，包括 83.30% 的单拷贝基因和 12.10% 的重复基因。同时，LTR 组装指数 LAI 为 11.31，为可参考水平 (表 2.4)。

表 2.4 川滇高山栎基因组组装结果评估

Table 3.4 Validation of *Quercus aquifolioides* genome assembly

物种	完整性		片段基 因	丢失基 因	LAI	LAI 标准	类别
	重复基 因	单拷贝 基因					
					-	$0 \leq \text{LAI} < 10$	Draft
<i>Quercus aquifol</i>	83.30%	12.10%	1.20%	3.40%	11.31	$10 \leq \text{LAI} < 20$	Reference
<i>ioides</i>					-	$20 \leq \text{LAI}$	Gold

(6) 基因组注释

结合从头预测和同源注释的预测方法，在川滇高山栎基因组中注释到约 666.10Mb 的重复序列，约占基因组的 58.04% (表 2.5)。其中，长末端重复序列为基因组中重复元件的主要类型，约占整个基因组的 40.86% (468.90Mb) (表 2.6)。

表 2.5 重复序列统计结果

Table 2.5 The statistics of repeat sequences

软件	重复大小 (bp)	占基因组的比例 (%)
Trf	43 647 205	3.80
RepeatProteinMask	159 181 410	13.87
RepeatMasker	641 856 328	55.93
Total	666 080 407	58.04

表 2.6 重复类型统计结果

Table 2.6 The statistics of repeat type

重复类型	RepeatProteinMask		RepeatMasker		Combined TEs	
	长度 (bp)	占基因组比例 (%)	长度 (bp)	占基因组比例 (%)	长度 (bp)	占基因组比例 (%)
DNA	9 805 880	0.85	30 729 163	2.68	33 316 855	2.90
LINE	37 482 002	3.27	39 393 516	3.43	47 689 830	4.16
LTR	111 646 016	9.73	466 774 684	40.68	468 900 718	40.86
SINE	0	0	174 003	0.02	172 094	0.01
Total	158 933 898	13.85	537 071 366	46.81	550 079 497	47.93

结合 de novo 预测、同源预测以及基于转录组数据的预测，共得到川滇高山栎基因组中 26 441 个蛋白质编码基因，平均基因长度为 5 583.47bp，平均蛋白质编码序列大小为 1 101.53bp，平均每个基因含有 5.55 个外显子（表 2.7）。

表 2.7 川滇高山栎与其它 13 个物种基因结构注释结果

Table 2.7 The statistics of gene structure annotation of *Quercus aquifolioides* and other 13 species

基因集	基因数量	平均基因长度(bp)	平均蛋白编码序列长度(bp)	每个基因平均外显子数	平均外显子长度(bp)	每个基因内含子数	平均内含子长度(bp)
<i>Quercus robur</i>	60 265	3 082.39	797.78	3.47	230.00	2.47	925.59
<i>Quercus suber</i>	61 281	4 107.12	1 362.58	6.40	212.74	5.40	507.87
<i>Malus domestica</i>	58 902	3 793.92	1 084.94	4.18	259.58	3.18	852.06
<i>Populus trichocarpa</i>	49 300	4 436.92	1 181.78	4.63	255.45	3.63	897.76
<i>Prunus persica</i>	54 659	4 353.49	1 195.75	4.66	256.86	3.65	864.01
<i>Vitis vinifera</i>	55 786	4 145.92	1 097.21	4.35	252.15	3.35	909.79
<i>Arabidopsis thaliana</i>	40 741	3 950.32	1 022.30	4.21	242.65	3.21	911.43
<i>Carica papaya</i>	18 003	4 339.19	1 218.03	5.51	221.15	4.51	692.40
<i>Fragaria vesca</i>	57 215	3 998.66	1 090.97	4.16	262.08	3.16	919.46
Homolog							
<i>Glycine max</i>	58 332	4 028.58	1 028.68	3.98	258.50	2.98	1 006.99
<i>Cucumis sativus</i>	45 124	4 619.49	1 108.06	4.48	247.58	3.48	1 010.45
<i>Oryza sativa</i>	52 931	3 015.54	944.96	3.45	273.94	2.45	845.36
<i>Ricinus communis</i>	60 391	3 464.05	903.89	3.66	247.25	2.66	964.14
De novo							
Geneid	74 534	5 953.52	734.18	4.23	173.67	3.23	1 617.15
Genscan	46 705	14 495.77	1 062.06	5.99	177.30	4.99	2 692.05
August	71 226	2 378.13	774.06	3.96	195.40	2.96	541.66
RNA-seq							
Stringtie	49 146	9 629.36	1 168.24	3.77	309.50	2.77	3 049.48
PASA	40 614	5 476.73	975.36	4.31	226.38	3.31	1 360.54
EVM							
<i>Quercus aquifolioides</i>	26 441	5 583.47	1 101.53	5.55	198.54	4.55	985.46

将基因结构预测得到的蛋白质序列与已知蛋白库进行比对，结果显示有 24 376 个基因被预测出功能，占基因总数的 92.19%（表 2.8）。此外通过与已知 ncRNA 库进行比对，预测得到川滇高山栎基因组中有 982 个 tRNAs、157 个 rRNAs、511 个 miRNAs 和 1 136 个 snRNAs。综上所述，我们得到了染色体水平上高质量的川滇高山栎参考基因组图谱（图 2.5）。

表 2.8 川滇高山栎基因组功能注释结果

Table 2.8 Functional annotation of genes of *Quercus aquifolioides*

数据库	数量	注释比例 (%)
KEGG	18 801	71.10
COG	14 454	54.66
GO	18 341	69.36
Swissprot	19 968	75.52
Trembl	21 151	79.99
Annotated	24 376	92.19

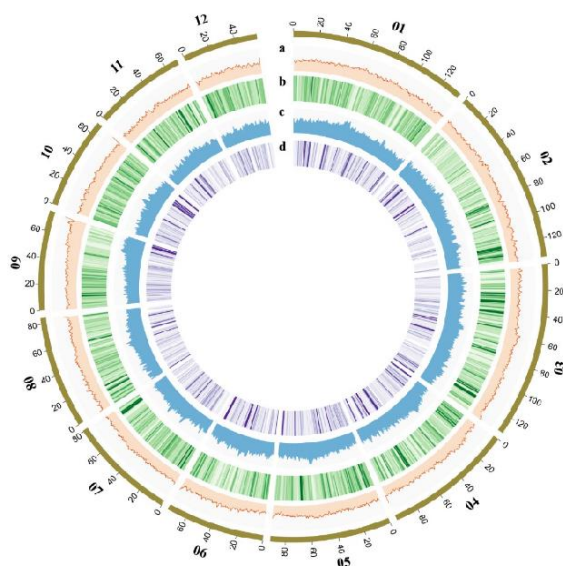


图 2.5 川滇高山栎基因组结构

Figure 2.5 Genomic structure of *Quercus aquifolioides*

注 a: a: GC 含量 (30.00~45.00%) ; b: 基因密度 (0~30.00%) ; c: 重复序列密度 (0~100.00%) ; d: 非编码 RNA 密度 (0~0.10%) 。

2.3.2 系统发育及分化时间估计

首先通过基因家族的聚类分析，鉴定得到 14 个物种共有的 318 个单拷贝基因家族。采用共有的 318 个单拷贝基因家族构建最大似然物种发育树，结果显示川滇高山栎与近缘物种欧洲栓皮栎、加州白栎、有柄栎聚为一枝。（图 2.6）。

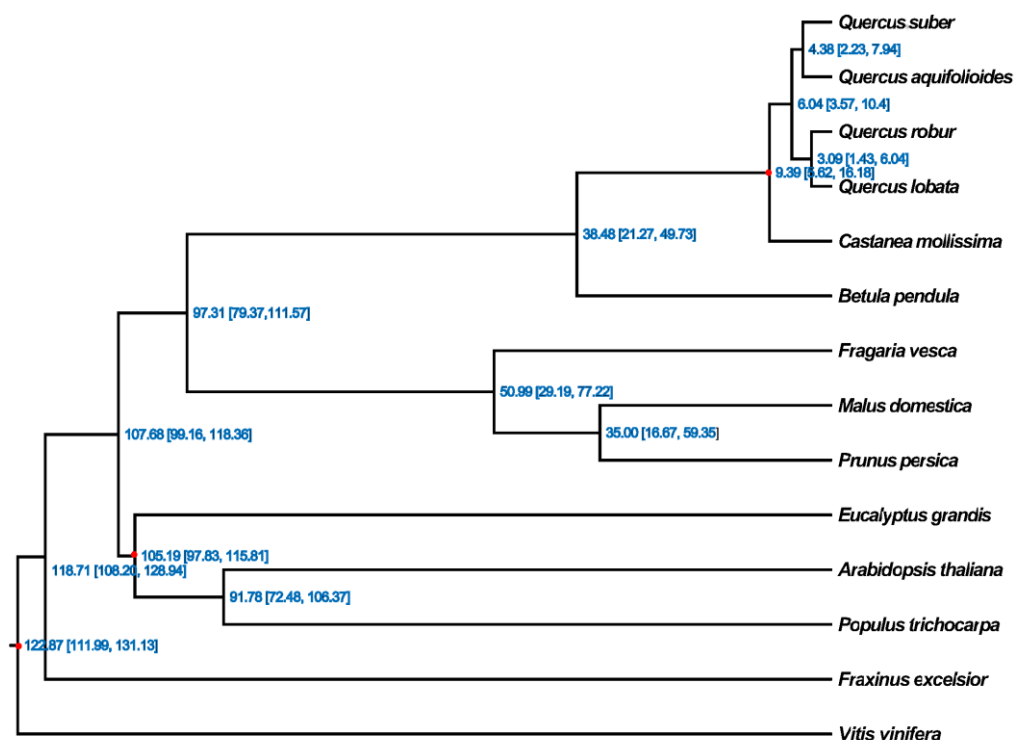


图 2.6 川滇高山栎分化时间估测

Figure 2.6 Estimation of divergence time of *Quercus aquifolioides*

2.3.3 正选择基因及其功能

为了评估川滇高山栎基因组的适应性进化，我们鉴定得到了基因组中 416 个正选择基因 ($P \leq 0.05$)，其中 310 个基因功能可以在 Swissprot 数据库中得到注释。为了进一步了解正选择基因的功能类型，将所有的正选择基因进行 KEGG 富集分析来探究正选择基因主要参与的代谢途径。以 $P\text{-value} < 0.05$ 为筛选条件，结果显示其主要富集在“糖基磷脂酰肌醇 (GPI)-锚定生物合成 (Glycosylphosphatidylinositol (GPI)-anchor biosynthesis)”、“花生四烯酸代谢 (Arachidonic acid metabolism)”、“液泡运输 (Vacuolar transport)”和“核苷酸切除修复 (Nucleotide excision repair)”四个 KEGG 代谢通路 (表 2.9)。

表 2.9 川滇高山栎正选择基因功能 KEGG 富集结果

Table 2.9 KEGG enrichment analysis of positive selection genes function in *Quercus aquifolioides* genome

KEGG ID	描述	P-value	基因数量
map00563	Glycosylphosphatidylinositol (GPI)-anchor biosynthesis	0.00	4
map00590	Arachidonic acid metabolism	0.01	3
map03420	Nucleotide excision repair	0.04	4

将正选择基因进行 GO 富集分析，以 P-value<0.05 为筛选条件，结果显示正选择基因主要集中在生物学过程和分子功能两个方面，尤其是 GO terms“RNA 修饰 (RNA modification)”、“假尿苷合酶活性 (Pseudouridine synthase activity)”、“液泡运输 (Vacuolar transport)”和“假尿苷合成 (Pseudouridine synthesis)”。其中 GO terms“RNA 修饰”是正选择基因富集最多的生物学过程，GO terms“假尿苷合酶活性”是富集最多的分子功能。同时，被富集的还包括与核苷酸修饰、加工和结合等相关的“DNA 模板转录，起始 (DNA-templated transcription, initiation)”和“RNA 加工 (RNA processing)”等生物学过程和“RNA 结合 (RNA binding)”、“受损 DNA 结合 (Damaged DNA binding)”等分子功能，其中与“核苷酸切除修复”相关的基因功能在 GO 和 KEGG 中均被富集。除此之外，与光合作用相关的“光系统 II 氧复合物 (Photosystem II oxygen evolving complex)”、与“质子跨膜转运 (Proton transmembrane transport)”相关的基因功能也被富集 (表 2.10)。

在富集的结果中，质子的跨膜转运是植物体能够进行正常物质交换、能量转换和信息传递的基础，包括了无机离子、激素及氨基酸等各种物质的跨膜转运。其中赖氨酸和组氨酸转运白 (Lysine-Histidine-Like Transporters, LHTs)、脯氨酸转运蛋白 (Proline Transporters, ProTs) 以及谷氨酰胺转运蛋白 (Glutamine Dumper Protein, GDU) 等氨基酸的跨膜转运蛋白在维持植物新陈代谢等生命活动中发挥重要作用。AtGDU 作为拟南芥中氨基酸的转运蛋白，同时 *AtGDU3* 基因的过表达可激活 SA 途径相关基因的表达从而增强转基因拟南芥的抗病性，因此我们对筛选出的川滇高山栎正选择 *QaGDU3* 基因通过基因克隆手段，对其分子功能进行验证，从而探究 *QaGDU3* 基因对川滇高山栎适应高山环境的作用。

表 2.10 川滇高山栎正选择基因功能 GO 富集结果

Table 2.10 GO enrichment analysis of positive selection genes function in *Quercus aquifolioides* genome

GO ID	功能描述	分类	P-value	基因数量
GO:0009451	RNA modification	BP	0.00	4
GO:0009982	pseudouridine synthase	MF	0.00	4

	activity			
GO:0007034	vacuolar transport	BP	0.00	3
GO:0001522	pseudouridine synthesis	BP	0.00	4
GO:0006506	GPI anchor biosynthetic	BP	0.00	3
	process			
GO:0004222	metalloen dopeptidase	MF	0.00	4
	activity			
GO:0004518	nuclease activity	MF	0.00	3
GO:0005730	nucleolus	CC	0.02	2
GO:0003723	RNA binding	MF	0.02	9
GO:0006289	nucleotide-excision repair	BP	0.02	2
GO:0000226	microtubule cytoskeleton	BP	0.03	2
	organization			
GO:0003684	damaged DNA binding	MF	0.03	2
GO:0008168	methyltransferase activity	MF	0.03	7
GO:0015992	proton transmembrane	BP	0.04	4
	transport			
GO:0006352	DNA-templated	BP	0.04	2
	transcription, initiation			
GO:0006396	RNA processing	BP	0.04	4
GO:0009654	photosystem II oxygen	CC	0.05	2
	evolving complex			

3 川滇高山栎 *QaGDU3* 基因克隆及其功能分析

3.1 试验材料

3.1.1 植物材料

本节研究所用川滇高山栎叶片与 2.1 相同，实验材料采集后干燥密封保存，后期用于实验室内 DNA 提取。野生型 (Wild-Type, WT) 拟南芥 (*Arabidopsis thaliana* (L.) Heynh.) 种子 (生态型为 Col-0) 由北京林业大学生态与自然保护学院分子生态实验室提供。

菌株及载体:

感受态根癌农杆菌 EHA105	北京康润诚业
DH5 α 大肠杆菌感受态细胞	上海唯地
pJG081 克隆载体	清华大学刘玉乐教授提供
13 个镰刀菌	北京林业大学生态与自然保护学院分子生态实验室提供

3.1.2 试验试剂

植物 DNA 提取试剂盒	TIANGEN
总 RNA 提取试剂 TRIGene	TIANGEN
Premix Taq (LA Taq Version 2.0)	宝日医
50 \times TAE Buffer	艾德莱
微柱浓缩 DNA 凝胶回收试剂盒	北京庄盟
氨苄青霉素、利福平、卡那霉素	上海生工
10 \times Buffer	北京全式金
100 μ m dATP	北京全式金
T4 DNA 聚合酶	北京全式金
0.1%BSA	北京全式金
BeyoRT™ II cDNA 合成试剂盒 (with gDNA Eraser)	碧云天
BeyoFast™ SYBR Green qPCR Mix (2X)	碧云天

3.1.3 实验试剂配制

(1) 抗生素

50mg/mL 卡那霉素 (Kana)：称取 10g 卡那霉素并使用足量的无菌水溶解，最后定容至 200mL，过滤灭菌分装至 1.5mL 无菌离心管中，-20°C 贮存。

50mg/mL 利福平 (Rif)：称取 10g 利福平并使用足量的二甲基亚砷 (DMSO) 溶解，最后定容至 200mL，过滤灭菌分装至 1.5mL 无菌离心管中，-20°C 贮存。

(2) 培养基

LB (Luria-Bertani) 液体培养基：称取 2g 蛋白胨 (Tryptone)，2g 氯化钠 (NaCl)，1g 酵母提取物 (Yeast extract)，加入足量的去离子水溶解，最终定容至 200mL，121°C 高压灭菌 30min 后使用。

LB (Luria-Bertani) 固体培养基：称取 2g 蛋白胨 (Tryptone)，2g 氯化钠 (NaCl)，1g 酵母提取物 (Yeast extract)，3g 琼脂粉 (Agar)，加入足量的去离子水溶解，最终定容至 200mL，121°C 高压灭菌 30min 后使用。

1/2MS (Murashige&Skoog) 固体培养基：0.44g 1/2MS，2g 蔗糖 (Sucrose)，1.6g 琼脂粉 (Agar)，加入去离子水溶解，pH 调至大于 5.8，定容至 200mL，121°C 高压灭菌 30min 后使用。

水琼脂固体培养基：2g 琼脂粉 (Agar)，加入去离子水溶解，定容至 200mL，121°C 高压灭菌 30min 后使用。

马铃薯葡萄糖琼脂 (Potato Dextrose Agar, PDA) 培养基：称取 100g 去皮马铃薯并切块，加入 500mL 去离子水煮沸至玻璃棒能戳动土豆块，过滤土豆得滤液，加入 7.5g 葡萄糖 (Glucose)，0.025g 氯霉素 (Chloramphenicol)，8.5g 琼脂粉 (Agar)，加入去离子水定容至 500mL，121°C 高压灭菌 30min 后使用。

马铃薯葡萄糖水 (Potato Dextrose Broth, PDB) 培养基：称取 100g 去皮马铃薯并切块，加入 500mL 去离子水煮沸至玻璃棒能戳动土豆块，过滤土豆得滤液，加入 7.5g 葡萄糖 (Glucose)，0.025g 氯霉素 (Chloramphenicol)，加入去离子水定容至 500mL，121°C 高压灭菌 30min 后使用。

(3) DNA 检测

1×TAE：称量 20mL 50×TAE，加入去离子水定容至 1L，混合均匀后备用。

3.2 研究内容与方法

3.2.1 *QaGDU3* 基因的克隆及表达载体构建

(1) 目的片段的扩增

根据川滇高山栎基因组测序的 *QaGDU3* CDS 序列，利用 SnapGene Viewer v5.3.1

设计对应引物 Sca495-F/R (附录 A)。使用植物基因组 DNA 提取试剂盒 (天根, 中国) 提取川滇高山栎干叶组织总 DNA, 具体操作步骤如图 3.1:

