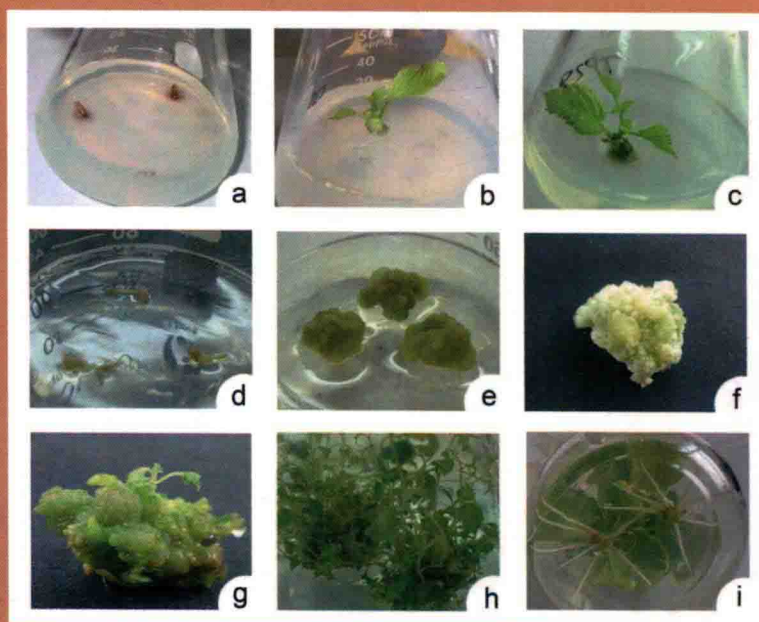


Cloning and Functional Analysis of *BpGT14* Gene and Promoter

白桦*BpGT14*基因和启动子的克隆及功能分析

曾凡锁 詹亚光 著



科学出版社

白桦 *BpGT14* 基因和启动子的 克隆及功能分析

曾凡锁 詹亚光 著

科学出版社

北京

内 容 简 介

本研究旨在明确糖基转移酶 *BpGT14* 基因在白桦生长发育中的功能,克隆了白桦 *GT14* 基因全长序列和启动子区,分析了其结构特征和表达模式,获得了 pRNAi-GG-*BpGT14* 转基因白桦。结果显示:RNAi 白桦茎横截面中木质部面积增加了 23%~47%,而韧皮部面积比例无明显变化;木质部中导管面积为对照的 1.82~2.16 倍,韧皮纤维面积比例无明显变化。干扰白桦纤维素含量无明显变化,半纤维素、果胶质含量显著降低,相比野生型分别减少了约 40%和 10%,说明 *BpGT14* 在植物半纤维素、果胶质等细胞壁多糖合成中具有重要作用。

本书对白桦和其他木本植物的遗传改良及功能基因的研究具有很好的借鉴意义。本书可供从事林木遗传改良和逆境分子生物学研究及教学的人员参考,也可用作林木遗传育种学相关专业研究生的学习参考书。

图书在版编目 (CIP) 数据

白桦 *BpGT14* 基因和启动子的克隆及功能分析/曾凡锁,詹亚光著. —北京:科学出版社,2018.6

ISBN 978-7-03-054437-7

I. ①白… II. ①曾… ②詹… III. ①白桦-基因克隆-研究 IV. ①S792.153

中国版本图书馆 CIP 数据核字(2017)第 222226 号

责任编辑:张会格 陈 新 高璐佳 / 责任校对:彭珍珍

责任印制:张 伟 / 封面设计:刘新新

科学出版社 出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

北京虎彩文化传播有限公司 印刷

科学出版社发行 各地新华书店经

*

2018 年 6 月第 一 版 开本: B5 (720×1000)

2018 年 6 月第一次印刷 印张: 10

字数: 200 000

定价: 108.00 元

(如有印装质量问题,我社负责调换)



前 言

糖基转移酶 (GT) 是一个重要的功能多样化酶家族, 与多糖和糖苷的合成密切相关。植物逐渐进化形成了 GT 大家族, 其在植物的生长和发育中进行糖基化的一系列反应。糖基化对于生物体是一个关键的反应, 截至目前, CAZy 数据库定义了超过 97 个 GT 家族。复杂的碳水化合物的合成在植物体内尤为重要, 这些物质控制植物体内各种分子功能、细胞与细胞的相互作用、信号及蛋白糖基化。寡糖和多糖的合成涉及了几百种不同的糖基转移酶, 它们可以将糖基从活化的供体分子转移到受体, 糖基化特异性的目标分子。