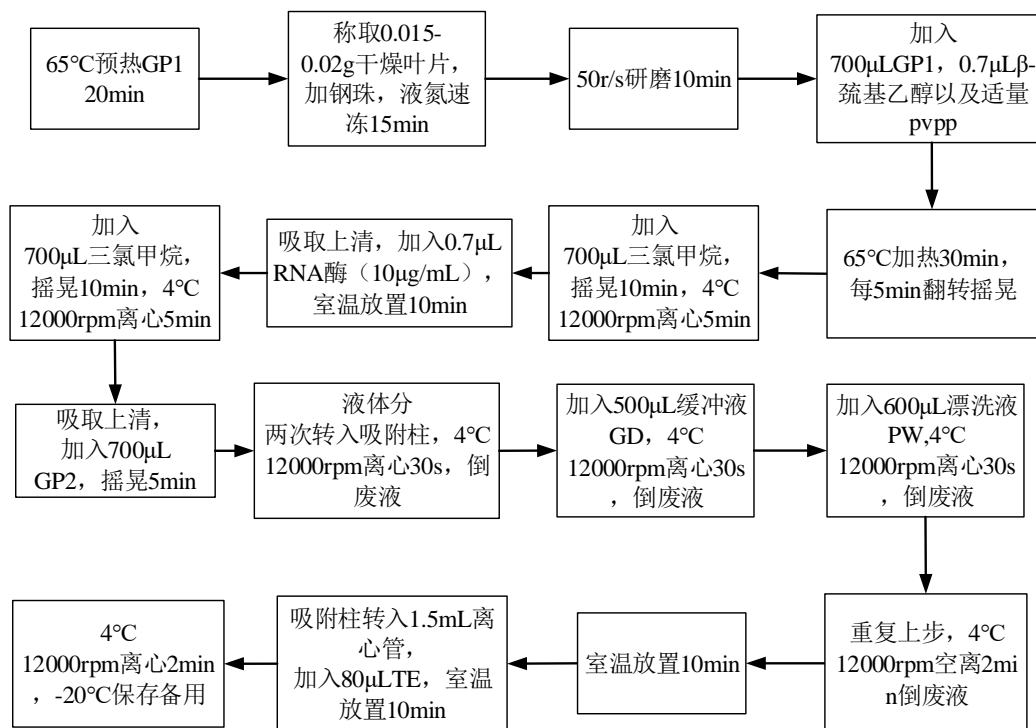


图 3.1 植物基因组 DNA 提取试剂盒提取 DNA 流程

Figure 3.1 DNA extraction process by Plant genomic DNA extraction kit

将提取得到的川滇高山栎 DNA 作为模板进行 *QaGDU3* 基因 PCR 扩增。PCR 反应体系参照 Premix Taq (LA Taq Version 2.0) 说明书。反应程序: 98°C 变性 10s, 56°C 退火 30s, 72°C 延伸 1min, 共 30 次循环, 16°C 保存。

(2) DNA 目的条带的回收

扩增产物用 1% 琼脂糖凝胶电泳检测, 并用微柱浓缩 DNA 凝胶回收试剂盒回收扩增得到的目的序列, 具体步骤如图 3.2:

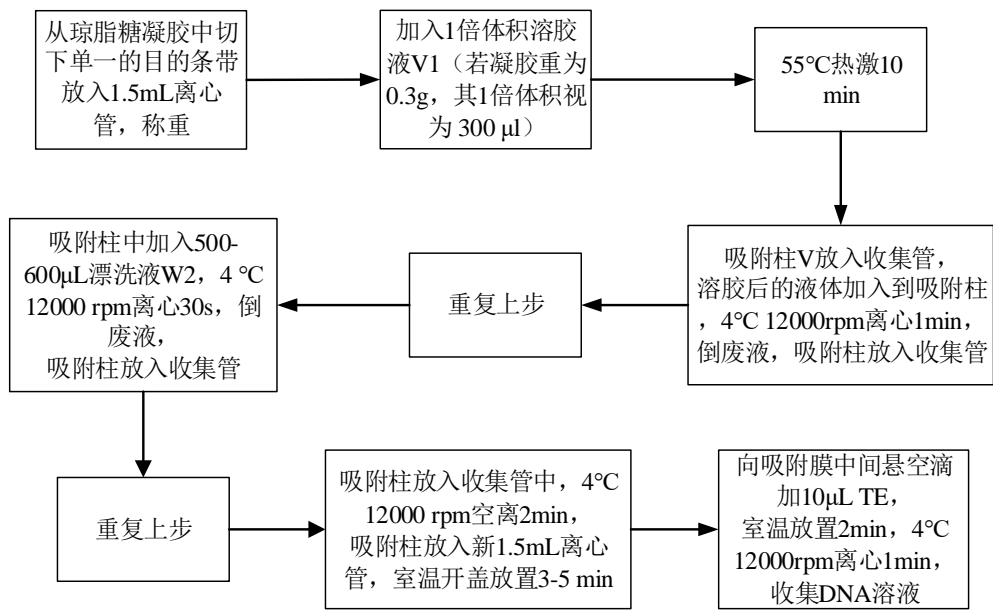


图 3.2 微柱浓缩 DNA 凝胶回收试剂盒回收 DNA 片段流程

Figure 3.2 Purification and recovery of DNA fragments process by V-ELUTE Gel Mini Purification Kit

(3) 构建 *QaGDU3*-pJG081 克隆载体

以 *QaGDU3* 基因片段纯化产物为模板，使用引物 Sca495-1F/R（附录 A）对目的基因片段添加不依赖于连接反应的克隆（Ligation-Independent cloning, LIC）接头。扩增的 PCR 反应体系及条件同目的片段的扩增。扩增产物用 1% 琼脂糖凝胶电泳检测，并用微柱浓缩 DNA 凝胶回收试剂盒回收扩增得到的目的序列。将回收后的产物按照 Zhang 和 Liu（2014）的方法与 pJG081 载体（图 3.3）进行 LIC 克隆。

LIC1 PCR 反应体系：

10×Buffer	2µL
dATP（100uM）	1µL
PCR 回收产物	10µL
T4 DNA 酶	0.4µL
0.1%BSA	2µL
ddH ₂ O	4.6µL
Total	20µL

室温加样，轻柔混合后短暂离心，反应程序：37°C 20min，75°C 20min，16°C 1min。

LIC2 PCR 反应体系：

LIC1 PCR 产物	6µL
pJG081 载体	4µL

Total 10 μ L
 反应程序：70 $^{\circ}$ C 10min，22 $^{\circ}$ C 10min，放置 4 $^{\circ}$ C 冰箱备用。

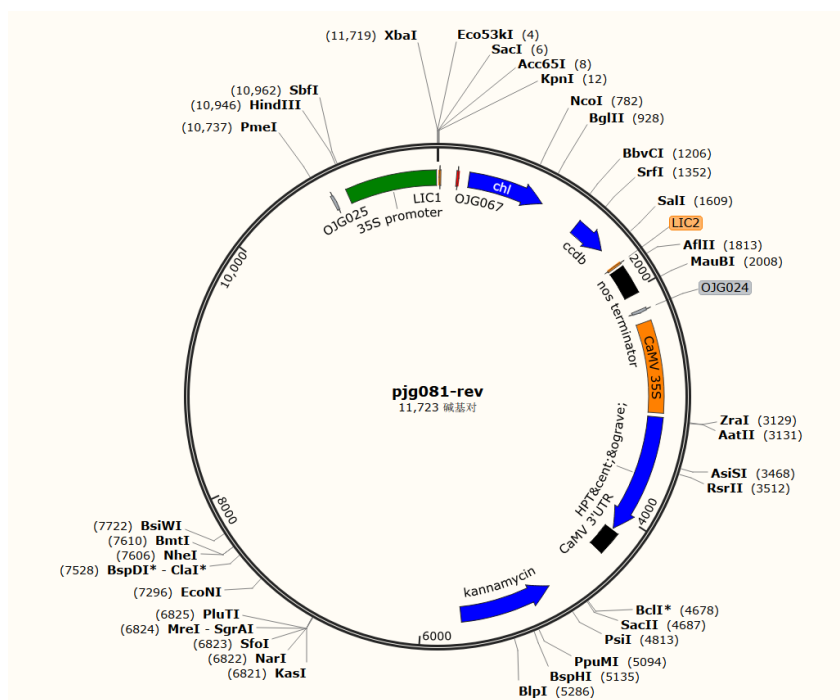


图 3.3 pJG081 载体结构图

Figure 3.3 pJG081 vector

(4) 连接产物转化 DH5 α 感受态

将大肠杆菌 DH5 α 感受态置于冰上解冻，取 5 μ L 连接产物至新的 1.5mL 离心管中，同时加 50 μ L 处于冰水混合状态的大肠杆菌 DH5 α 感受态，轻弹混匀后冰浴 30min；42 $^{\circ}$ C 热激 90s；冰上静置 2~3min；随后在超净工作台中加 1mL 无抗生素的 LB 液体培养基；220rpm，37 $^{\circ}$ C 振荡培养 1h，12 000rpm 离心 1min；弃上清至 100 μ L，使用移液枪反复吹打重悬菌体，并涂布于含有 50mg/L Kana 的 LB 固体培养基上，37 $^{\circ}$ C 倒置培养 24h，待培养基长出单菌落。

(5) 阳性克隆检测

使用无菌枪头挑取培养基上的单菌落于新的含有 50mg/L Kana 的 LB 固体培养基上进行划线，37 $^{\circ}$ C 倒置培养 12h 至长出菌落后，使用引物 Sca495-1F/R（附录 A）进行菌落 PCR 扩增。反应程序：95 $^{\circ}$ C 预变性 5min；95 $^{\circ}$ C 变性 30s，56 $^{\circ}$ C 退火 30s，72 $^{\circ}$ C 延伸 50s，共 28 个循环；72 $^{\circ}$ C 延伸 5min，16 $^{\circ}$ C 保存。经 1% 琼脂糖凝胶电泳挑取条带与目的片段一致的菌落加 10 μ L TE 稀释为菌液，送至北京睿博生物科技有限公司进行测序，使用 SnapGene Viewer 5.3.1 将测序结果与 *QaGDU3* 的 CDS 序列进行比对。

(6) 阳性质粒转化农杆菌 EHA105 及验证

将感受态根癌农杆菌 EHA105 置于冰上解冻，取 3 μ L 测序无误后的 pJG081-*QaGDU3* 质粒（公司返回）加入到 30 μ L 处于冰水混合态的感受态根癌农杆菌 EHA105 中，轻弹混匀，冰上静置 5min。然后依次置于液氮 5min，37 $^{\circ}$ C 热激 5min。加入 1mL 无抗生素的 LB 液体培养基，置于 28 $^{\circ}$ C，220rpm 摇床培养 2h。12 000rpm 离心 1min，弃上清至 100 μ L，涡旋混匀重悬菌体并涂布于含 50mg/L Kana 和 50mg/L Rif 的 LB 固体培养基上，28 $^{\circ}$ C 倒置培养 3 天，挑取培养基上单克隆菌株，以引物 Sca495-1F/R（附录 A）进行菌落 PCR 检测阳性转化菌株。

3.2.2 *QaGDU3* 基因的生物信息学分析

使用 ExPASy ProtParam 和 ProtScale (Wilkins *et al.*, 1999) 推测 *QaGDU3* 蛋白质的基本理化性质并预测蛋白亲水性，使用 TMHMM (Moller *et al.*, 2001) 对 *QaGDU3* 蛋白的跨膜区域进行预测。并用 SignalP v4.1 Server (Nielsen, 2017) 预测其蛋白质信号肽。分别使用 Predict Protein (Bernhofer *et al.*, 2021) 和 Robetta (Yang *et al.*, 2020) 预测 *QaGDU3* 蛋白的二级和三级结构，使用 ProtComp9.0 预测 *QaGDU3* 蛋白的亚细胞定位，并使用 MotifScan (Pagni *et al.*, 2007) 进行蛋白质的功能位点预测。分别使用 NCBI 在线网站 conserved domains (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) 及 BLASTp (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&BLAST_SPEC=&LINK_LOC=blasttab&LAST_PAGE=blastp) 进行蛋白质结构域预测和蛋白序列比对，通过 DNAMAN v9.0 软件对蛋白序列进行多重比对，同时使用 MEME (Bailey *et al.*, 2015) 查找 motif (附录 B)。

3.2.3 *QaGDU3* 基因的系统发育分析

为了探究 *GDU3* 在 *GDU* 家族中的进化位置，我们对拟南芥和四个栎属植物中 *GDU* 家族进行了系统发育分析。首先从 TAIR 网站 (<https://www.arabidopsis.org>) 中下载 AtGDU1、AtGDU2、AtGDU3、AtGDU4、AtGDU5、AtGDU6、AtGDU7 七个氨基酸序列，通过 NCBI 在线网站 BLASTp 进行比对，得到四个栎属植物中七个 AtGDU 同源序列，使用 MEGA v7 软件 (Kumar *et al.*, 2016) 参数设置为 Bootstrap method 和 500 bootstrap，构建拟南芥和四个栎属植物 *GDU* 家族最大似然 (Maximum Likelihood, ML) 系统发育树。

进一步通过 Phytozome v13 (Goodstein *et al.*, 2012) 在线网站查找不同物种中 *QaGDU3* 同源序列，利用 MEGA v7 软件构建最大似然 (ML) 系统发育树，参数设置为 Bootstrap method 和 500 bootstrap，探究 *QaGDU3* 与其它物种同源蛋白在进化上的关系。

3.2.4 *QaGDU3* 基因遗传转化拟南芥与鉴定

(1) 拟南芥种植及花序轴浸染法进行遗传转化

使用放置于 4°C 冰箱的野生型拟南芥 (WT) 种子, 按土: 蛭石: 珍珠岩为 3:1:1 的比例配置营养土, 装盆播种拟南芥, 待拟南芥长出两片子叶后, 将其移栽至每个盆的四角处, 待拟南芥抽薹开花后进行花序侵染试验。

首先挑取含目的基因重组质粒的单克隆阳性根癌农杆菌 EHA105 菌落至 5mL 含 50mg/L Kana 和 50mg/L Rif 的 LB 液体培养基中, 28°C, 220rpm 振荡培养 12h, 并与终浓度 25% 的甘油混合后保存于 -70°C 冰箱存菌。然后取活化后的 2mL 菌液至于 200mL 含 50mg/L Kana 和 50mg/L Rif 的 LB 液体培养基中扩大培养, 28°C, 220rpm 振荡 12h 后转入 50mL 离心管中, 25°C, 3000r/min 离心 15min, 离心后收集沉淀。沉淀使用农杆菌悬浮液 (0.44gMS, 5g 蔗糖, 5 μ L BREAKTHRU®S 233 (Evonik)) 定容至 50mL 重悬。将待转化的拟南芥植株 (36 株) 剪去已经开过的花与籽粒荚, 留下花蕾平放, 将其花蕾部分浸入含有转化液的 50mL 离心管中侵染 30s, 水平置于 22°C 黑暗条件下培养 24h 后, 将植株正立, 并置于正常光照条件下继续培养, 待种子成熟后, 收集转基因拟南芥种子。

(2) *QaGDU3* 转基因拟南芥植株的筛选和鉴定

将存于 4°C 的 T0 代的种子取出, 使用配制的消毒液 (10% CLOROX 和 0.05% 吐温) 消毒 15min, 无菌水冲洗 3-5 次后转移至在含 25mg/L 潮霉素的 1/2MS 固体平板培养基上, 放入 25°C, 14h 光照/10h 黑暗的温室中培养, 筛选得到转基因植株 (T1 代), 待其幼苗长出两片真叶时, 将其单株移植到土壤基质中, 置于温度 22°C 和相对湿度 80% 的培养条件下培养。待 1 月龄时使用改良的 CTAB 法提取叶片 DNA, 具体操作步骤如图 3.4:

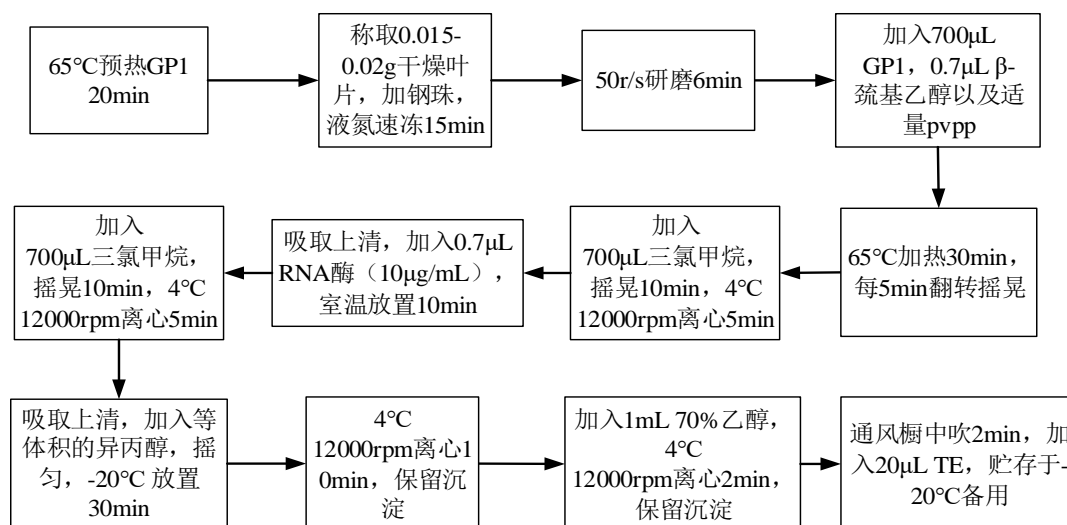


图 3.4 改良的 CTAB 法提取植物叶片 DNA 步骤

Figure 3.4 Extraction of DNA from plant leaves process by modified CTAB method

以 Sca495-1F/R 为引物（附录 A）对提取的 DNA 进行 PCR 检测，鉴定是否为阳性植株并进行标记，观察 T1 代植株生长特征。

3.2.5 荧光定量分析

(1) cDNA 样品准备

使用总 RNA 提取试剂 TRIGene 提取阳性植株叶片的 RNA，具体步骤如图 3.5:

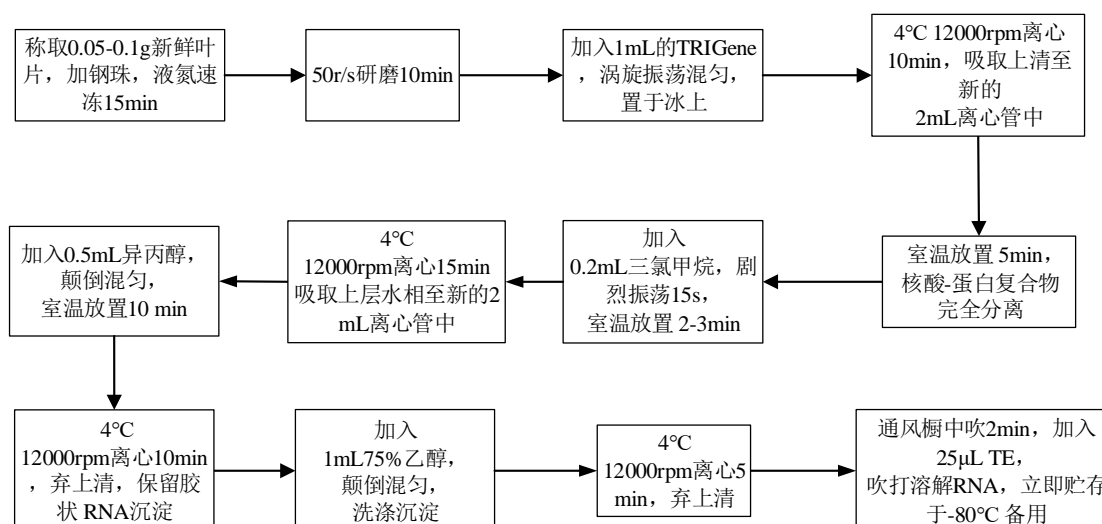


图 3.5 TRIGene 提取植物叶片 RNA 步骤

Figure 3.5 Extraction of plant leaves RNA process by TRIGene

使用 NanoDrop 2000C 超微量分光光度计测量 RNA 样品的浓度 (ng/μL) 及 OD260/OD280 值, 并使用 BeyoRT™ II cDNA 合成试剂盒 (with gDNA Eraser) 合成 cDNA 模板, 具体步骤如下:

根据公式计算每个样品中所需 RNA 模板体积:

$$\text{RNA 模板体积 } (\mu\text{L}) = 1000 / \text{RNA 样品浓度 } (\text{ng}/\mu\text{L}) \quad (2-1)$$

将模板 Total RNA, 5X gDNA Eraser Buffer 在冰上解冻, 5X RT Buffer、10X RT Primer Mix、DEPC-treated Water 在室温解冻, 将每种溶液混匀, 在高速离心机中 4°C 3000rpm 条件下离心 30s 以使所有液体沉降于管底, 后置于冰上备用。

去除 RNA 样品中的基因组 DNA 反应体系:

5X gDNA Eraser Buffer	2μL
RNA 模板	计算得到
DEPC-treated Water	To 10μL
总体积	10μL

在 PCR 仪上, 37°C 孵育 2min。迅速置于冰上放置, 直接加入反转录反应体系:

5X RT Buffer	4μL
10X RT Primer Mix	2μL
BeyoRT™ II M-MLV 反转录酶 (RNase H -)	2μL
DEPC-treated water	2μL
总体积 (2X Master Mix)	10μL

充分混匀后, 放入 PCR 仪中 42°C 孵育 60min, 随后 80°C 孵育 10min, 置于 -20°C 保存备用。

(2) 引物设计

利用 Beacon Designer 软件设计定量引物 AtUBQ5q、Sca495-rt、PR1-RT、ACD6-RT、PCC1-RT、WAK1-RT、ALD-RT、FMO1-RT、CBP60g-RT、SARD1-RT、SID2-RT、PAD4-RT 和 EDS1-RT (附录 A), 引物设计选择 SYBR Green Design, 扩增片段长度在 100~150bp 之间。

(3) 荧光定量 PCR 反应

应用 qTOWER 2.0 实时荧光定量 PCR 仪进行荧光定量 (qRT-PCR) 试验, 反应体系参照 BeyoFast™ SYBR Green qPCR Mix (2X) 说明书, 反应程序为: 预变性 95°C 2min, 变性 95°C 15s, 退火 60°C 30s, 72°C 延伸 30s, 40 次循环, 每个样品进行 3 次技术重复。以 *AtUBQ5* (ubiquitin) 为内参基因 (Thieme *et al.*, 2015), 分别对 *QaGDU3* 基因及 SA 途径中相关基因: *AtPR1*、*AtACD6*、*AtPCC1*、*AtWAK1*、*AtALD*、*AtFMO1*、*AtCBP60g*、*AtSARD1*、*AtSID2*、*AtPAD4* 和 *AtEDS1* 的表达量进行测定, 使

用 $2^{-\Delta\Delta Ct}$ 法计算每个样品目的基因的相对表达量 (Livak and Schmittgen, 2001), 同时计算 $2^{-(\Delta\Delta Ct-SE)}$ 和 $2^{-(\Delta\Delta Ct+SE)}$ 值作为误差线的上下限, SE 为 $\Delta\Delta Ct$ 的标准误。

3.2.6 *QaGDU3* 转基因拟南芥植株莲座大小测定

待转基因 T1 代植株成熟后, 单株系收集 33 株阳性 *QaGDU3* 转基因拟南芥种子并做好标记, 放入 37°C 烘箱干燥 2d 后, 加干燥剂至于 4°C 冰箱保存。按照 *QaGDU3* 基因表达量测定结果, 选择 *QaGDU3* 基因表达量高、中和低三种类型的植株进行莲座大小测定。取每个类型适量的种子, 使用配制的消毒液 (10% CLOROX 和 0.05% 吐温) 消毒 15min, 无菌水冲洗 3-5 次后转移至在含 25mg/L 潮霉素的 1/2MS 固体平板培养基上筛选。待其长至 17d 时, 每个株系分别随机选取 12 株阳性植株进行拍照, 并使用 Image J v1.53 软件 (Abràmoff *et al.*, 2004) 中加载的 RosetteTracker 程序包 (De Vylder *et al.*, 2012) 分别对每个植株的莲座直径进行测量, 并使用 SPASS v22.0 (Nie *et al.*, 1975) 进行显著性统计分析。

3.2.7 *QaGDU3* 转基因拟南芥对镰刀菌根腐病抗性研究

(1) *QaGDU3* 转基因拟南芥根部对镰刀菌根腐病的抗性

对实验室分离获得的共计 13 株镰刀菌 (附录 C), 采用分生孢子悬浮液进行活体灌根接种, 探究 *QaGDU3* 转基因拟南芥根部对镰刀菌根腐病的抗性。取 10mL PDB 液体培养基至于 50mL 离心管中并加入 10mL 无菌水稀释 2 倍, 夹取 PDA 培养基上的镰刀菌菌丝放入稀释后的 PDB 液体培养基中, 28°C, 220rpm 振荡培养 5d 后, 用移液枪吸取微量孢子悬浮液, 滴于灭菌载玻片上, 采用五点取样法, 镜检统计 5 个视野中孢子个数, 根据公式:

$$\text{孢子浓度} = \text{孢子总个数} / 25 \times 10^4 \quad (2-2)$$

计算孢子浓度。待分生孢子悬浮液浓度达 10^6 个/mL 时, 使用灭菌后带有滤布的锥形瓶分别对每个菌液进行过滤, 去除菌丝。对水琼脂培养基中的转基因 T2 代拟南芥根部灌根接种分生孢子悬浮液, 每株接种 10 μ L, 每个类型的 *QaGDU3* 转基因拟南芥重复 3 次 (3 株), 同时接种无菌水为空白对照, 放置于 25°C, 14h 光照/10h 黑暗的培养箱中培养, 观察转基因 T2 代及野生型拟南芥根部发病情况。同时为了评估不同类型的拟南芥对镰刀菌根腐病的抗病程度, 我们将疾病等级分别设置为 1 级 (0–25% 枯萎的叶子/根)、2 级 (25–50% 枯萎的叶子/根)、3 级 (50–75% 枯萎的叶子/根) 和 4 级 (75–100% 枯萎的叶子/根) 分别对转基因 T2 代及野生型拟南芥根部和叶片疾病程度进行统计 (Zhang *et al.*, 2016)。

(2) *QaGDU3* 转基因拟南芥叶片对镰刀菌根腐病的抗性

待土壤基质中的拟南芥幼苗生长至一月龄时，选择 *QaGDU3* 基因表达量高和野生型两种类型的株系，分别随机剪下 90 片叶片，放置在铺有湿滤纸的培养皿中。选择 *QaGDU3* 转基因拟南芥根部对镰刀菌根腐病抗性表型较为明显的菌株，将过滤后去除菌丝的菌液至于 1.5mL 离心管中，在高速离心机中 4°C 12 000rpm 条件下离心 5 min，使用注射器吸取上清，微孔滤膜过滤孢子，将滤液至于新的 1.5mL 离心管中，对滤纸上的转基因 T2 代和野生型拟南芥叶片分别接种菌液，每株接种 10 μ L，重复 15 次（15 片叶片），同时接种无菌水和 PDB 液体培养基为空白对照，放置于 25°C，14h 光照/10h 黑暗的培养箱中培养，观察转基因 T2 代和野生型拟南芥叶片发病情况（Thatcher *et al.*, 2009）。

(3) 植物材料的筛选

选择 *QaGDU3* 基因表达量高、中和低三种类型的植株，取每种类型适量的种子消毒后转移至含 25mg/L 潮霉素的 1/2MS 固体平板培养基上，同时设置野生型 WT 株系作为对照，待其生长出两片子叶时，每种类型的株系分别选取 4-6 株幼苗，转移至土壤基质中，每盆分别移栽 2 株转基因和 2 株野生型拟南芥植株进行对照，并至于 25°C 温室中进行培养；待 1/2MS 固体平板培养基上的幼苗生长至半月龄时，按表达量分别转移至水琼脂培养基中，同时转移等量的野生型拟南芥作为对照，进行标记。

3.3 研究结果

3.3.1 *QaGDU3* 基因的克隆及序列分析结果

以川滇高山栎 DNA 序列为模板，通过 PCR 扩增，结合 1% 琼脂糖凝胶电泳检测结果显示：在 Marker 大约 500bp 至 700bp 之间有 1 条清晰明亮的特异条带，无拖尾，与预期大小相同，推测该条带即为目的片段（图 3.6）。将此条带回收，通过 LIC 技术将此目的片段克隆至双元表达载体 pJG081，筛选阳性克隆进行测序，通过序列比对结果显示扩增序列与 *QaGDU3* 基因编码序列一致，*QaGDU3* 基因 cDNA 全长为 519bp，编码 172 个氨基酸（图 3.7）。将构建好的 pJG081-*QaGDU3* 质粒转入农杆菌 EHA105，筛选获得携带有重组质粒的菌落，用于后续侵染拟南芥实验。

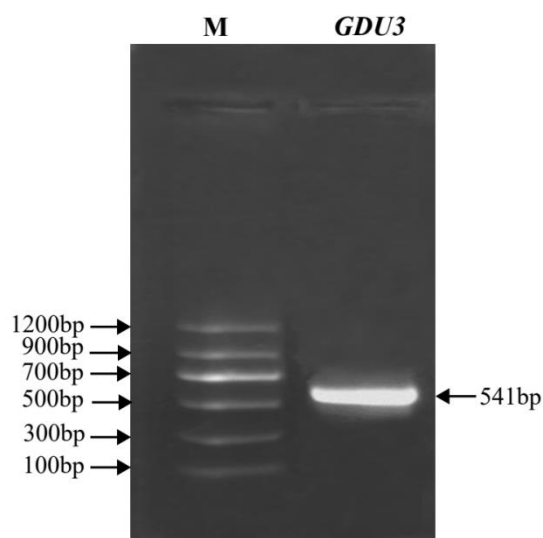


图 3.6 *QaGDU3* 基因扩增结果；M: DNA Marker II
Figure 3.6 The amplification result of *QaGDU3*; M: DNA Marker II

```

10      20      30      40      50      60
1      ATGGCCGCGAGCAAGAGAGGCCATTCAACATGAATGAAACAGCAAGATCTCCAACAGGGAGC
1      M A A A R E P F N M N E T A R S P T G S

70      80      90      100     110     120
61     ATGGGCACACACAACACTCACCATGGCATTCTCCTGTGCCATACTTGTGGTGGGCTG
21     M G T Q Q H S P W H S P V P Y L F G G L

130     140     150     160     170     180
121    GCAGCCATGTTGGGTCTCATTGCTTTTGCTCTTCTCATCCTTGCATGCTCCTACTGGAGA
41    A A M L G L I A F A L L I L A C S Y W R

190     200     210     220     230     240
181    CTCCTCTGGGTTCCCTTGACGGAGAAGGCAATTCAGAGAGAGACCTTGAGGCTGGCCAAGAT
61    L S G F L D G E G N S E R D L E A G Q D

250     260     270     280     290     300
241    GGTGAGGCCAAGCCTAATGATTCGCATAAACAACCACCAAGTGTGTTGAAGAGAAGTTCCTTG
81    G E A K P N D S H K Q P P V F E E K F L

310     320     330     340     350     360
301    GTGATTATGGCTGGTGAGGCTAAGCCAACGTATTTGGCTACCCCCATGTCAAGTAGGTCA
101   V I M A G E A K P T Y L A T P M S S R S

370     380     390     400     410     420
361    TCTTCTTTGGGTGATACTACTTGTAGCAGCTGTGGGAGCGAGAAGAGTGACAAGTCTGTG
121   S S F G D T T C S S C G S E K S D K S V

430     440     450     460     470     480
421    GAAATGACTGAGACAGAGACAGCCAAGCAGGCACAGATACCAGTACCACAACAAAGTGAC
141   E M T E T E T A K Q A Q I P V P Q Q S D

490     500     510
481    AACCAACGAGAGCGAGAGAGAGAGAGGCTCAAGCCAATGA
161   N N E S E R E R G S S Q *
    
```

图 3.7 *QaGDU3* 的 CDS 和氨基酸序列
Figure 3.7 The CDS and amino acid sequence of *QaGDU3*

3.3.2 *QaGDU3* 基因的生物信息学分析结果

根据 *QaGDU3* 氨基酸序列，采用 ExPASy ProtParam 软件预测得到 *QaGDU3* 蛋白

质的基本理化性质, 结果显示: *QaGDU3* 蛋白分子量为 18 583.56D, 理论等电点为 4.79, 推测该蛋白为酸性蛋白。预测半衰期为 30.00h, 不稳定指数为 65.67, 属于不稳定蛋白质; 其中丝氨酸 (Ser) 含量最高, 达 12.80%, 其次谷氨酸 (Glu) 和丙氨酸 (Ala) 含量均为 9.30%, 带负电残基数的总数 (丙氨酸 (Asp)+谷氨酸 (Glu)) 为 23 个, 带正电残基数的总数 (精氨酸 (Arg)+赖氨酸 (Lys)) 为 14 个, 疏水性平均值预测为-0.60, 表明 *QaGDU3* 蛋白为亲水性蛋白 (表 3.1)。ProtScale 预测结果整体显示负值多于正值, 因此推测 *QaGDU3* 蛋白为亲水蛋白; 此外, *QaGDU3* 蛋白第 50 位精氨酸 (Arg) 疏水性最强, 第 164 位丝氨酸 (Ser) 亲水性最强, 疏水性分值分别为 3.40 和-3.42 (图 3.8A)。进一步表明 *QaGDU3* 蛋白为亲水蛋白。综上, 推测 *QaGDU3* 蛋白可能是一个亲水不稳定酸性蛋白。

表 3.1 *QaGDU3* 蛋白信息
Table 3.1 The *QaGDU3* protein information

蛋白特性	数值
氨基酸数	172
相对分子质量/kD	18 583.56
理论等电点	4.79
不稳定系数	65.67
亲水性平均系数	-0.60

使用 SignalP v4.1 Server 信号肽预测结果发现 *QaGDU3* 蛋白 N 末端并没有信号肽 (图 3.8B)。TMHMM 跨膜区预测发现 *QaGDU3* 蛋白在约 33 到 55 个氨基酸处存在一个跨膜域, 其包含约 23 个氨基酸, 表明 *QaGDU3* 为跨膜蛋白 (图 3.8C)。通过 ProtComp9.0 亚细胞定位预测进一步发现 *QaGDU3* 蛋白存在于细胞质膜的概率最大为 9.48%; 而在其它细胞器的存在概率为 0.00%~0.21% (表 3.2)。综上, 说明 *QaGDU3* 蛋白很可能作为跨膜蛋白在植物中发挥作用。

表 3.2 *QaGDU3* 蛋白亚细胞定位预测
Table 3.2 Subcellular localization prediction of *QaGDU3*

基因在细胞中表达位置	存在概率 (%)
细胞质膜	9.48
线粒体	0.21
叶绿体	0.16
液泡	0.11
细胞质	0.00
内质网	0.00
细胞核	0.00

通过 PredictProtein 软件对蛋白质二级结构进行预测，结果显示 QaGDU3 蛋白具有 14 个蛋白质-蛋白质作用位点，二级结构由 26.16% 的 α -螺旋、2.33% 的延伸链和 71.51% 的无规卷曲构成， α -螺旋和无规卷曲是构成 QaGDU3 蛋白主要的结构元件（图 3.8D）。通过 Robetta 网站对三级结构预测，采用 RoseTTAFold 建模方法，三级结构模型置信度为 51.00%， α -螺旋和无规卷曲主要构成了 QaGDU3 蛋白的三级结构，与二级结构预测结果相似（图 3.8E）。

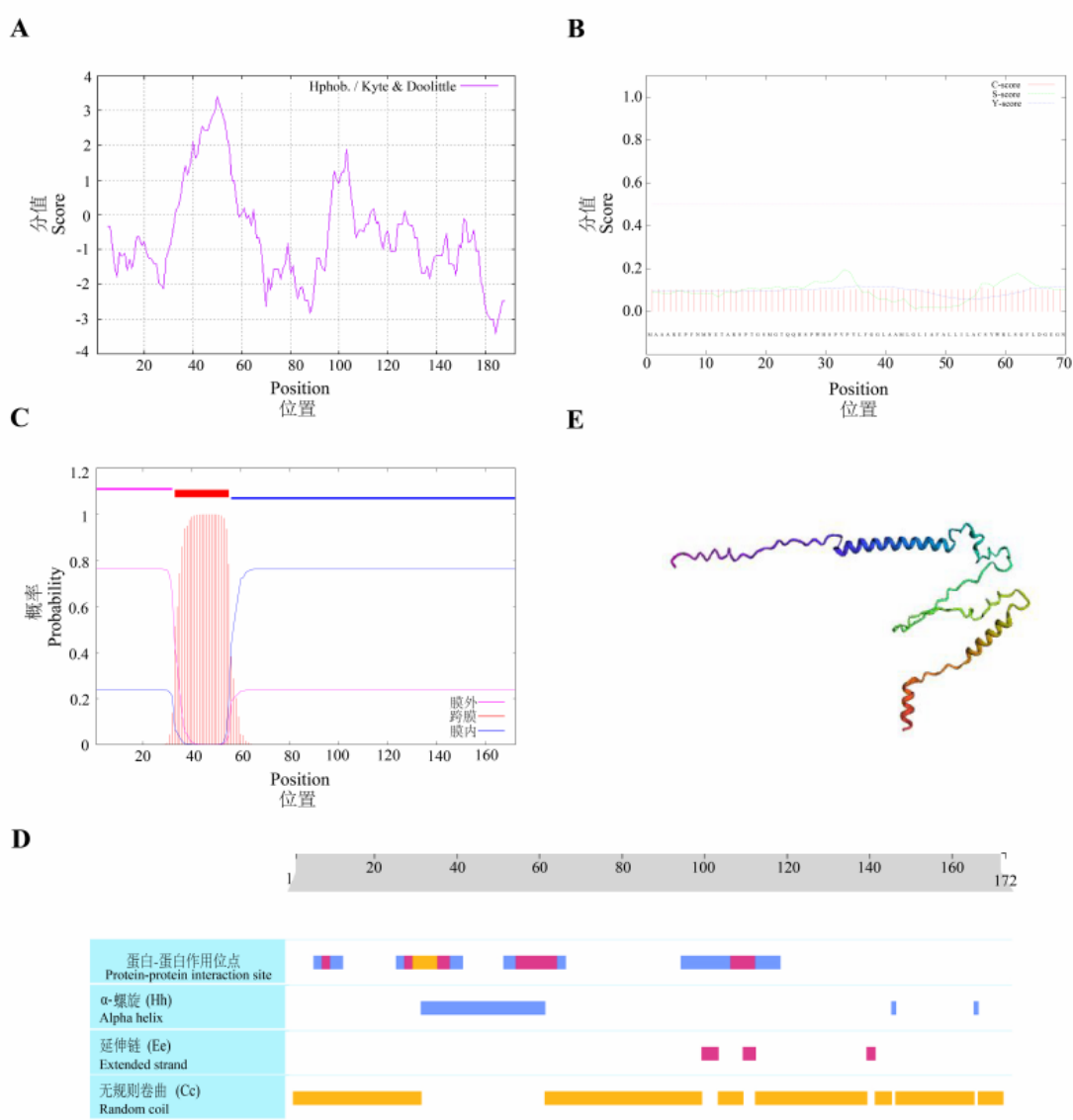


图 3.8 QaGDU3 蛋白基本性质及结构预测；QaGDU3 蛋白亲水性（A）、信号肽（B）、跨膜结构（C）、二级结构及蛋白作用位点（D）和三级结构（E）

Figure 3.8 Basic properties and structure prediction of QaGDU3 protein; QaGDU3 protein hydrophobicity and hydrophilicity (A), signal peptide (B), transmembrane structure (C), secondary structure and protein action site (D) and tertiary structure (E)

使用 NCBI 中 CD-search 分析 *QaGDU3* 蛋白保守结构域，结果显示 *QaGDU3* 蛋白属于阴离子渗透酶 ArsB/NhaD 超家族成员（图 3.9）。通过 MotifScan 软件预测蛋白的功能位点，结果显示 *QaGDU3* 蛋白具有 3 个天冬酰胺-连接糖基化位点（ASN-glycosylation site）、4 个 CK2-磷酸化位点（CK2-phospho site）、3 个肉豆蔻基位点（myristyl site）和 8 个 PKC-磷酸化位点（PKC-phospho site），氨基酸详细位置见表 3.3。

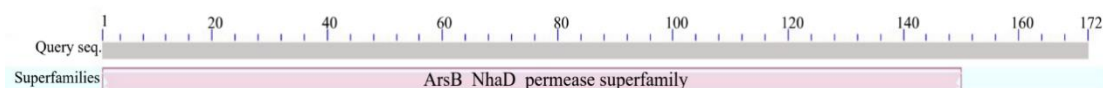


图 3.9 *QaGDU3* 蛋白保守域预测；ArsB_NhaD_permease superfamily：阴离子渗透酶 ArsB/NhaD 超家族

Figure 3.9 Prediction of the *QaGDU3* protein conserved domain; ArsB_NhaD_permease superfamily: Anion permease ArsB/NhaD superfamily

表 3.3 *QaGDU3* 蛋白功能位点预测

Table 3.3 Prediction of functional sites of *QaGDU3*

功能位点	氨基酸位置		氨基酸序列
	起始	终止	
天冬酰胺-连接糖基化位点	11	14	NETA
	86	89	NDSH
	162	165	ESER
酪蛋白激酶II磷酸化位点	71	74	ERDL
	122	125	FGDT
	143	146	ETET
	164	167	ERER
肉豆蔻基	19	24	SMGTQQ
	38	43	GLAAML
	132	137	SEKSDK
蛋白激酶 C 磷酸化位点	13	15	ARS
	71	73	ERD
	88	90	HKQ
	117	119	SRS
	133	135	EKS
	136	138	DKS
	147	149	AKQ
	164	166	ERE