细胞壁的成分中, 90% 左右是多糖, 主要为纤维素、半纤维素和果胶类, 糖基转移酶参与细胞壁组分 (纤维素、半纤维素和果胶) 的合成, 催化各种分子的糖基化, 因此对植物细胞壁的发育有着重要影响。GT 家族在植物中的重要功能为催化植物细胞壁多糖的合成。糖基转移酶 14 家族属于糖基转移酶超家族, 该家族在植物细胞壁合成中扮演着重要角色。对于生物乙醇的生产, 细胞壁是最丰富的纤维素生物质来源, 因此, 对糖基转移酶 14 调控细胞壁生长发育的研究对于生物能源的开发具有重要的研究意义。

本书主要研究了糖基转移酶 *BpGT14* 基因在白桦生长发育中的功能。第 1 章简要介绍了糖基转移酶的分类与功能研究进展, 同时分析了启动子和 DNA 甲基化修饰的研究概况和国内外的发展趋势。第 2 章主要克隆了白桦 *GT14* 基因全长序列, 命名为 *BpGT14* 基因 (JQ409354), 利用生物信息学对其理化性质进行了分析, 揭示了该基因在不同部位、不同时间及对逆境响应的表达模式。应用染色体步移技术 Sitefinding-PCR 的方法克隆获得了该基因上游的启动子序列, 大小为 2169bp。第 3 章通过酵母单杂交筛选启动子 MYBPLANT 元件结合候选蛋白, 筛选得到了多个可能与其互作的候选蛋白, 包括糖基化类 MYB 转录因子、GPI 锚定蛋白及多个功能未知蛋白等。第 4 章研究筛选得到了一个含有 PHD 及 DUF3594 结构域的候选蛋白, 属于 Alfin-like 4 家族蛋白, 将其基因命名为 *BpAL4* 基因, 该基因全长 762bp、编码 252 个氨基酸。其 DNA 结合位点 PHD 结构域及功能位点 DUF3594 结构域共同作用来调节基因的表达。第 5 章主要进行了白桦 *BpGT14* 基因表达载体及 RNA 干扰载体的构建及遗传转化, 获得了 pRNAi-GG-*BpGT14* 转基因白桦, RNAi 白桦茎横截面中木质部面积增加了 23%~47%, 而韧皮部面积比例无明显变化。木质部中导管面积为对照的 1.53~2.83 倍, 韧皮纤维面积比例无变化。

干扰白桦半纤维素、果胶质的含量显著降低，相比野生型分别减少了约 40% 和 10%，这说明 *BpGT14* 在植物半纤维素、果胶质等细胞壁多糖合成中具有重要作用。第 6 章主要利用亚硫酸盐测序法，分析了白桦愈伤再生途径中 *BpGT14* 基因、启动子及编码区 DNA 甲基化水平及顺式元件甲基化位点的动态变化，并分析了 *BpGT14* 基因表达量的变化趋势。

本书的研究和出版得到了国家自然科学基金（31200463）和中央高校基本科研业务费专项资金项目等课题资助，特致诚挚谢意。同时，特别感谢参加上述课题的工程师贾洪柏、高级工程师齐凤慧及各位付出辛勤汗水的研究生——李蕾蕾、孙丰坤、周姗、李晓一、李思达、陈晓慧。

我们希望本书的出版为本研究领域的科研工作者和学生提供有价值的指导和参考。由于分子生物学领域发展迅速，本书难免存在不足或疏漏，敬请读者批评指正。

著 者

2017 年 3 月 20 日

目 录

前言

1 绪论	1
1.1 糖基转移酶简介	1
1.2 糖基转移酶功能	1
1.2.1 植物防御反应	1
1.2.2 植物次生代谢产物修饰	2
1.2.3 植物次生细胞壁合成	2
1.2.4 植物激素平衡	2
1.3 糖基转移酶 14 研究进展	3
1.3.1 GT14/DUF266 (GT14-like) 蛋白的鉴定	3
1.3.2 GT14 和 GT14-like 基因系统发育的关系	4
1.3.3 GT14 和 GT14-like 基因的进化	4
1.3.4 GT14 和 GT14-like 基因的表达模式分析	4
1.3.5 GT14 蛋白质三维结构及亚细胞定位	9
1.4 植物启动子研究	11
1.4.1 启动子结构及功能	11
1.4.2 启动子克隆方法	11
1.4.3 启动子的应用	12
1.5 转录因子研究	12
1.5.1 转录因子简介	12
1.5.2 转录因子分类及功能	12
1.6 转录因子与启动子互作研究方法	13
1.6.1 酵母单杂交	13
1.6.2 免疫共沉淀	14
1.7 植物 DNA 甲基化的作用方式及功能	14
1.7.1 植物 DNA 甲基化的作用方式	14
1.7.2 植物 DNA 甲基化的功能	15