3.3.3 *QaGDU3* 基因家族的进化结果

通过 DNAMAN v9.0 软件将 *QaGDU3* 氨基酸序列与拟南芥 *AtGDU3* (*A.thaliana*,

AT5G57685.1) 及近缘物种峡谷栎 QIGDU3 (*Q. lobata*, rna-XM_031080093.1)、欧洲栓皮栎 QsGDU3 (*Q. suber*, rna-XM_024054781.1) 和有柄栎 QrGDU3 (*Q. robur*, Qrob_T0462130.2) 的氨基酸序列进行多重比对, 结果显示它们的氨基酸序列平均一致度为 86.20%。使用 MEME 在线网站查找 motif 结果显示, 它们的氨基酸序列有 2 处保守的 motif1 和 motif2, 且具有 *GDU* 基因家族特有的“VIMAG”结构域, 其中四个栎属物种特有 1 处保守的 motif3 (图 3.10)。

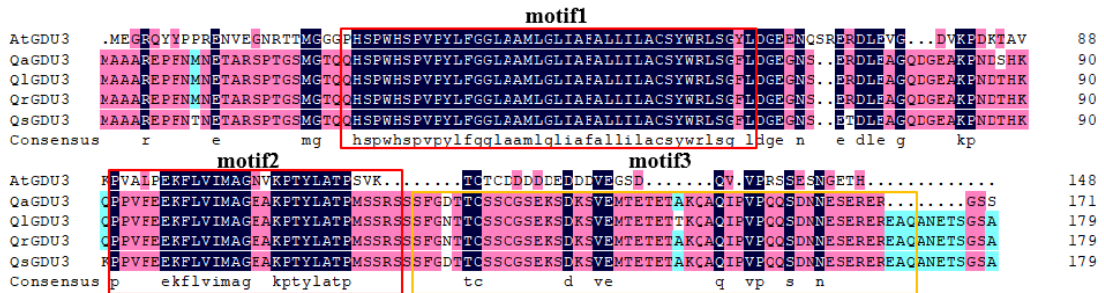


图 3.10 QaGDU3 与 QIGDU3, QrGDU3, QsGDU3 和 AtGDU3 蛋白的多序列比对及模体预测
Figure 3.10 Multiple sequence alignment and motif prediction of QaGDU3, QIGDU3, QrGDU3, QsGDU3 and AtGDU3

同时对四个栎属物种及拟南芥的 *GDU3* 蛋白质三级结构进行比较, 结果发现它们的三级结构基本相似, *QaGDU3*、*QIGDU3*、*QsGDU3* 和 *QrGDU3* 的蛋白三级结构均具有 3 个 α -螺旋, *AtGDU3* 的蛋白三级结构具有 2 个 α -螺旋。此外, 四个栎属物种的蛋白三级结构 N 末端均形成 α -螺旋, 而 *AtGDU3* 蛋白 N 末端未形成 α -螺旋 (图 3.11)。

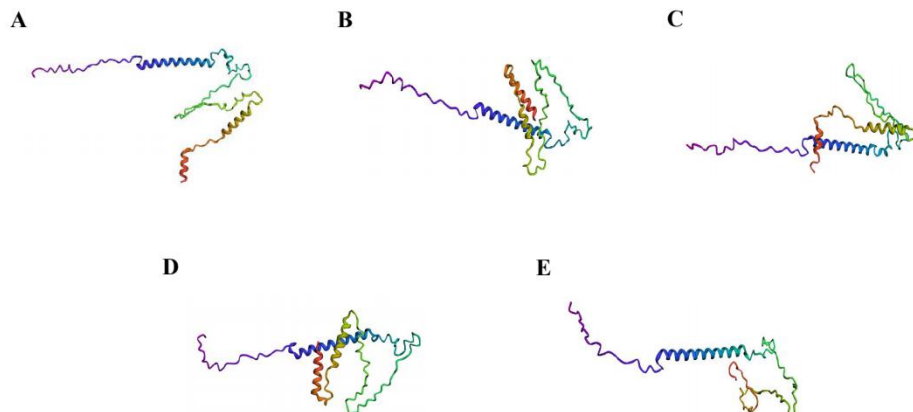


图 3.11 QaGDU3 (A) 与 QIGDU3 (B)、QrGDU3 (C)、QsGDU3 (D) 及 AtGDU3 (E) 蛋白的三位结构预测比较

Figure 3.11 Comparison of tertiary structure prediction of QaGDU3 (A), QIGDU3 (B), QrGDU3 (C), QsGDU3 (D) and AtGDU3 (E)

拟南芥和四个栎属植物的 GDU 家族系统发育分析发现，在四个栎属植物中都具有 GDU3、GDU5 和 GDU7，其中欧洲栓皮栎和加州白栎基因组中分别有两个 *AtGDU5* 的同源基因（图中以 *QsGDU-1*、*QsGDU-2*、*QlGDU-1*、*QlGDU-2* 表示）。在 GDU 家族中，GDU3 进化速率较快，GDU3 与 GDU2 的亲缘关系最近，与 GDU7 的亲缘关系最远。四个栎属植物和拟南芥的 GDU3 聚为一枝，且相较于 *AtGDU3*，四个栎属植物中的 GDU3 进化速率较慢（图 3.12）。

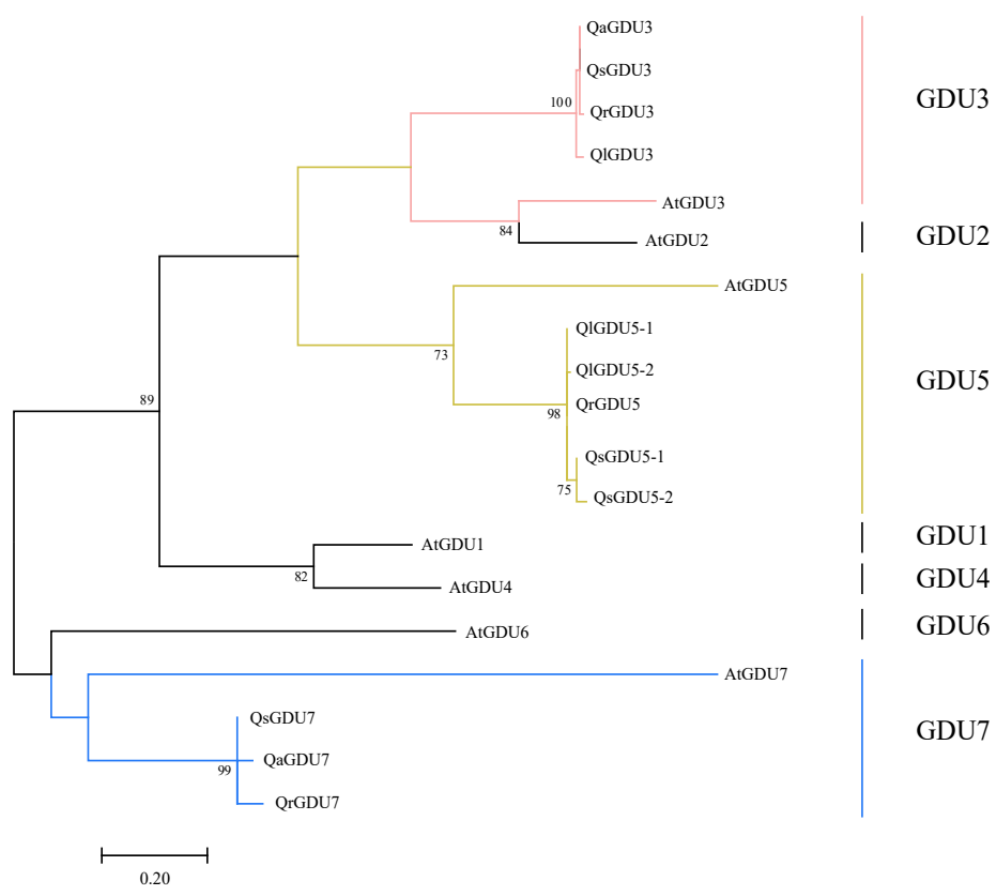


图 3.12 GDU 家族的系统发育分析

Figure 3.12 Phylogenetic analysis of GDU

注 a: 分支长度表示进化率，节点处的数字表示支持率，分支颜色粉色：GDU3；黄色：GDU5；蓝色：GDU7。

为了进一步研究 *QaGDU3* 蛋白与其它物种在进化上的亲缘关系，使用 phytozome v13 网站将 *QaGDU3* 的氨基酸序列与不同物种的同源序列进行比对，发现 GDU3 最早在种子植物中出现。选取种子植物中不同科且比对结果一致性较高的 24 个物种，包括裸子植物：西部红柏 (*Thuja plicata*) 和云杉 (*Picea sitchensis*)，单子叶植物：二穗短柄草 (*Brachypodium distachyon*)、玉米 (*Zea mays*)、水稻 (*Oryza sativa*) 和高粱 (*Sorghum bicolor*) 及双子叶植物：拟南芥 (*Arabidopsis thaliana*)、葡萄 (*Vitis vinifera*)、毛果杨 (*Populus trichocarpa*)、无油樟

(*Amborella trichopoda*)、蓝花耧斗菜 (*Aquilegia coerulea*)、番茄 (*Solanum lycopersicum*)、大豆 (*Glycine max*)、柠檬桉 (*Corymbia citriodora*)、菘蓝 (*Isatis tinctoria*)、陆地棉 (*Gossypium hirsutum*)、柑橘 (*Citrus clementina*)、豇豆 (*Vigna unguiculata*)、可可 (*Theobroma cacao*)、美国山核桃 (*Carya illinoensis*)、美洲板栗 (*Castanea dentate*)、欧洲栓皮栎 (*Quercus suber*)、峡谷栎 (*Quercus lobata*) 和有柄栎 (*Quercus robur*) 的 *QaGDU3* 的同源序列，使用 MEGA v7 软件中最大似然法 (ML) 构建系统进化树。结果表现出 *GDU3* 蛋白在裸子植物和被子植物中可明显区分，在被子植物中的单子叶和双子叶植物中也可明显区分。*QaGDU3* 与其近缘物种 *QlGDU3*、*QsGDU3* 和 *QrGDU3* 聚在一支，且进化时间较为一致。其中 *QaGDU3* 与 *QsGDU3* 亲缘关系最近，与云杉 (*ABK21085.1*) 的亲缘关系最远 (图 3.13)。

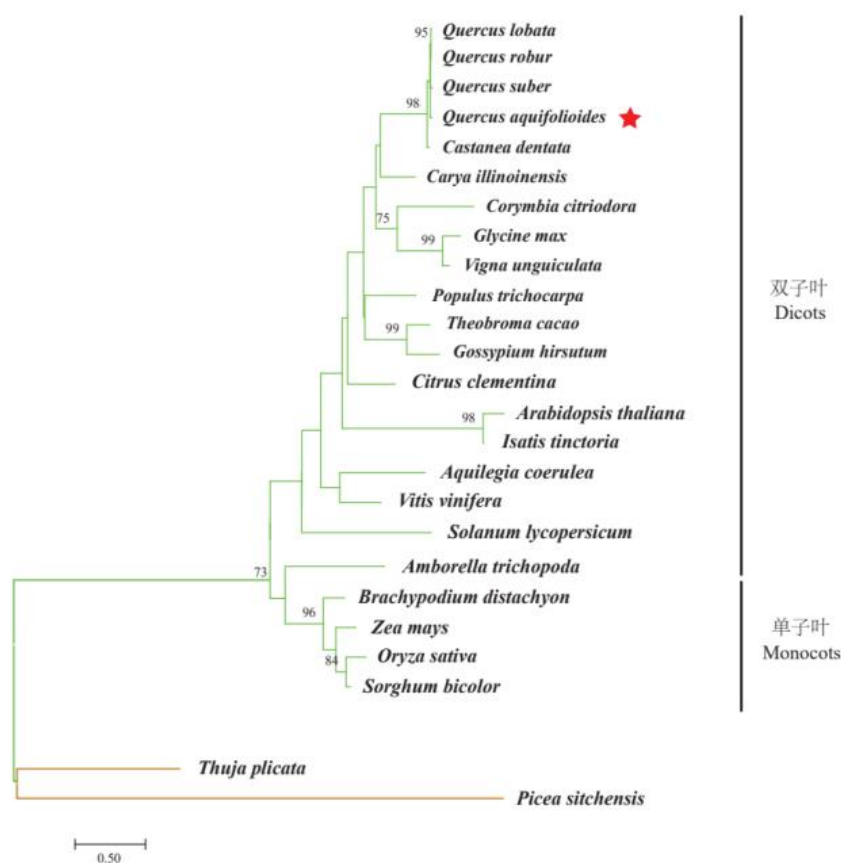


图 3.13 *QaGDU3* 蛋白的系统发育分析

Figure 3.13 Phylogenetic analysis of *QaGDU3*

注 a: 分支长度表示进化率，节点处的数字表示支持率，分支颜色棕色：裸子植物；绿色：被子植物。

3.3.4 *QaGDU3* 遗传转化拟南芥鉴定

收获经花序侵染法遗传转化的拟南芥 T0 代株系种子后，通过含潮霉素的 1/2MS 固体培养基进行抗性筛选，总共获得 33 株 T1 代 *QaGDU3* 转基因拟南芥抗性幼苗，其阳性表达率为 0.38%（图 3.14A）。将 T1 代 *QaGDU3* 转基因拟南芥抗性幼苗培养至 1 月龄时，观察其生长情况发现：与同时期的野生型（WT）拟南芥相比，T1 代拟南芥叶片卷曲且呈深绿色，大小明显小于野生型植株（图 3.14B）。通过改良的 CTAB 方法提取 T1 代 33 株拟南芥的叶片 DNA，进行 PCR 检测以鉴定 *QaGDU3* 转基因拟南芥中的阳性植株，琼脂糖凝胶电泳结果显示：所检测的 33 株 T1 代拟南芥均在大约 500bp 至 700bp 之间有 1 条清晰明亮的特异条带，无拖尾，均含有目的基因，说明 *QaGDU3* 基因已经成功转入拟南芥植株（图 3.14C）。

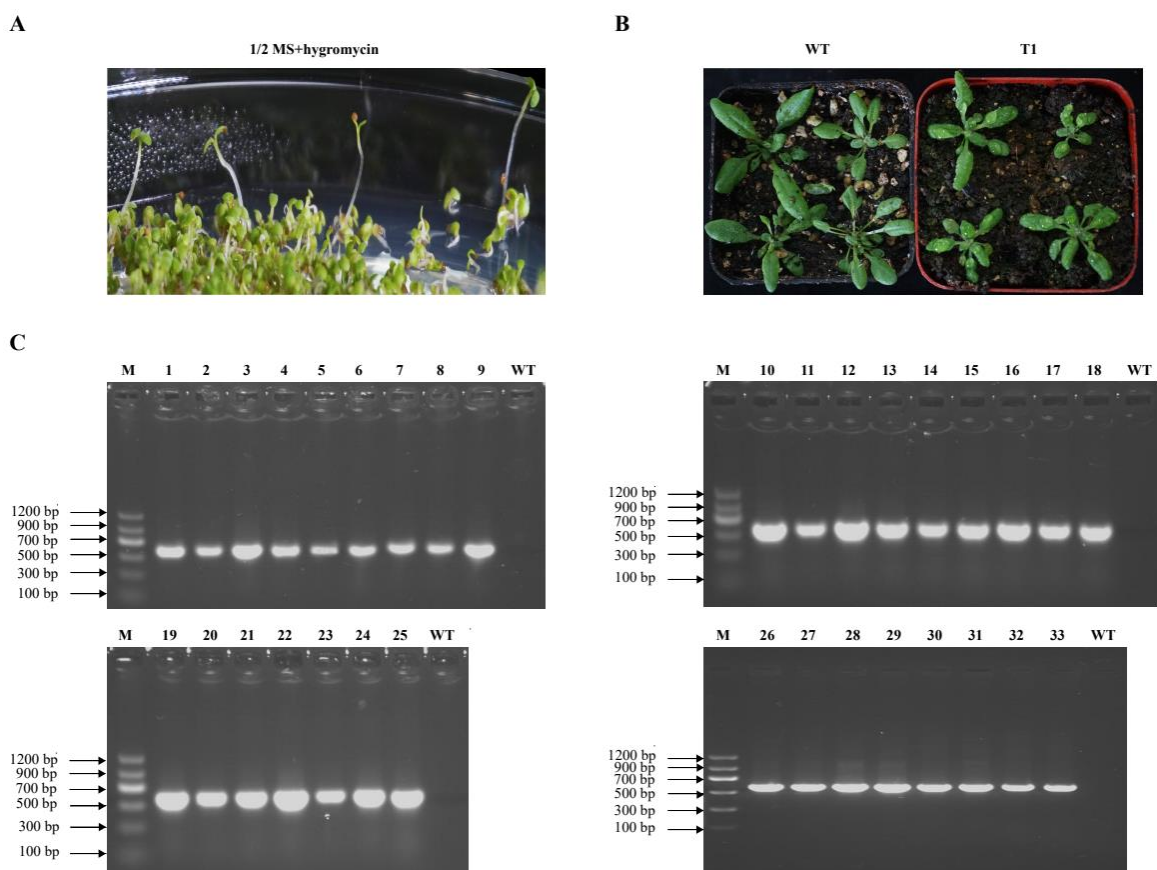


图 3.14 T1 代转基因拟南芥幼苗 PCR 鉴定

Figure 3.14 PCR verification of T1 generation transgenic *Arabidopsis thaliana* seedlings

3.3.5 *QaGDU3* 基因表达量与莲座大小关系

以拟南芥 *AtUBQ5* 基因作为内参基因，对转基因拟南芥的 *QaGDU3* 基因表达水

平进行定量分析。结果表明：3号转基因拟南芥株系表达量最高，2号转基因拟南芥株系表达量最低（图 3.15）。

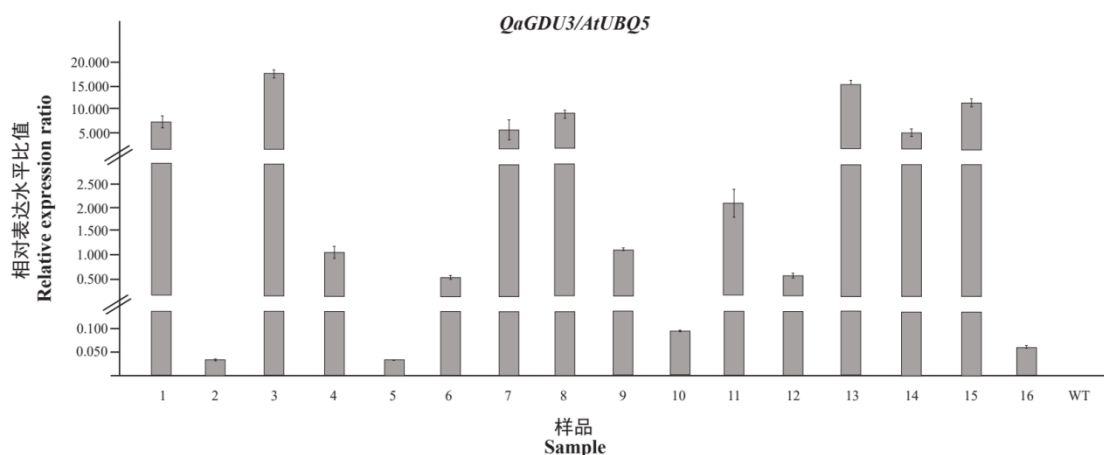


图 3.15 *QaGDU3* 与 *AtUBQ5* 基因的相对表达量比值
Figure 3.15 Relative expression ratio of *QaGDU3* to *AtUBQ5* gene

根据 *QaGDU3* 基因表达量测定结果，选取转基因拟南芥表达量高（编号为 3）、中（编号为 4）、低（编号为 2）的株系，观察并统计其 T2 代 *QaGDU3* 转基因拟南芥莲座大小（n=12/株系）。根据测量结果发现，2 号转基因拟南芥株系莲座叶直径最大，3 号转基因拟南芥株系莲座叶直径最小，且莲座叶直径随着 *QaGDU3* 基因表达量升高而变小（图 3.16）。

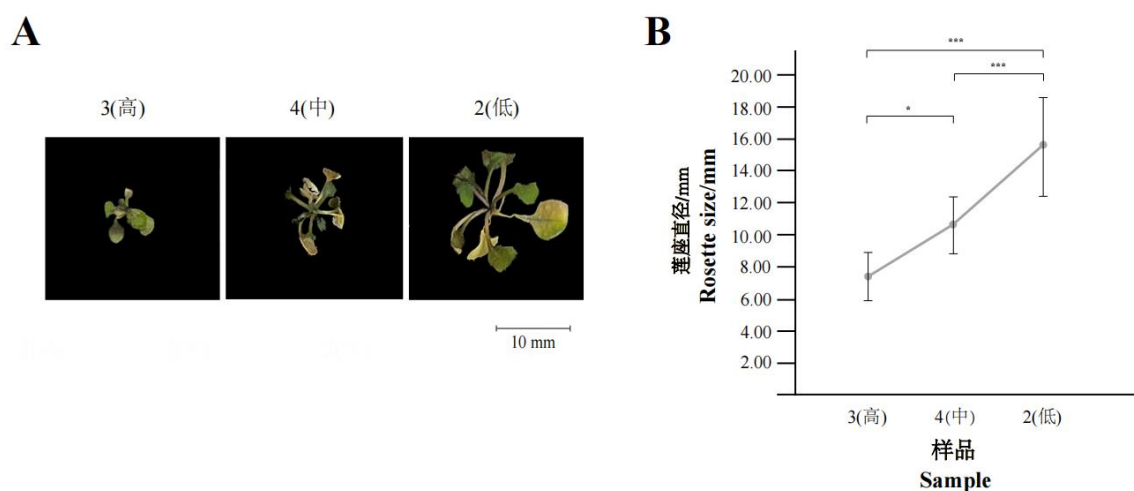


图 3.16 在含潮霉素的 1/2 MS 筛选培养基上生长 17 天的 T2 代转基因拟南芥幼苗 (A) 及其莲座直径统计 (B)

Figure 3.16 T2 generation transgenic *Arabidopsis thaliana* seedlings screened on 1/2 MS medium with hygromycin for 17 days (A) and their rosette size statistics (B)

注 a: *表示差异显著 (P<0.05); ***表示差异极显著 (P<0.001)。

3.3.6 水杨酸途径中相关基因表达量测定

选取转基因拟南芥 *QaGDU3* 表达量高（编号为3）、中（编号为4）、低（编号为2）和野生型（WT）拟南芥，以 *AtUBQ5* 基因作为内参基因，对转基因和野生型拟南芥中与 SA 途径相关基因 *AtALD*、*AtFMO1*、*AtCBP60g*、*AtPRI*、*AtPCC1*、*AtSARD1*、*AtSID2*、*AtPAD4*、*AtEDS1*、*AtACD6* 和 *AtWAK1* 的表达水平进行定量分析。结果表明，除 *AtEDS1* 基因外，三个株系中 *QaGDU3* 表达量最低的 2 号株系与 SA 途径相关基因的表达水平与野生型拟南芥相当。随 *QaGDU3* 表达量升高，转基因拟南芥中 *AtALD*、*AtFMO1*、*AtCBP60g*、*AtPRI*、*AtPCC1*、*AtSARD1*、*AtSID2*、*AtPAD4*、*AtEDS1*、*AtACD6* 和 *AtWAK1* 的表达水平均呈现升高趋势。这些结果显示 *QaGDU3* 的过量表达会进一步激活 SA 途径相关基因的表达（图 3.17）。

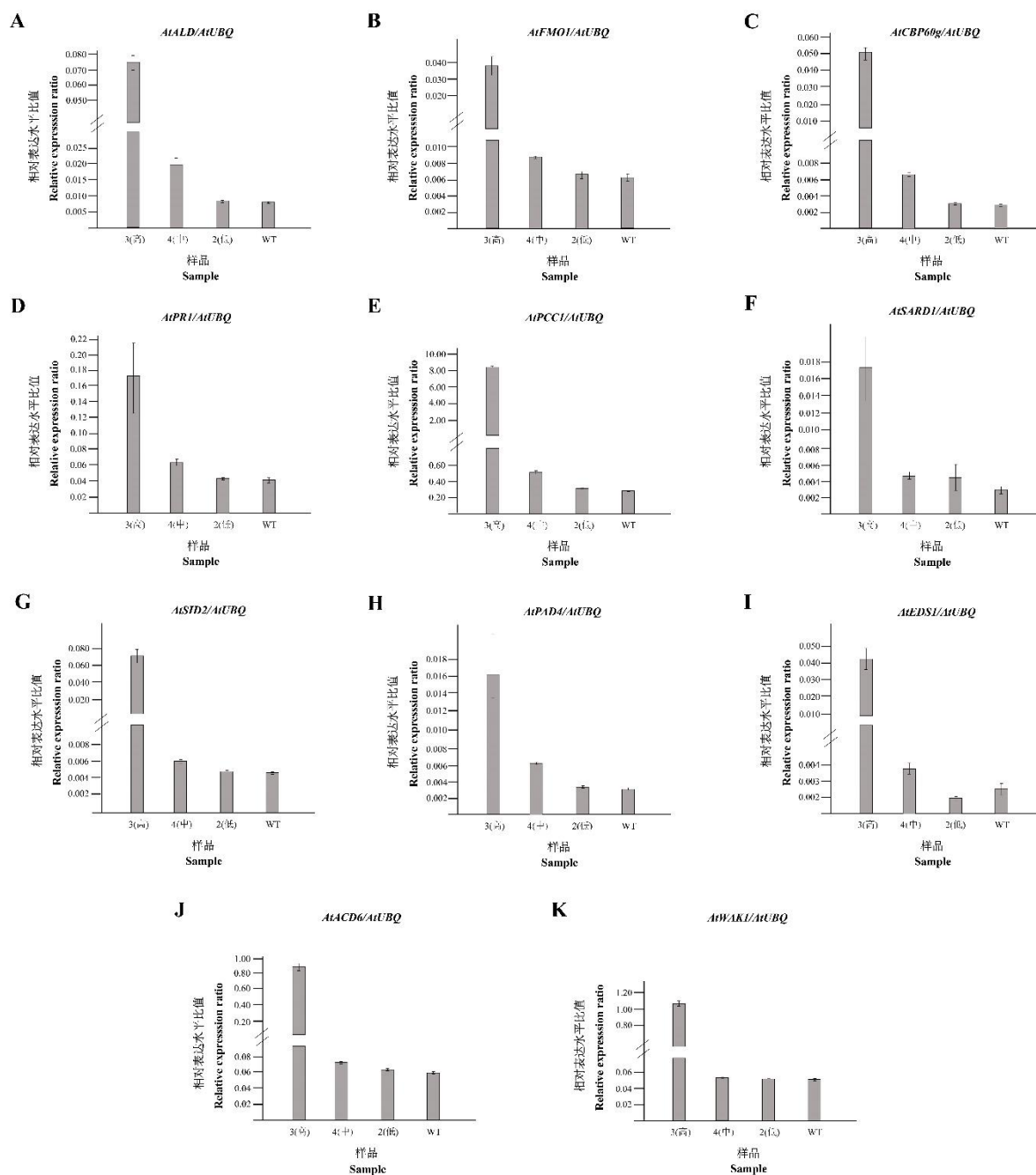


图 3.17 *AtALD* (A)、*AtFMO1* (B)、*AtCBP60g* (C)、*AtPRI* (D)、*AtPCC1* (E)、*AtSARD1* (F)、*AtSID2* (G)、*AtPAD4* (H)、*AtEDS1* (I)、*AtACD6* (J) 和 *AtWAK1* (K) 的相对表达量比值

Figure 3.17 Relative expression ratio of *AtALD* (A)、*AtFMO1* (B)、*AtCBP60g* (C)、*AtPRI* (D)、*AtPCC1* (E)、*AtSARD1* (F)、*AtSID2* (G)、*AtPAD4* (H)、*AtEDS1* (I)、*AtACD6* (J) and *AtWAK1* (K) to *AtUBQ5* gene

3.3.7 *QaGDU3* 基因抗镰刀菌根腐病研究结果

本研究共使用 13 株具有致病力的镰刀菌对 *QaGDU3* 转基因拟南芥进行抗性实

验，其中包括：1个尖孢镰刀菌（*Fusarium oxysporum*）、5个层生镰刀菌（*Fusarium proliferatum*）、1个锐顶镰刀菌（*Fusarium acuminatum*）、2个甜瓜镰刀菌（*Fusarium roseum*）、3个茄病镰刀菌（*Fusarium solani*）和1个芬芳镰刀菌（*Fusarium redolens*）（附录 C），采用分生孢子悬浮液灌根法研究 *QaGDU3* 转基因拟南芥对镰刀菌根腐病的抗性。*QaGDU3* 表达量高（编号为3）、中（编号为4）、低（编号为2）和野生型（WT）四种类型的拟南芥在接种 13 株菌株后表型如图 3.18 所示。观察发现，*QaGDU3* 转基因拟南芥对 2A 菌株抗病性最弱，接种第 1 周后即出现发病症状，导致植株叶部萎蔫，根部腐病，且病症与同时期野生型拟南芥相同；其次 *QaGDU3* 转基因拟南芥对 1A、1C、1D、2B、2C、3B 和 3D 菌株抗病性较弱，接种第 2 周出现发病症状，植株叶部出现轻微萎蔫，根部出现轻微腐病。*QaGDU3* 转基因拟南芥对菌株 1B、2D、2E、3A 和 3C 抗病性相对较强，接种第 2 周其植株叶部较为健康，根部无明显的腐病现象。

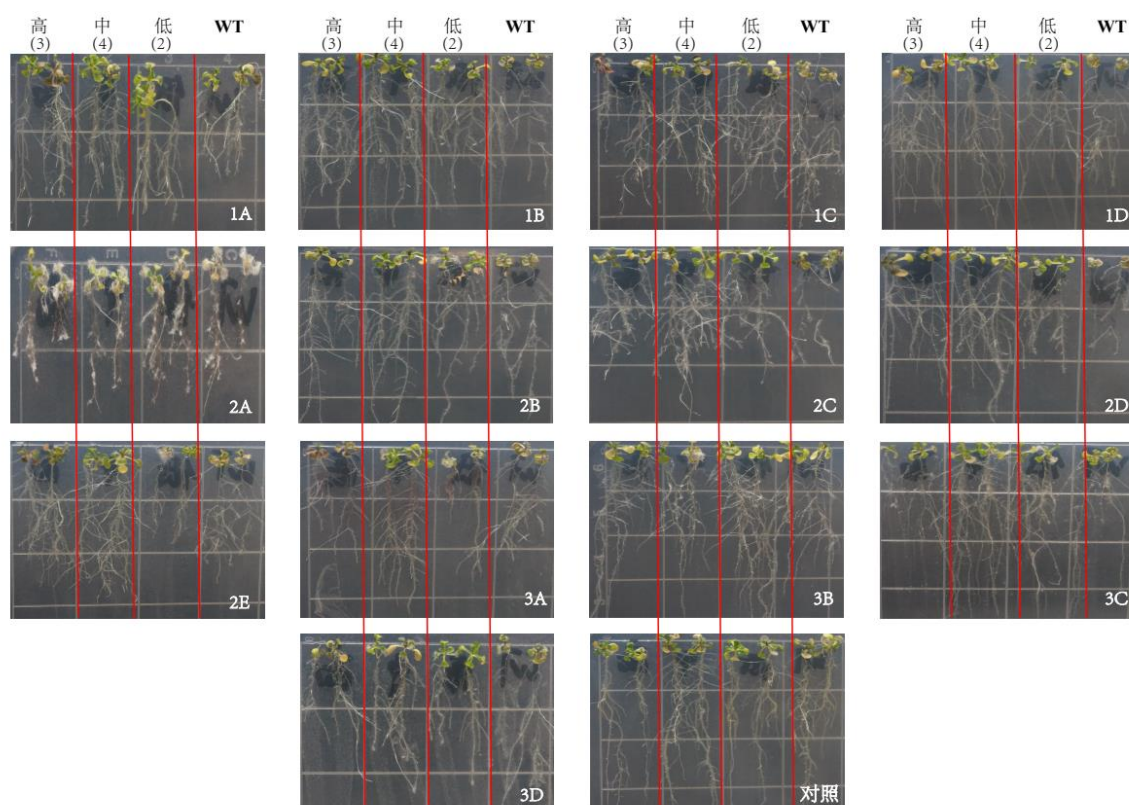


图 3.18 T2 代转基因拟南芥对 13 株镰刀菌菌株抗性检测

Figure 3.18 T2 generation *Arabidopsis thaliana* seedlings resistance tests to thirteen *Fusarium* strains

为了进一步探究 *QaGDU3* 转基因拟南芥对 1B、2D、2E、3A 和 3C 五个菌株抗病性，分别对 *QaGDU3* 转基因表达量高（编号为 3）和野生型（WT）两种类型的

拟南芥叶片进行侵染实验。在接种五种过滤的菌液 5 天时，野生型叶片接种部位均产生明显的褪绿和局部细胞死亡的现象，并在 7 天内发展为严重的褪绿和坏死，而 *QaGDU3* 转基因表达量高的叶片除了部分出现轻微的衰退变黄外，无明显的坏死现象（图 3.19）。接种菌株 1B、2D、2E、3A 和 3C 后野生型拟南芥叶片的坏死率分别为 73.33%、80.00%、60.00%、46.67% 和 53.33%，而对应的 *QaGDU3* 转基因拟南芥叶片坏死率分别为 20.00%、20.00%、46.67%、26.67% 和 26.67%，这些结果进一步证明了 *QaGDU3* 转基因拟南芥对茄病镰刀菌（*Fusarium solani*）、尖孢镰刀菌（*Fusarium oxysporum*）和层生镰刀菌（*Fusarium proliferatum*）具有一定的抗性。

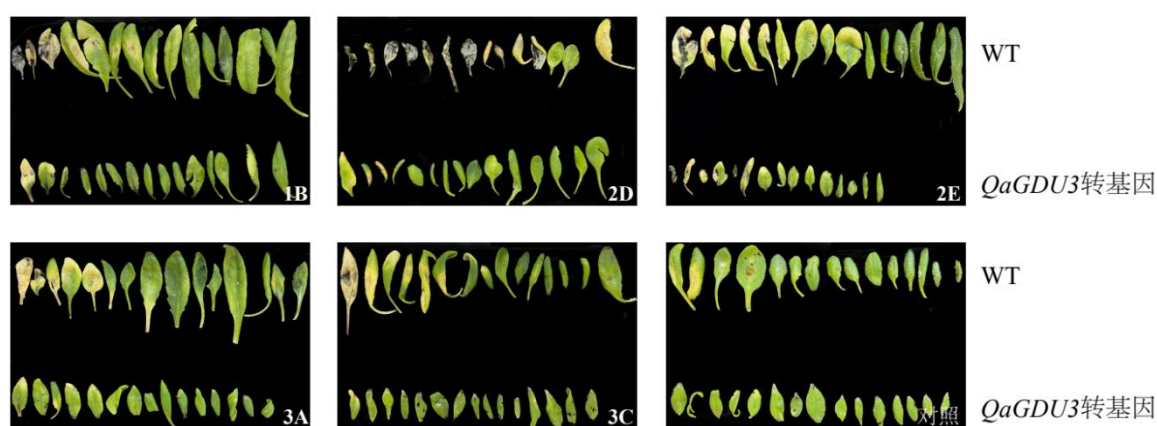


图 3.19 T2 代转基因拟南芥叶片对 5 株镰刀菌菌株抗性检测

Figure 3.19 T2 generation *Arabidopsis thaliana* leaves resistance tests to five *Fusarium* strain

接下来我们对接菌后不同类型 *QaGDU3* 转基因拟南芥中表型差异较大的 1B 菌进行重复实验，每种类型的拟南芥设置 20 个重复（20 株），分别基于叶片和根部发病情况统计疾病指数，进一步探究 *QaGDU3* 转基因拟南芥对镰刀菌根腐病抗性的影响。结果显示在接种 1B 菌液 2 周左右，四种类型的拟南芥根部和叶片分别不同程度出现了发病症状，相较于其它类型的拟南芥，表达量高（编号为 3）的 *QaGDU3* 转基因拟南芥株系发病症状程度较轻，植株叶部仅出现轻微萎蔫，根部仅出现轻微腐病（图 3.20）。进一步结合疾病指数统计发现随着 *QaGDU3* 基因的表达量升高，叶片和根部的疾病指数越低，即抗性越强（图 3.21）。

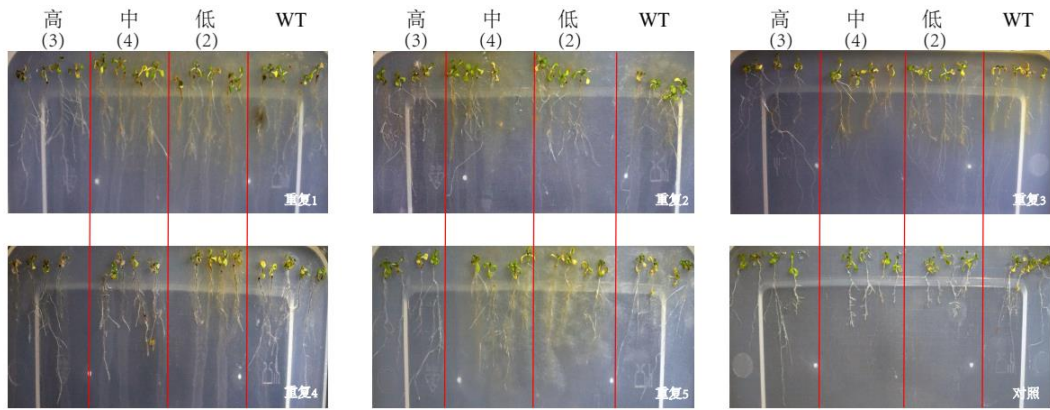


图 3.20 T2 代转基因拟南芥对茄病镰刀菌 (1B) 菌株抗性检测

Figure 3.20 T2 generation *Arabidopsis thaliana* seedlings resistance tests to *Fusarium solani* (1B)

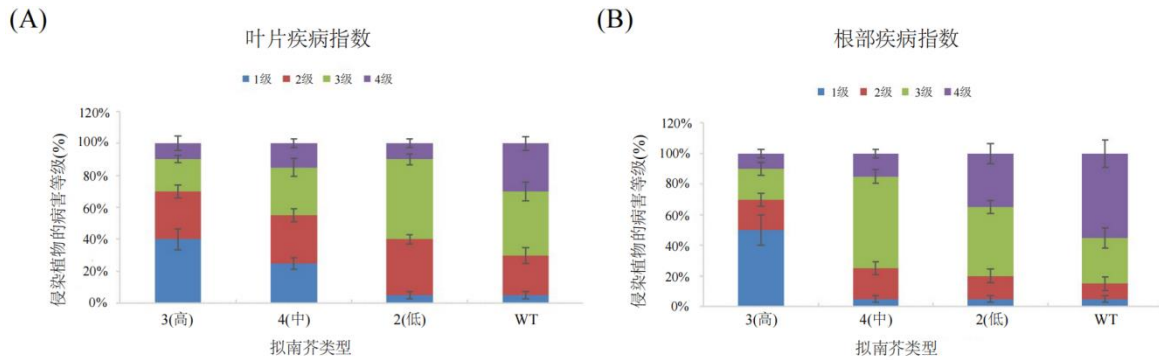


图 3.21 接种茄病镰刀菌 (1B) 后 T2 代转基因及野生型拟南芥叶片和根部的疾病等级

Figure 3.21 The disease grades were evaluated with leaves and roots of T2 generation and wild-type *Arabidopsis thaliana* seedlings infected with *Fusarium solani* (1B)

4. 讨论

4.1 川滇高山栎基因组特性

栎属植物是森林生态系统的重要组成部分，其基因组数据可以为该属植物的分类、起源和进化以及育种等方面的研究提供重要的遗传信息（胡花蕾等, 2019）。本研究基于前期实验室调查结果的基础之上基于二代 Illumina 和三代 Nanopore 的测序数据，结合 Hi-C 技术，组装得到了川滇高山栎染色体水平的高质量参考基因组，基因组大小约为 1.15Gb，contig N50 约为 1.20Mb，为组装得到的第一个染色体水平上的壳斗科栎属冬青栎组植物基因组，同时也是已测序栎属植物基因组中最大的（有柄栎：814.28Mb, Plomion *et al.*, 2018；欧洲栓皮栎：953.30Mb, Ramos *et al.*, 2018；加州白栎：846.07Mb, Sork *et al.*, 2016；蒙古栎：809.83Mb, Ai *et al.*, 2020），96.02%的组装基因组序列能够固定在 12 条染色体上，是栎属植物中组装至染色体水平上较高质量的参考基因组。

基因组的连续性、准确性和完整性为评估基因组组装质量的重要指标。川滇高山栎基因组 contig N50 大小约为 1.20Mb，与 contig N50 大小为 69.35kb 的有柄栎（Plomion *et al.*, 2018）、96.59kb 的欧洲栓皮栎（Ramos *et al.*, 2018）和 80.92kb 的加州白栎（Sork *et al.*, 2016）相比更大，这证明了川滇高山栎基因组较高的连续性。研究中使用 BUSCO 和 LAI 两个指标对基因组组装质量进行了评估。BUSCO 评估显示，在组装的川滇高山栎基因组中可以检测到大约 95.40%的完整基因，LAI 评分显示基因组符合可参考质量（Ou *et al.*, 2018）。总体来说，本研究组装得到了染色体水平上较高质量的川滇高山栎参考基因组，为栎属植物高杂合、高重复的基因组组装提供了重要参考。

在川滇高山栎基因组中注释得到了 26 441 个蛋白编码基因，相较于有柄栎、加州白栎和欧洲栓皮栎注释得到的蛋白编码基因数量较少，其它三种栎属植物与川滇高山栎蛋白质编码基因数量的差异可归因于基因组组装时使用的组装技术和基因组注释所使用的预测方法的差异（Ma *et al.*, 2021）。功能注释结果显示，川滇高山栎基因组中有 7.81%基因未注释到其功能，这些基因可能是新基因，可能参与调控植物体内关键的生物合成和次生代谢，与生物对其栖息地的适应性相关（Yang *et al.*, 2009）。基因组重复序列在一定程度上决定了基因组的大小以及复杂程度，川滇高山栎基因组中重复序列约为 666.10Mb，占基因组的 58.04%，比~359.40Mb 的有柄栎（Plomion *et al.*, 2018）和~123.50Mb 的加州白栎（Sork *et al.*, 2016）基因组中的重复序列含量都大。其中长末端片段重复序列是含量最丰富的重复类型，占整个基因组的 40.86%。川滇高山栎基因组重复序列含量较高，因此对于基因组的组装更具有挑战性，同时基因组中较高含量的重复序列可能也是川滇高山栎能够适应极端高山生

境的原因之一 (Guo *et al.*, 2018)。

综上所述, 我们组装得到了连续性和基因覆盖率较高的川滇高山栎基因组, 为进一步系统发育分析提供了基础, 也为其它栎属植物基因组的组装提供了重要参考。

4.2 川滇高山栎对高山环境的适应性进化

本研究基于 12 个木本植物和 2 个草本植物构建了全基因组水平的系统进化树, 其拓扑结构显示麻栎组的欧洲栓皮栎和冬青栎组的川滇高山栎亲缘关系较近, 且两个白栎组物种有柄栎和加州白栎聚在一支, 验证了之前的研究结果 (Hipp *et al.*, 2020)。同时基于叶绿体片段和核微卫星分子标记进行种群遗传学分析表明我们进行全基因组测序的这棵川滇高山栎属于横断山脉-川西高原谱系 (Du *et al.*, 2017), 考虑到川滇高山栎的分化时间与青藏高原从中新世晚期到上新世晚期最近一次快速抬升的时间较为相近 (Xing and Ree, 2017)。因此我们推测, 横断山脉-川西高原谱系的川滇高山栎的物种形成可能由喜马拉雅山的快速隆升引发的, 并导致川滇高山栎进化出能够适应极端环境特化的表型和生理性状 (Xing and Ree, 2017)。

在川滇高山栎基因组中鉴定得到 416 个正选择基因, 功能富集显示这些基因功能集中在 RNA 和受损 DNA 结合, 这些基因可以选择性的非共价作用于一个 RNA 或受损 DNA 分子或其中的一部分, 从而形成广泛的调节网络, 帮助维持细胞稳态 (Lukong *et al.*, 2008); 同时与 RNA 修饰相关的基因可以引起 RNA 分子内的一个或多个核苷酸的共价改变, 以产生具有不同于遗传编码序列的 RNA 分子。与光系统 II 氧复合物功能相关的基因能够与外在蛋白质结合锰、钙和氯形成离子簇, 可催化川滇高山栎植物体内的水分解为氧气和氢离子, 从而参与光系统 II 的调节, 影响川滇高山栎的光合作用等生理活动 (Kochhar *et al.*, 1996)。

与“质子跨膜转运 (Proton transmembrane transport)”相关的基因是植物体能够进行正常物质交换、维持植物新陈代谢等生命活动的重要保证, 也是生物膜进行能量转换和信息传递的基础 (Zheng *et al.*, 1999), 因此后续我们筛选相关的基因对其分子功能进行进一步验证; 除此之外参与核苷酸切除修复过程的正选择基因, 可以将受损 DNA 周围的一小部分作为寡核苷酸从 DNA 螺旋中移除, 而 DNA 螺旋中留下的小间隙由 DNA 聚合酶和 DNA 连接酶填充, 核苷酸切除修复可识别多种底物, 包括由紫外线照射和化合物引起的损伤 (Luo *et al.*, 2012), 这些基因可能是使川滇高山栎能够分布在其它树种很难适应的土壤贫瘠、寒冷干旱、降水量少、紫外辐射强等立地条件差的高山极端环境中的原因之一。

4.3 *QaGDU3* 基因特征

蛋白质由结构单元或结构域组成, 几乎可以独立于蛋白质的其余部分折叠并执行特定功能 (Orengo and Thornton, 2005)。其中跨膜蛋白是一类结构独特的蛋白,

它在植物细胞中广泛存在，并发挥重要的生理功能（Litou *et al.*, 2008）。植物体内的跨膜蛋白具有一些共同的理化性质。植物体内的跨膜区由强疏水的氨基酸组成，以保证跨膜蛋白能够穿过膜的磷脂双分子层。一般来说，跨膜蛋白包括 1~20 个跨膜区，每个跨膜区大约由 20 个氨基酸组成。在植物中，许多跨膜蛋白以特定的 α 螺旋或 β 折叠等结构嵌入膜的磷脂双分子层中并稳定的发挥其作用（宋江华和张立新, 2009）。“VIMAG”为 *AtGDU* 蛋白的一个特有的疏水功能的结构域，该保守的结构域可以与其它蛋白相互作用并构成一个完整的跨膜结构域，从而使 *AtGDU* 蛋白与位于膜上的蛋白质相互作用，并参与植物体内氨基酸转运的各种途径（Pratelli and Pilot, 2006）。本研究从川滇高山栎基因组中克隆获得 *QaGDU3* 基因，同时对该基因进行结构域分析，表明 *QaGDU3* 蛋白属于跨膜蛋白，具有保守的“VIMAG”蛋白结构域，主要由 α -螺旋和无规卷曲构成，能够稳定的附着在细胞膜两侧与位于膜上的蛋白质相互作用。进一步通过序列比对发现 *QaGDU3* 与其近缘物种 *QlGDU3*、*QrGDU3*、*QsGDU3* 的蛋白序列有三个共有的 motif，证明了 *GDU3* 蛋白序列在栎属植物中有较高的保守性。四个栎属植物的 *GDU3* 蛋白与 *AtGDU3* 蛋白 N 端的 24 个氨基酸序列差异导致了栎属植物的蛋白三维结构 N 末端均形成 α -螺旋，而 *AtGDU3* 蛋白 N 末端未形成 α -螺旋。在 *GDU* 家族中 *GDU3* 的进化速率较快且四个栎属植物的 *GDU3* 蛋白聚在一支，进化时间较为一致，说明其具有共同的起源；同时，*GDU3* 蛋白从种子植物中开始出现，在单子叶中的分化时间早于双子叶，与 Pratelli 和 Pilot（2006）的研究结论一致。

跨膜蛋白处于细胞与外界的交界处，能独立或者与其它相关的蛋白构成跨膜蛋白复合体，并执行各种生物学功能。在植物体中，跨膜蛋白的作用主要体现在以下三个方面：首先，植物体内的跨膜蛋白能够构成各种离子跨膜的通道，把维持植物生存所需的营养物质和无机电解质输入细胞，同时将有毒或没用的代谢产物排出细胞，从而影响细胞与外界之间物质运输；其次，植物体内的一些跨膜蛋白可以构成各种信号分子、激素和其它底物的受体，从而参与膜内外的物质交换、介导细胞与外界之间能量和信号的转导；除此之外，一些植物跨膜蛋白可构成呼吸链和转运蛋白，并参与植物细胞内的生理生化代谢（宋江华和张立新, 2009; 常秋霜, 2015）。本研究通过结构域功能分析表明 *QaGDU3* 蛋白属于阴离子渗透酶 *ArsB/NhaD* 超家族成员，可横跨生物膜转移钠、砷酸盐、铋酸盐、硫酸盐和有机阴离子；既能独立作为膜转运蛋白，也可作为 ATP 驱动的阴离子泵通道形成亚基发挥作用（Kuroda *et al.*, 1997）。

分泌是植物体内一个基本的生理过程，可帮助植物体处理溶质、提高营养获取和吸引其它生物。在植物的叶片中，通常由位于叶缘，接近导管的末端的水孔介导物质的输出。植物的吐水现象是指植物从未受伤的叶片尖端或边缘的水孔向外分泌含有离子、代谢物和蛋白质的汁液。水孔可以从质外体和木质部汁液中获取离子和

有机分子，并通过吐水现象主动分泌汁液（Komarnytsky *et al.*, 2000; Ingo *et al.*, 2003; Madsen *et al.*, 2016）；或者由于细胞渗透压变化导致汁液的被动流出。研究表明施肥的黑麦草和茶树会从水孔中分泌氨基酸（Greenhill and Chibnall, 1934; Yoneda *et al.*, 2019）；在拟南芥中，*AtGDUI* 基因表达量增强可诱导水孔分泌谷氨酰胺，且游离氨基酸含量随着 *AtGDUI* 基因过表达强度的增加而增加，同时植株大小随着 *AtGDUI* 基因的 mRNA 积累而减小（Pilot *et al.*, 2004）。本研究通过 qRT-PCR 分析验证了转基因植物 *QaGDUI3* 基因的表达量，同时结合转基因拟南芥莲座大小测量结果发现 *QaGDUI3* 转基因拟南芥 T2 代株系莲座叶直径随着 *QaGDUI3* 基因表达量升高而变小，此结果与 *AtGDUI3* 基因的过表达会导致植株变小的研究结果达成一致（Pratelli and Pilot, 2006），预示着 *QaGDUI3* 可能与 *AtGDUI3* 在氨基酸代谢中具有相似的作用。

4.4 *QaGDUI3* 基因对川滇高山栎适应性影响

水杨酸（SA）是一种重要的植物激素，也是一种植物合成的抗病关键信号分子，可激活病程相关基因的表达、建立局部和系统获得性抗性，从而对抗多种病原微生物的侵害。本研究通过 qRT-PCR 测定了 *QaGDUI3* 转基因拟南芥中与 SA 途径相关基因的表达，结果表明 *QaGDUI3* 的超量表达会进一步激活 SA 途径相关基因的表达。其中 *AtACD6* 的功能获得性突变导致 *AtEDS1*、*AtPAD4* 和 *AtPRI* 的表达增加，并诱导高水平的 SA 积累（Song *et al.*, 2004），可能在植物防御和/或细胞死亡中发挥作用，从而帮助川滇高山栎有效地抵御病原物的侵害（Rate *et al.*, 1999）。*AtALDI* 可以与免疫组分 *AtPAD4* 和 *ICS1*（一种 SA 合成酶）共同作用来增强植物抗病性，并参与 SA 相关的植物-病原体的防御反应（Cecchini *et al.*, 2015）。*AtSARD1* 和 *AtCBP60g* 通过控制 *ICS1* 表达并诱导 SA 生物合成，参与植物的防御反应，*AtCBP60g* 在防御反应的早期发挥作用，而 *AtSARD1* 在 *AtCBP60g* 作用后扮演更重要的角色（Wang *et al.*, 2011）。除此之外，*AtFMO1* 可以和 *AtACD1*、*AtPAD4* 和 *AtEDS1* 共同激活 SA 途径参与防御性反应（Mishina and Zeier, 2006）。总的来说，*QaGDUI3* 基因的过表达通过激活 SA 途径相关基因的表达，来增加植物体中 SA 的含量，从而引发川滇高山栎对病原菌及病毒的防御反应（Yalpani and Raskin, 1993），有效地抑制病原物对自身的侵害，增强其对高山环境的适应性。*QaGDUI3* 基因与 SA 途径相关基因表达的紧密关系，可能是川滇高山栎能够适应快速变化的高山环境并进化出物种特异的抵抗病原体的基本策略。

利用镰刀病原菌菌液侵染植物并观察植物发病症状，不仅能够鉴定不同种类镰刀菌的致病力，也能够为植物抗病种质资源的筛选提供参考。本研究结合 *QaGDUI3* 转基因拟南芥根部和叶片分别对部分镰刀菌具有初步的抗性结果，推测 *QaGDUI3* 基因可能与植物的抗性相关。进一步发现 *QaGDUI3* 转基因拟南芥对茄病镰刀菌具有抗性，由于茄病镰刀菌为栎属植物的内共生菌（Lynch *et al.*, 2013），因此我们推测在

川滇高山栎中 *QaGDU3* 基因可能使川滇高山栎能够抵御茄病镰刀菌等各种病原菌的侵害，从而增强其对高山环境的适应性，但对于在川滇高山栎中 *QaGDU3* 基因具体的抗病分子机制还需进一步的实验和分析进行研究。

5 结论与展望

5.1 结论

本研究结合三代 Nanopore 和二代 illumina 测序数据, 利用 Hi-C 技术组装得到了染色体水平上高质量的川滇高山栎基因组序列图谱, 并注释得到川滇高山栎蛋白编码基因。利用系统发育和比较基因组学分析估计了川滇高山栎的分化时间, 与此同时通过对川滇高山栎基因组中正选择基因及功能的分析, 初步探究了川滇高山栎对高山环境的适应性进化。进一步通过分子克隆等手段对正选择 *QaGDU3* 基因进行鉴定及功能分析。得出的主要结论如下:

- (1) 组装得到川滇高山栎基因组大小为 1.15Gb, contig 数量为 1 445 个, contig N50 为 1.20Mb。1 101.96Mb (96.02%) 的基因组序列组装到 12 条假染色体上。基因组注释结果显示 58.04% 的川滇高山栎基因组由重复序列组成, 其中长末端重复序列是含量最丰富的重复类型, 占川滇高山栎基因组的 40.86%。在川滇高山栎基因组中预测得到 26 441 个蛋白质编码基因, 其中 92.19% 基因功能被注释。
- (2) 川滇高山栎与近缘物种欧洲栓皮栎、加州白栎、有柄栎聚为一枝, 且川滇高山栎的物种形成可能由喜马拉雅山的快速隆升引发的, 导致川滇高山栎为适应极端的环境而进化出特化的表型和生理性状。
- (3) 在川滇高山栎基因组中鉴定得到了 416 个正选择基因, GO 和 KEGG 功能富集结果表明这些基因功能集中在“RNA 和受损 DNA 结合”、“光系统 II 氧复合物”和“核苷酸切除修复”等方面, 这些基因可能通过调控川滇高山栎体内生理生化代谢过程, 进一步在川滇高山栎对高海拔适应方面发挥作用。
- (4) “VIMAG”为 GDU 蛋白的一个特有的疏水、功能保守的结构域。且转基因 *QaGDU3* 拟南芥 T2 代株系莲座大小随着 *QaGDU3* 基因表达量升高而变小。
- (5) *QaGDU3* 基因的过表达能够激活 SA 途径相关基因的表达, 同时 *QaGDU3* 转基因拟南芥对茄病镰刀菌具有抗性, 因此推测在川滇高山栎中 *QaGDU3* 基因与在转基因拟南芥中具有相同的作用, 通过增加川滇高山栎中 SA 的含量, 引发川滇高山栎对病原菌及病毒的防御反应, 增强其对高山环境的适应性。

5.2 展望

本论文组装得到了染色体水平高质量川滇高山栎参考基因组, 进一步对正选择 *QaGDU3* 基因进行了分子克隆和功能验证, 但依然存在一些问题需要解决, 下一步可从以下几个方面进行更深入的研究和分析:

- (1) 本研究仅基于川滇高山栎的全基因组数据进行分析, 后期可结合重测序和转录组等多组学数据, 更全面的探究并解析川滇高山栎及栎属植物的适应性进化的

遗传信息。

- (2) 目前已经初步确定了 *QaGDU3* 基因功能，然而关于 *QaGDU3* 基因在川滇高山栎中的调控研究还是空白，下一步可以通过筛选调控 *QaGDU3* 基因表达的转录因子，进一步阐明与川滇高山栎抗性相关基因的分子调控机制。
- (3) 本研究仅在异源表达载体中初步验证了 *QaGDU3* 转基因拟南芥对部分镰刀菌的抗性，接下来可通过测定与植物抗性相关的生理指标来筛选出具有较高抗病性状的 *QaGDU3* 转基因拟南芥植株。同时可结合相关实验进一步验证 *QaGDU3* 基因具体在川滇高山栎中的作用机制，为川滇高山栎重要抗病性状的遗传解析和遗传改良提供参考，同时为研究川滇高山栎抗病机理和分子育种提供材料。

参考文献

- 常秋霜. 拟南芥跨膜转运蛋白相关基因 T-DNA 插入突变体的筛选及表型鉴定[D]. 南京大学. 2015.
- 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.
- 程晓彬, 张旭, 何操, 杨平. 川滇高山栎繁育技术研究[J]. 林业实用技术. 2014(11): 38-40.
- 端木圻. 我国栎属资源的综合利用[J]. 河北林果研究, 1994, 9(2): 177-181.
- 方香玲, 张彩霞, 南志标. 紫花苜蓿镰刀菌根腐病研究进展[J]. 草业学报. 2019, 28(12): 169-183.
- 郭玉霞, 南志标, 王成章, 李春杰, 沈禹颖, 严学兵. 苜蓿根部入侵真菌研究进展[J]. 草业学报. 2009, 18(05): 243-249.
- 胡花蕾, 杨健, 李鹏, 刘彦群, 秦利. 柞树基因组研究进展[J]. 蚕业科学. 2019, 45(03): 431-440.
- 黄亚洲. 川滇高山栎组织碳氮代谢与萌生更新能力的关系研究[D]. 西南大学硕士学位论文. 2011.
- 宋江华, 张立新. 植物跨膜蛋白研究进展[J]. 生物学杂志. 2009, 26(06): 62-64.
- 王超, 卢杰, 周晨霓, 姚慧芳, 加央朗珍, 拉巴. 藏东南川滇高山栎叶功能性状海拔分布特征[J]. 森林与环境学报. 2021, 41(04): 366-372.
- 王国严, 罗建, 徐阿生, 庄颖梅. 西藏色季拉山川滇高山栎种群结构与动态[J]. 林业科学研究. 2011, 24(3): 292-299.
- 王国严, 徐阿生. 川滇高山栎研究综述[J]. 四川林业科技. 2008, 29(02): 23-29.
- 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.
- 王玉垚. 基于干旱及缺氧候选基因的生态适应研究[D]. 北京林业大学. 2018.
- 谢静, 朱万泽, 周鹏, 赵广, 王尧尧. 川滇高山栎灌丛营养元素积累与生物循环特征[J]. 西北植物学报. 2014, 34(05): 1032-1038.
- 杨玉坡, 李承彪, 管中天. 四川森林[M]. 北京: 中国林业出版社, 1992.
- Abràmoff M D, Magalhães P J, Ram S J. Image processing with ImageJ[J]. Biophotonics international, 2004, 11(7): 36-42.
- Ai W, Liu Y, Mei M, Zhang X, Tan E, Liu H, Han X, Zhan H, Lu X. A chromosome-scale genome assembly of the Mongolian oak (*Quercus mongolica*)[J]. Molecular Ecology Resources, <https://doi.org/10.1111/1755-0998.13616>, 2020.
- Aitken S N, Whitlock M C. Assisted gene flow to facilitate local adaptation to climate change[J]. Annual Review of Ecology Evolution and Systematics, 2013, 44(1): 367.
- Akilli S, Ulubaş Serçe Ç, Katiırciođlu Y Z, Maden S. Phytophthora dieback on narrow leaved ash in the Black Sea region of Turkey[J]. Forest Pathology, 2013. 43(3): 252-256.
- Alidadi A, Kowsari M, Javan-Nikkhah M, Jouzani G R, Rastaghi M E. New pathogenic and endophytic fungal species associated with Persian oak in Iran[J]. European Journal of Plant Pathology, 2019.

- 155(3): 1017-1032.
- Bailey T L, Johnson J, Grant C E, Noble W S. The MEME Suite[J]. *Nucleic Acids Research*, 2015, 43(W1): W39-W49.
- Bernhofer M, Dallago C, Karl T, Satagopam V, Heinzinger M, Littmann M, Olenyi T, Qiu J, Schütze K, Yachdav G. PredictProtein-predicting protein structure and function for 29 years[J]. *Nucleic acids research*, 2021, 49(W1): W535-W540.
- Canaguier A, Grimplet J, Di Gaspero G, Scalabrin S, Duchêne E, Choisine N, Mohellibi N, Guichard C, Rombauts S, Le Clainche I. A new version of the grapevine reference genome assembly (12X. v2) and of its annotation (VCost. v3) [J]. *Genomics Data*, 2017, 14: 56.
- Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis[J]. *Molecular Biology and Evolution*, 2000, 17(4): 540-552.
- Cavender Bares J. Diversity, distribution and ecosystem services of the North American oaks[J]. *International oaks*, 2016, 27: 37-48.
- Cecchini N M, Jung H W, Engle N L, Tschaplinski T J, Greenberg J T. ALD1 regulates basal immune components and early inducible defense responses in *Arabidopsis*[J]. *Molecular Plant-Microbe Interactions*, 2015, 28(4): 455-466.
- Chen H, Zhang Z, Teng K, Lai J, Zhang Y, Huang Y, Li Y, Liang L, Wang Y, Chu C. Up-regulation of LSB1/GDU3 affects geminivirus infection by activating the salicylic acid pathway[J]. *The Plant Journal*, 2010, 62(1): 12-23.
- Chi C C, Childers W R, Hanson E W. Penetration and subsequent development of three *Fusarium* species in alfalfa and red clover[J]. *Phytopathology*, 1964, 54: 434-437.
- Daccord N, Celton J, Linsmith G, Becker C, Choisine N, Schijlen E, Van de Geest H, Bianco L, Micheletti D, Velasco R. High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development[J]. *Nature genetics*, 2017, 49(7): 1099-1106.
- De Vylder J, Vandenbussche F, Hu Y, Philips W, Van Der Straeten D. Rosette tracker: an open source image analysis tool for automatic quantification of genotype effects[J]. *Plant Physiology*, 2012, 160(3): 1149-1159.
- Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation[J]. *Science*, 2002, 295(5558): 1306-1311.
- Denk T, Grimm GW, Manos PS, Deng M, Hipp AL. Oaks physiological ecology. Exploring the functional diversity of genus *Quercus* L.[M]. Switzerland: Springer, 2018: 13–38.
- Dong X. The role of membrane-bound ankyrin-repeat protein ACD6 in programmed cell death and plant defense[J]. *Science's STKE*, 2004, 2004(221): e6.
- Dosskey M, Schultz R C, Isenhardt T M. Riparian Buffers for Agricultural Land[J]. *Agroforestry Notes*, 1997.

- Dow M, Newman M A, von Roepenack E. The induction and modulation of plant defense responses by bacterial lipopolysaccharides[J]. Annual Review of Phytopathology, 2000, 38: 241-261.
- Doyle J. A rapid DNA isolation procedure for small quantities of fresh leaf tissue[J]. Phytochemical Bulletin, 1987, 19, 11-15.
- Du F K, Hou M, Wang W, Mao K, Hampe A. 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, Wang Y, Ueno S, Lafontaine G. Contrasted patterns of local adaptation to climate change across the range of an evergreen oak, *Quercus aquifolioides*[J]. Evolutionary Applications, 2020, 13(9): 2377-2391.
- Edreva A, Kostoff D. A novel strategy for plant protection: Induced resistance[J]. Journal of cell and Molecular Biology, 2004, 3(2): 61-69.
- Falk A, Feys B J, Frost L N, Jones J D, Daniels M J, Parker J E. EDS1, an essential component of R gene-mediated disease resistance in *Arabidopsis* has homology to eukaryotic lipases[J]. Proceedings of the National Academy of Sciences, 1999, 96(6): 3292-3297.
- Feng L, Zheng Q J, Qian Z Q, Yang J, Zhang Y P, Li Z H, Zhao G F. 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.
- Goodstein D M, Shu S, Howson R, Hellsten U, Putnam N, Rokhsar D S. Phytozome: a comparative platform for green plant genomics[J]. Nucleic Acids Research, 2012, 40(Database issue): D1178-D1186.
- Greenhill A W, Chibnall A C. The exudation of glutamine from perennial rye-grass[J]. Biochemical Journal, 1934, 28(4): 1422-1427.
- Gugger P F, Ikegami M, Sork V L. Influence of late Quaternary climate change on present patterns of genetic variation in valley oak, *Quercus lobata* Nee[J]. Molecular Ecology, 2013, 22(13): 3598-3612.
- Guo X, Hu Q, Hao G, Wang X, Zhang D, Ma T, Liu J. The genomes of two *Eutrema* species provide insight into plant adaptation to high altitudes[J]. DNA Research, 2018, 25(3): 307-315.
- Gururani M A, Venkatesh J, Upadhyaya C P, Nookaraju A, Pandey S K, Park S W. Plant disease resistance genes: Current status and future directions[J]. Physiological and Molecular Plant Pathology, 2012, 78(51): 51-65.
- Hanada K, Zou C, Lehti-Shiu M D, Shinozaki K, Shiu S H. Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli[J]. Plant Physiology, 2008, 148(2): 993-1003.
- Hao H, Jiang C, Wu J, Zong C, Zeng D, Alimubike N. The effects of scatter-hoarding rodents on Mongolian oak (*Quercus mongolica*) acorn dispersal and seedling distribution in urban woodland[J].

- Urban Ecosystems, 2017, 20(6): 1277-1283.
- He Z H, Cheeseman I, He D, Kohorn B D. A cluster of five cell wall-associated receptor kinase genes, Wak1-5, are expressed in specific organs of *Arabidopsis*[J]. Plant Molecular Biology, 1999, 39(6): 1189-1196.
- He Z U, He D, Kohorn B D. Requirement for the induced expression of a cell wall associated receptor kinase for survival during the pathogen response[J]. Plant Journal, 1998, 14(1): 55-63.
- Hedges S B, Dudley J, Kumar S. TimeTree: a public knowledge-base of divergence times among organisms[J]. Bioinformatics, 2006, 22(23): 2971-2972.
- Herrero S, Gonzalez E, Gillikin J W, Velez H, Daub M E. Identification and characterization of a pyridoxal reductase involved in the vitamin B6 salvage pathway in *Arabidopsis*[J]. Plant Molecular Biology, 2011, 76(1-2): 157-169.
- Hipp A L, Manos P S, Hahn M, Avishai M, Bodénès C, Cavender Bares J, Crowl A A, Deng M, Denk T, Fitz Gibbon S. Genomic landscape of the global oak phylogeny[J]. New Phytologist, 2020, 226(4): 1198-1212.
- Hubert F, Grimm G W, Jousset E, Berry V, Franc A, Kremer A. Multiple nuclear genes stabilize the phylogenetic backbone of the genus *Quercus* [J]. Systematics and Biodiversity, 2014, 12(4): 405-423.
- Ingo, Grunwald, Ines, Rupprecht, Gadi, Schuster, Klaus, Kloppstech. Identification of guttation fluid proteins: the presence of pathogenesis-related proteins in non-infected barley plants[J]. Physiologia Plantarum, 2003, 119(2): 192-202.
- Jacob A, Lancaster J, Buhler J, Harris B, Chamberlain R D. Mercury BLASTP: Accelerating Protein Sequence Alignment[J]. ACM Transactions on Reconfigurable Technology and Systems, 2008, 1(2): 9.
- Jaillon O, Aury J M, Noel B, Weissenbach J, Quétier F, Wincker P. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla[J]. Nature, 2007, 449(7161): 463-467.
- Jirage D, Tootle T L, Reuber T L, Frost L N, Feys B J, Parker J E, Ausubel F M, Glazebrook J. *Arabidopsis thaliana* PAD4 encodes a lipase-like gene that is important for salicylic acid signaling[J]. Proceedings of the National Academy of Sciences, 1999, 96(23): 13583-13588.
- Jupe F, Pritchard L, Etherington G J, Mackenzie K, Cock P J, Wright F, Sharma S K, Bolser D, Bryan G J, Jones J D. Identification and localisation of the NB-LRR gene family within the potato genome[J]. BMC Genomics, 2012, 13: 75.
- Katoh K, Standley D M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability[J]. Molecular Biology and Evolution, 2013, 30(4): 772-780.
- Kochhar A, Khurana J P, Tyagi A K. Nucleotide sequence of the psbP gene encoding precursor of 23-kDa

- polypeptide of oxygen-evolving complex in *Arabidopsis thaliana* and its expression in the wild-type and a constitutively photomorphogenic mutant[J]. *DNA Research*, 1996, 3(5): 277-285.
- Komarnytsky S, Borisjuk N V, Borisjuk L G, Alam M Z, Raskin I. Production of recombinant proteins in tobacco guttation fluid[J]. *Plant Physiology*, 2000, 124(3): 927-934.
- Kremer A, Ronce O, Robledo-Arnuncio J J, Kuparinen A, Gerber S, Schueler S. Long-distance gene flow and adaptation of forest trees to rapid climate change[J]. *Ecology Letters*, 2012, 15(4): 378-392.
- Kremer A. Microevolution of European temperate oaks in response to environmental changes[J]. *Comptes Rendus Biologies*, 2016, 339(7-8): 263-267.
- Kumar S, Stecher G, Tamura K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets[J]. *Molecular Biology and Evolution*, 2016, 33(7): 1870-1874.
- Kuroda M, Dey S, Sanders O I, Rosen B P. Alternate energy coupling of ArsB, the membrane subunit of the Ars anion-translocating ATPase[J]. *Journal of Biological Chemistry*, 1997, 272(1): 326-331.
- Leath K T. *Fusarium* root rot of forage species: Pathogenicity and host range[J]. *Phytopathology*, 1978, 68(6): 826-831.
- Lesur I, Durand J, Sebastiani F, Gyllenstrand N, Bodénès C, Lascoux M, Kremer A, Vendramin G G, Plomion C. A sample view of the pedunculate oak (*Quercus robur*) genome from the sequencing of hypomethylated and random genomic libraries[J]. *Tree Genetics and Genomes*, 2011, 7(6): 1277-1285.
- Li C, Guan Z, Liu D, Raetz C. Pathway for lipid A biosynthesis in *Arabidopsis thaliana* resembling that of *Escherichia coli*[J]. *Proceedings of the National Academy of Sciences*, 2011, 108(28): 11387-11392.
- Li L, Stoeckert C J, Roos D S. OrthoMCL: identification of ortholog groups for eukaryotic genomes[J]. *Genome Research*, 2003, 13(9): 2178-2189.
- Linaldeddu B T, Sirca C, Spano D, Franceschini A. Variation of endophytic cork oak - associated fungal communities in relation to plant health and water stress[J]. *Forest Pathology*, 2011. 41(3): 193-201.
- Litou Z I, Bagos P G, Tsirigos K D, Liakopoulos T D, Hamodrakas S J. Prediction of cell wall sorting signals in gram-positive bacteria with a hidden markov model: application to complete genomes[J]. *Journal of bioinformatics and computational biology*, 2008, 6(02): 387-401.
- 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.
- Liu X, Bush D R. Expression and transcriptional regulation of amino acid transporters in plants[J]. *Amino acids*, 2006. 30(2): 113-120.
- Livak K J, Schmittgen T D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method[J]. *Methods*, 2001, 25(4): 402-408.

- Lu H, Liu Y, Greenberg J T. Structure-function analysis of the plasma membrane- localized *Arabidopsis* defense component ACD6[J]. *Plant Journal*, 2005, 44(5): 798-809.
- Lu H, Rate D N, Song J T, Greenberg J T. ACD6, a novel ankyrin protein, is a regulator and an effector of salicylic acid signaling in the *Arabidopsis* defense response[J]. *Plant Cell*, 2003, 15(10): 2408-2420.
- Lu H, Salimian S, Gamelin E, Wang G, Fedorowski J, Lacourse W, Greenberg J T. Genetic analysis of *acd6-1* reveals complex defense networks and leads to identification of novel defense genes in *Arabidopsis*[J]. *Plant Journal*, 2009, 58(3): 401-412.
- Lukong K E, Chang K W, Khandjian E W, Richard S. RNA-binding proteins in human genetic disease[J]. *Trends in Genetics*, 2008, 24(8): 416-425.
- Luo D, Bernard D G, Balk J, Hai H, Cui X. The DUF59 family gene *AE7* acts in the cytosolic iron-sulfur cluster assembly pathway to maintain nuclear genome integrity in *Arabidopsis*[J]. *Plant Cell*, 2012, 24(10): 4135-4148.
- Lynch S C, Zambino P J, Mayorquin J S, Wang D H, Eskalen A. Identification of new fungal pathogens of coast live oak in California[J]. *Plant Disease*, 2013. 97(8): 1025-1036.
- Ma D, Guo Z, Ding Q, Zhao Z, Shen Z, Wei M, Gao C, Zhang L, Li H, Zhang S. Chromosome-level assembly of the mangrove plant *Aegiceras corniculatum* genome generated through Illumina, PacBio and Hi-C sequencing technologies[J]. *Molecular Ecology Resources*, 2021, 21(5): 1593-1607.
- Madsen S R, Nour-Eldin H H, Halkier B A. Collection of Apoplastic Fluids from *Arabidopsis thaliana* Leaves[J]. *Methods in Molecular Biology*, 2016, 1405: 35-42.
- Malhotra R, Laursen S B, Willis A C, Sim R B. Localization of the receptor-binding site in the collectin family of proteins[J]. *Biochemical Journal*, 1993, 293 (Pt1): 15-19.
- Meng H H, Su T, Gao X Y, Li J, Jiang X L, Sun H, Zhou Z K. Warm-cold colonization: response of oaks to uplift of the Himalaya-Hengduan Mountains[J]. *Molecular Ecology*, 2017, 26(12): 3276-3294.
- Mishina T E, Zeier J. The *Arabidopsis* flavin-dependent monooxygenase FMO1 is an essential component of biologically induced systemic acquired resistance[J]. *Plant Physiology*, 2006, 141(4): 1666-1675.
- Moller S, Croning M D, Apweiler R. Evaluation of methods for the prediction of membrane spanning regions[J]. *Bioinformatics*, 2001, 17(7): 646-653.
- Moreira A C, Martins J. Influence of site factors on the impact of *Phytophthora cinnamomi* in cork oak stands in Portugal[J]. *Forest Pathology*, 2005. 35(3): 145-162.
- Mun J H, Yu H J, Park S, Park B S. Genome-wide identification of NBS-encoding resistance genes in *Brassica rapa*[J]. *Molecular Genetics and Genomics*, 2009, 282(6): 617-631.
- Myburg A A, Grattapaglia D, Tuskan G A, Hellsten U, Hayes R D, Grimwood J, Jenkins J, Lindquist E, Tice H, Bauer D. The genome of *Eucalyptus grandis*[J]. *Nature*, 2014, 510(7505): 356-362.
- Nawrath C, Heck S, Parinthewong N, Metraux J P. EDS5, an essential component of salicylic acid-

- dependent signaling for disease resistance in *Arabidopsis*, is a member of the MATE transporter family[J]. *Plant Cell*, 2002, 14(1): 275-286.
- Newman M A, Daniels M J, Dow J M. The activity of lipid A and core components of bacterial lipopolysaccharides in the prevention of the hypersensitive response in pepper[J]. *Molecular Plant-Microbe Interactions*, 1997, 10(7): 926-928.
- Nie N H, Bent D H, Hull C H. SPSS: Statistical package for the social sciences[M]. American: McGraw-Hill New York, 1975.
- Nielsen H. Predicting Secretory Proteins with SignalP[J]. *Methods in Molecular Biology*, 2017, 1611: 59-73.
- Orengo C A, Thornton J M. Protein families and their evolution-a structural perspective[J]. *Annual Review of Biochemistry*, 2005, 74: 867-900.
- Ou S, Chen J, Jiang N. Assessing genome assembly quality using the LTR Assembly Index (LAI)[J]. *Nucleic Acids Research*, 2018, 46(21): e126.
- Pagni M, Ioannidis V, Cerutti L, Zahn-Zabal M, Jongeneel C V, Hau J, Martin O, Kuznetsov D, Falquet L. MyHits: improvements to an interactive resource for analyzing protein sequences[J]. *Nucleic Acids Research*, 2007, 35(Web Server issue): W433-W437.
- Peng B, Kong H, Li Y, Wang L, Zhong M, Sun L, Gao G, Zhang Q, Luo L, Wang G. OsAAP6 functions as an important regulator of grain protein content and nutritional quality in rice[J]. *Nature Communications*, 2014. 5(1): 1-12.
- Petit R J, Aguinagalde I, de Beaulieu J, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M. Glacial refugia: hotspots but not melting pots of genetic diversity[J]. *Science*, 2003, 300(5625): 1563-1565.
- Petit R J, Brewer S, Bordács S, Burg K, Cheddadi R, Coart E, Cottrell J, Csaikl U M, van Dam B, Deans J D. Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence[J]. *Forest Ecology and Management*, 2002, 156(1-3): 49-74.
- Petit R J, Carlson J, Curtu A L, Loustau M L, Plomion C, González-Rodríguez A, Sork V, Ducousso A. Fagaceae trees as models to integrate ecology, evolution and genomics[J]. *New Phytologist*, 2013, 197(2): 369-371.
- Pilot G, Stransky H, Bushey D F, Pratelli R, Ludewig U, Wingate V P, Frommer W B. Overexpression of GLUTAMINE DUMPER1 leads to hypersecretion of glutamine from Hydathodes of *Arabidopsis* leaves[J]. *Plant Cell*, 2004, 16(7): 1827-1840.
- 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.
- Plomion C, Aury J M, Amselem J, Salse J, Quesneville H, Kremer A. Decoding the oak genome: public

- release of sequence data, assembly, annotation and publication strategies[J]. *Molecular Ecology Resources*, 2016, 16(1): 254-265.
- Pratelli R, Pilot G. The plant-specific VIMAG domain of Glutamine Dumper1 is necessary for the function of the protein in *Arabidopsis*[J]. *FEBS Letters*, 2006, 580(30): 6961-6966.
- Pratelli R, Voll L M, Horst R J, Frommer W B, Pilot G. Stimulation of nonselective amino acid export by Glutamine Dumper proteins[J]. *Plant Physiology*, 2010, 152(2): 762-773.
- Ragazzi A, Vagniluca S, Moricca S, Vigniluca S. European expansion of oak decline: involved microorganisms and methodological approaches[J]. *Phytopathologia Mediterranea*, 1995: 207-226.
- Ramos A M, Usié A, Barbosa P, Barros P M, Capote T, Chaves I, Simões F, Abreu I, Carrasquinho I, Faro C. The draft genome sequence of cork oak[J]. *Scientific data*, 2018, 5(1): 1-12.
- Rate D N, Cuenca J V, Bowman G R, Guttman D S, Greenberg J T. The gain-of-function *Arabidopsis* *acd6* mutant reveals novel regulation and function of the salicylic acid signaling pathway in controlling cell death, defenses, and cell growth[J]. *Plant Cell*, 1999, 11(9): 1695-1708.
- Ryals J A, Neuenschwander U H, Willits M G, Molina A, Steiner H Y, Hunt M D. Systemic acquired resistance[J]. *Plant Cell*, 1996, 8(10): 1809-1819.
- Saez A, Rodrigues A, Santiago J, Rubio S, Rodriguez P L. HAB1-SWI3B interaction reveals a link between abscisic acid signaling and putative SWI/SNF chromatin-remodeling complexes in *Arabidopsis*[J]. *Plant Cell*, 2008, 20(11): 2972-2988.
- Salojärvi J, Smolander O, Nieminen K, Rajaraman S, Safronov O, Safdari P, Lamminmäki A, Immanen J, Lan T, TansKanaen J. Genome sequencing and population genomic analyses provide insights into the adaptive landscape of silver birch[J]. *Nature Genetics*, 2017, 49(6): 904-912.
- Salse J. Ancestors of modern plant crops[J]. *Current Opinion in Plant Biology*, 2016, 30: 134-142.
- Samac D A, Rhodes L H, Lamp W O. Compendium of Alfalfa Diseases and Pests, Third Edition || PART I: Infectious Diseases[M]. American: The American Phytopathological Society Publications, 2016.
- Sauerbrunn N, Schlaich N L. PCC1: a merging point for pathogen defence and circadian signalling in *Arabidopsis*[J]. *Planta*, 2004, 218(4): 552-561.
- Schenk P M, Kazan K, Wilson I, Anderson J P, Richmond T, Somerville S C, Manners J M. Coordinated plant defense responses in *Arabidopsis* revealed by microarray analysis[J]. *Proceedings of the National Academy of Sciences*, 2000, 97(21): 11655-11660.
- Segarra S, Mir R, Martinez C, Leon J. Genome-wide analyses of the transcriptomes of salicylic acid-deficient versus wild-type plants uncover Pathogen and Circadian Controlled 1 (PCC1) as a regulator of flowering time in *Arabidopsis*[J]. *Plant, Cell and Environment*, 2010, 33(1): 11-22.
- Shang-Guan K, Wang M, Htwe N, Carlson R W, Stacey G, Liang Y. Lipopolysaccharides trigger two successive bursts of reactive oxygen species at distinct cellular locations[J]. *Plant Physiology*, 2018, 176(3): 2543-2556.

- Shulaev V, Sargent D J, Crowhurst R N, Borodovsky M, Veilleux R E, Folta K M. The genome of woodland strawberry (*Fragaria vesca*) [J]. *Nature Genetics*, 2011, 43(2): 109-116.
- Sollars E S, Harper A L, Kelly L J, Bancroft I, Caccamo M, Buggs R J. Genome sequence and genetic diversity of European ash trees[J]. *Nature*, 2017, 541(7636): 212-216.
- Song J T, Lu H, McDowell J M, Greenberg J T. A key role for ALD1 in activation of local and systemic defenses in *Arabidopsis*[J]. *Plant Journal*, 2004, 40(2): 200-212.
- Sork V L, Cokus S J, Fitz-Gibbon S T, Zimin A V, Salzberg S L. High-quality genome and methylomes illustrate features underlying evolutionary success of oaks[J]. 2021.
- Sork V L, Fitz-Gibbon S T, Puiu D, Crepeau M, Gugger P F, Sherman R, Stevens K, Langley C H, Pellegrini M, Salzberg S L. First draft assembly and annotation of the genome of a California endemic oak *Quercus lobata* Nee (Fagaceae)[J]. *G3-Genes Genomes Genetics* (Bethesda), 2016, 6(11): 3485-3495.
- Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models[J]. *Bioinformatics*, 2006, 22(21): 2688-2690.
- Sun T, Liang W, Zhang Y, Li X. Negative regulation of resistance protein-mediated immunity by master transcription factors SARD1 and CBP60g[J]. *Journal of Integrative Plant Biology*, 2018, 60(11): 1023-1027.
- Tegeder M, Offler C E, Frommer W B, Patrick J W. Amino acid transporters are localized to transfer cells of developing pea seeds[J]. *Plant Physiology*, 2000. 122(2): 319-326.
- Tegeder M. Transporters for amino acids in plant cells: some functions and many unknowns[J]. *Current Opinion in Plant Biology*, 2012. 15(3): 315-321.
- Thatcher L F, Manners J M, Kazan K. *Fusarium oxysporum* hijacks COI1-mediated jasmonate signaling to promote disease development in *Arabidopsis*[J]. *The Plant Journal*, 2009. 58(6): 927-939.
- Thieme C J, Rojas-Triana M, Stecyk E, Paz-Ares J, Scheible W, Kragler F. Endogenous *Arabidopsis* messenger RNAs transported to distant tissues[J]. *Nature Plants*, 2015, 1(4): 15025.
- Tuskana G A, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray)[J]. *Science*, 2006, 313(5793): 1596-1604.
- Uknes S, Mauch-Mani B, Moyer M, Potter S, Williams S, Dincher S, Chandler D, Slusarenko A, Ward E, Ryals J. Acquired resistance in *Arabidopsis*[J]. *Plant Cell*, 1992, 4(6): 645-656.
- Verde I, Abbott A G, Scalabrin S, Schmutz J, Morgante M, Rokhsar D S. The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution[J]. *Nature Genetics*, 2013, 45(5): 487-494.
- Vujanovic V, Cogliastro A, St-Arnaud M, Neumann P, Gagnon D. First report of *Fusarium solani* canker and wilt symptoms on red oak (*Quercus rubra*) in Quebec, Canada[J]. *Plant disease*, 1999. 83(1): 78.

- Wan D, Li R, Zou B, Zhang X, Cong J, Wang R, Xia Y, Li G. Calmodulin-binding protein CBP60g is a positive regulator of both disease resistance and drought tolerance in *Arabidopsis*[J]. *Plant Cell Reports*, 2012, 31(7): 1269-1281.
- Wang L, Tsuda K, Truman W, Sato M, Nguyen L V, Katagiri F, Glazebrook J. CBP60g and SARD1 play partially redundant critical roles in salicylic acid signaling[J]. *Plant Journal*, 2011, 67(6): 1029-1041.
- Wildermuth M C, Dewdney J, Wu G, Ausubel F M. Isochorismate synthase is required to synthesize salicylic acid for plant defence[J]. *Nature*, 2001, 414(6863): 562-565.
- Wilkins M R, Gasteiger E, Bairoch A, Sanchez J C, Williams K L, Appel R D, Hochstrasser D F. Protein identification and analysis tools in the ExpASY server[J]. *Methods in Molecular Biology*, 1999, 112: 531-552.
- Willis J C, Shaw H K A. A dictionary of the flowering plants and ferns[J]. *American Fern Journal*, 1968, 58(3): 436.
- Wipf D, Ludewig U, Tegeder M, Rentsch D, Koch W, Frommer W B. Conservation of amino acid transporters in fungi, plants and animals[J]. *Trends in Biochemical Sciences*, 2002, 27(3): 139-147.
- Xing Y, Liu Y, Zhang Q, Nie X, Sun Y, Zhang Z, Li H, Fang K, Wang G, Huang H. Hybrid de novo genome assembly of Chinese chestnut (*Castanea mollissima*)[J]. *GigaScience*, 2019, 8(9): z112.
- Xing Y, Ree R H. Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot[J]. *Proceedings of the National Academy of Sciences*, 2017, 114(17): E3444-E3451.
- Xu X, Dimitrov D, Shrestha N, Rahbek C, Wang Z. A consistent species richness–climate relationship for oaks across the Northern Hemisphere[J]. *Global Ecology and Biogeography*, 2019, 28(8): 1051-1066.
- Yalpani N, Raskin I. Salicylic acid: a systemic signal in induced plant disease resistance[J]. *Trends in Microbiology*, 1993, 1(3): 88-92.
- Yang J, Anishchenko I, Park H, Peng Z, Ovchinnikov S, Baker D. Improved protein structure prediction using predicted interresidue orientations[J]. *Proceedings of the National Academy of Sciences*, 2020, 117(3): 1496-1503.
- Yang X, Jawdy S, Tschaplinski T J, TusKana G A. Genome-wide identification of lineage-specific genes in *Arabidopsis*, *Oryza* and *Populus*[J]. *Genomics*, 2009, 93(5): 473-480.
- Yang Z, Wong W S, Nielsen R. Bayes empirical bayes inference of amino acid sites under positive selection[J]. *Molecular Biology and Evolution*, 2005, 22(4): 1107-1118.
- Yang Z. PAML 4: phylogenetic analysis by maximum likelihood[J]. *Molecular Biology and Evolution*, 2007, 24(8): 1586-1591.
- Yoneda Y, Kuramoto N, Kawada K. The role of glutamine in neurogenesis promoted by the green tea amino acid theanine in neural progenitor cells for brain health[J]. *Neurochemistry International*, 2019, 129: 104505.
- Yu G, Wang L G, Han Y, He Q Y. clusterProfiler: An R package for comparing biological themes among

- gene clusters[J]. OMICS: A Journal of Integrative Biology, 2012, 16(5): 284-287.
- Zapata L, Ding J, Willing E M, Koornneef M, Ossowski S, Schneeberger K. Chromosome-level assembly of *Arabidopsis thaliana* Ler reveals the extent of translocation and inversion polymorphisms[J]. Proceedings of the National Academy of Sciences, 2016, 113(28): E4052-E4060.
- Zeidler D, Zahringer U, Gerber I, Dubery I, Hartung T, Bors W, Hutzler P, Durner J. Innate immunity in *Arabidopsis thaliana*: lipopolysaccharides activate nitric oxide synthase (NOS) and induce defense genes[J]. Proceedings of the National Academy of Sciences, 2004, 101(44): 15811-15816.
- Zerbini F M, Briddon R W, Idris A, Martin D P, Moriones E, Navas-Castillo J, Rivera-Bustamante R, Roumagnac P, Varsani A, Ictv R C. ICTV Virus Taxonomy Profile: Geminiviridae[J]. Journal of General Virology, 2017, 98(2): 131-133.
- Zhang H, Liu Y. VIGS Assays[J]. Bio-Protocol. 2014, 4(5): e1057.
- Zhang J, Nielsen R, Yang Z. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level[J]. Molecular Biology and Evolution, 2005, 22(12): 2472-2479.
- Zhang T, Zhao Y, Zhao J, Wang S, Jin Y, Chen Z, Fang Y, Hua C, Ding S, Guo H. Cotton plants export microRNAs to inhibit virulence gene expression in a fungal pathogen[J]. Nature Plants, 2016, 2(10): 1-6.
- Zhang Y, Xia R, Kuang H, Meyers B C. The Diversification of Plant NBS-LRR Defense Genes Directs the Evolution of MicroRNAs That Target Them[J]. Molecular Biology and Evolution, 2016, 33(10): 2692-2705.
- Zhao H, Ma H, Yu L, Wang X, Zhao J. Genome-wide survey and expression analysis of amino acid transporter gene family in rice (*Oryza sativa* L.)[J]. Public Library of Science one, 2012, 7(11): e49210.
- Zheng H, Von Mollard G F, Kovaleva V, Stevens T H, Raikhel N V. The Plant Vesicle-associated SNARE AtVTI1a Likely Mediates Vesicle Transport from the Trans-Golgi Network to the Prevacuolar Compartment[J]. Molecular Biology of the Cell, 1999, 10(7): 2251-22
- Zhu B, Si L, Wang Z, Jingjie Zhu Y Z, Shangguan Y, Lu D, Fan D, Li C, Lin H, Qian Q. Genetic control of a transition from black to straw-white seed hull in rice domestication[J]. Plant Physiology, 2011, 155(3): 1301-1311.

附录

附录 A 川滇高山栎 *QaGDU3* 克隆引物详细信息Appendix A Detailed information for cloning *QaGDU3*

用途	引物名称	上游引物序列 (5'-3')	下游引物序列 (5'-3')	扩增片段长度 (base pair)	
<i>QaGDU3</i> 基因克隆 及表达载体构建	Sca495	AAAGAAAGCAATG GCCGCAGC	GAGCCTGAGGTC TCATTGGCTTG	541	
	Sca495-1	CGACGACAAGACC GTGACCATGGCCGC AGCAAGAGAGCC	GAGGAGAAGAG CCGTTCATTGGC TTGAGCCTCTCT CTCTC	519	
	AtUBQ5q	ACACCAAGCCGAA GAAGA	TCCACAGGTTGC GTTAGG	128	
	Sca495-rt	AAGCCTAATGATTC GCATAA	AGAAGATGACCT ACTTGACA	116	
荧光定量 PCR	PR1-RT	GTAGGTGCTCTTGT TCTTC	CTCTTAGTTGTTTC TGCGTAG	153	
	ACD6-RT	TTGGTGGAATGAGT AATGG	TGAAGAATTGAA TCTCCTGTAT	103	
	PCC1-RT	AATTGGTTATCCGA CTAGAGA	ACAACCTCATTAT GGCTTCG	100	
	WAK1-RT	CCTTGTTGGTCTATG AGTTC	GCGACTTCTATT GCTATCTT	115	
	qRT-PCR	ALD-RT	AGGATTAGTGGGTT TGGA	ATTGGTATTAGA AGTGGAAGAG	117
		FMO1-RT	CCACCTTACCTTGA GATATTAG	CACCGATGAACC TFACTTC	109
		CBP60G-RT	GCTCGTAACTTAAC ATCAAG	TCTCGGACAATT CTTCTTATC	107
		SARD1-RT	ACTACTCCGACACT GAAG	TCATCAACAAGG ATAATCTCAA	113
		SID2-RT	TCTCCAACCTCATT CACT	GTCCTTGTCTCC ACTGTA	113
		PAD4-RT	AGAAGCAGCAATG AACAA	AACATCAGACGA ACAGAAC	101
EDS1-RT		AATGGCTGTGAGGA GTAG	TTGGATTGCGTAT GAAGTATT	113	

附录 B *QaGDU3* 基因生物信息分析软件及网址Appendix B Websites or softwares used in *QaGDU3* bioinformatics analysis

分析	软件/工具	网址
蛋白基本理化性质分析	ExPASy ProtParam	https://web.expasy.org/protparam/
蛋白亲疏水性预测	Prostate	https://web.expasy.org/protscale/
跨膜结构预测	TMHMM v2.0	https://services.healthtech.dtu.dk/service.php?TMHMM-2.0
信号肽预测	SignalP Server v4.1	https://services.healthtech.dtu.dk/service.php?SignalP-4.1
蛋白二级结构预测	Predict Protein	https://predictprotein.org/
蛋白三级结构预测	Robetta	https://robetta.bakerlab.org
蛋白亚细胞定位预测	ProtComp v9.0	http://www.softberry.com/
功能位点预测	MotifScan	https://myhits.sib.swiss/cgi-bin/motif_scan/
蛋白结构域预测	NCBI	https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi
蛋白序列比对	NCBI blastp	https://blast.ncbi.nlm.nih.gov/Blast.cgi
蛋白多重序列比对	DNAMAN v9.0	-
蛋白模体预测	MEME v5.4.1	https://meme-suite.org/meme/doc/meme.html
<i>QaGDU3</i> 同源序列比对	Phytozome v13	https://phytozome-next.jgi.doe.gov/blast-search

附录 C 13 个镰刀菌的分类地位

Appendix C Species of *Fusarium* strains

菌株编号	菌株种类	菌株编号	菌株种类
1A	锐顶镰刀菌 <i>Fusarium acuminatum</i>	2D	茄病镰刀菌 <i>Fusarium solani</i>
1B	茄病镰刀菌 <i>Fusarium solani</i>	2E	尖孢镰刀菌 <i>Fusarium oxysporum</i>
1C	层生镰刀菌 <i>Fusarium proliferatum</i>	3A	层生镰刀菌 <i>Fusarium proliferatum</i>
1D	甜瓜镰刀菌 <i>Fusarium roseum</i>	3B	层生镰刀菌 <i>Fusarium proliferatum</i>
2A	层生镰刀菌 <i>Fusarium proliferatum</i>	3C	茄病镰刀菌 <i>Fusarium solani</i>
2B	甜瓜镰刀菌 <i>Fusarium roseum</i>	3D	芬芳镰刀菌 <i>Fusarium redolens</i>
2C	层生镰刀菌 <i>Fusarium proliferatum</i>		

个人简介

杨霖，女，1996年7月生，籍贯河南省郑州市。2015年9月至2019年7月就读于河南师范大学生命科学学院生态学专业，取得理学学士学位。2019年9月至2022年7月就读于北京林业大学生态与自然保护学院生态学专业，导师为杜芳教授，研究方向为分子生态学。

硕士期间，主修了分子生态理论方法、高级森林生态学、生态学研究方法、生态学数据分析—R语言、多元统计分析等课程，成绩良好。硕士期间主要进行川滇高山栎 *QaGDU3* 基因生态功能研究，并以第一作者发表 CSCD 学术论文一篇。

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。以第一作者或通讯作者发表SCI论文24篇，发明专利1项，论文引用865次（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。

成果目录清单

[1]杨霖,尹康权,杜芳.川滇高山栎 *QaGDU3* 基因的克隆及遗传转化[J].植物科学学报. (已被植物科学学报期刊接收, CSCD 源刊)

致谢

惊风飘白日，光景西迟流。林花谢了春红太匆匆，无奈朝来寒雨晚来风。你我相聚在银杏大道满是落叶的秋，离别在绿意盎然蝉鸣不绝的夏。时间沉淀了研究生三年美好的回忆，与君一程，荣幸之至！

谁言寸草心，报得三春晖。花有重开时，人无再少年，三年研究生求学于异乡，独在异乡为异客，每逢佳节，对窗望月，乡愁便化作一张小小的车票。虽不能回家陪伴，但父母默默的支持和关怀、每次的关心问候都化作我不断前行的动力。

古之学者必严其师，师严然后道尊。研究生期间师从杜芳教授，从论文的选题到内容的写作及修改都离不开杜老师的指导，杜老师严谨的治学态度和专业的学术思维，让我对生态学专业有了更深层次的了解和认识，即便日后没有选择继续从事科研工作，但杜老师对科研的热爱和认真负责的工作态度深深的影响着我，获益匪浅，受用终身。乃知古学者，尊师复重友。感谢北京林业大学草业与草原学院尹康权老师对我在基因功能验证实验上的帮助，包括载体构建、引物设计和整个分子功能验证实验体系的构建和完成。感谢四川大学生命科学学院的马涛老师、王婧老师及王毓博师兄对我在生物信息分析和服务器学习使用上的帮助。感谢云南大学邓敏老师提供川滇高山栎种子，感谢北京林业大学园林学院赵惠恩老师提供了部分实验仪器。感谢在论文开题、中期、匿名评审及答辩阶段百忙之中对本文进行审阅并提出宝贵意见和建议的各位老师。金科玉律真师范，化雨春风入我庐，衷心的祝愿各位老师身体健康，平安喜乐！

白日去如箭，达者惜今阳。三年学习生涯，亦感谢同窗各位同学。感谢李月娟师姐教我阅读基因组文献和处理数据；感谢许阳师兄、王天瑞师兄和刘可可帮助采样；感谢燕语师姐帮助构建文库；感谢祁敏师姐、罗艳君和师妹原明星在作图和文章写作中注意事项的帮助；感谢苏蔚师姐和许阳师兄帮助实验室共同建设 `oaks of china` 网站；感谢康艺馨在生活上的帮助；感谢师弟秦伟、兰波在服务器分析中的帮助；感谢师妹张艺帮助温室植物栽培；感谢师妹李思扬和杨健在实验中的帮助。此外还要感谢 7 号楼 1421 全体宿舍成员在生活上的陪伴和帮助，同时也要感谢一直陪伴我成长的男朋友，沿途不枉为少年，希望我们被鲜花拥簇，看大海漫漫，热烈且自由，赤诚又勇敢。

春风拂碎冰雪，山水致以澄明。最后我要感谢北京林业大学，在“知山知水，树木树人”严谨求实的校风校训下，我度过了三个春夏秋冬，银杏大道见证了我们的脚步，时光沉淀了我们美好的回忆。

文末将至，书之有尽，致谢难穷。敬求学之路流觞，往事飞扬，敬少年无惧无畏，敬前路彷徨，也敬一路流浪；敬乌云镶金边的光芒，且敬浩荡满襟的饱霜。敬你滩边娇艳的玫瑰与山间弥漫的薄雾，敬我翡翠耀眼的理想与时休不落的初心。生

活从来都是泥沙俱下鲜花与荆棘并存，愿日后心怀感恩，不忘少年的朝气与梦想，手持利刃，披荆斩棘，乘风破浪，跨过重重的山丘，翻越连绵的四季。在下一个凛冬蔓延之际，找寻属于我们的璀璨星光。雨打霓裳，何必抱怨风雨，少年终会相遇，玫瑰终将绽放。