1.8	木本植物不同发育阶段 DNA 甲基化水平和模式的变化	16
1.8.1	DNA 甲基化与木本植物的年龄	17
1.8.2	DNA 甲基化与木本植物的花期、休眠和育性	18
1.9	木本植物离体繁殖过程中的 DNA 甲基化模式重建	18
1.9.1	组织培养植株与野生植株的 DNA 甲基化水平和模式差异	19
1.9.2	再生过程、体胚发生、继代培养及外植体状态与 DNA 甲基化	20
1.10	植物 DNA 甲基转移酶及与甲基化有关的蛋白质	21
1.10.1	MET1 家族	21
1.10.2	CMT 家族	21
1.10.3	DRM 家族	22
1.11	白桦的生物学特性	23
1.12	研究意义	23
2	白桦 <i>BpGT14</i> 基因及其启动子的克隆及生物信息学分析	25
2.1	实验材料	25
2.1.1	植物材料	25
2.1.2	菌株及载体	25
2.1.3	数据分析	26
2.2	实验方法	26
2.2.1	白桦 <i>BpGT14</i> 基因生物信息学分析	26
2.2.2	白桦茎段悬浮细胞非生物胁迫处理	26
2.2.3	转录表达定量分析	27
2.2.4	数据处理	27
2.2.5	白桦 <i>BpGT14</i> 启动子的克隆	27
2.2.6	启动子生物信息学分析	30
2.2.7	植物表达载体的构建	30
2.2.8	植物表达载体农杆菌的转化	32
2.2.9	<i>BpGT14</i> 启动子在烟草中的表达活性	33
2.2.10	启动子在白桦细胞中的表达活性	33
2.3	结果与分析	34
2.3.1	白桦 <i>BpGT14</i> 基因的生物信息学分析	34
2.3.2	白桦 <i>BpGT14</i> 基因时空特异性表达分析	37
2.3.3	白桦茎段悬浮细胞 <i>BpGT14</i> 基因非生物胁迫下表达分析	38

2.3.4	白桦 <i>BpGT14</i> 基因启动子克隆	39
2.3.5	白桦 <i>BpGT14</i> 基因启动子序列元件分析	40
2.3.6	启动子 pXGUS-P 及 pXGFP-P 植物表达载体的构建及鉴定	41
2.3.7	植物表达载体转化农杆菌的鉴定	42
2.3.8	非生物胁迫下启动子在烟草中的表达活性	42
2.3.9	启动子在白桦细胞中的表达活性	43
2.4	讨论	45
2.5	本章小结	49
3	酵母单杂交筛选启动子区及 MYBPLANT 元件结合候选蛋白	51
3.1	实验材料	51
3.2	实验方法	51
3.2.1	启动子 1156bp 片段 PCR 扩增	51
3.2.2	诱饵载体 pAbAi 线性化及与启动子片段的连接	51
3.2.3	MYBPLANT 元件的构建	52
3.2.4	酵母感受态的制备	53
3.2.5	重组诱饵载体 pAbAi 转化酵母感受态细胞	53
3.2.6	诱饵酵母的鉴定	54
3.2.7	报告基因在诱饵酵母中本底表达水平的测定	54
3.2.8	白桦文库 cDNA 的合成	55
3.2.9	酵母单杂交文库的构建及筛选	57
3.2.10	阳性克隆菌株的鉴定	58
3.2.11	阳性克隆 cDNA 插入片段的生物信息学分析	58
3.2.12	启动子候选互作蛋白 <i>BpARF2</i> 基因的非生物胁迫响应	59
3.3	结果与分析	59
3.3.1	重组诱饵载体的构建	59
3.3.2	重组诱饵质粒转化酵母感受态细胞	59
3.3.3	报告基因在诱饵酵母中本底表达水平的测定	59
3.3.4	白桦文库 cDNA 的合成	60
3.3.5	酵母单杂交文库的构建及筛选	61
3.3.6	阳性克隆 cDNA 插入片段的生物信息学分析	61
3.3.7	启动子互作候选蛋白 <i>BpARF2</i> 基因非生物胁迫响应	64
3.3.8	MYB-pAbAi 重组诱饵载体的构建	67

3.3.9	MYBPLANT 元件结合候选蛋白的筛选	68
3.4	讨论	68
3.5	本章小结	70
4	启动子 MYBPLANT 元件结合候选蛋白的鉴定	72
4.1	实验材料	72
4.2	实验方法	72
4.2.1	候选蛋白互作强弱初步鉴定	72
4.2.2	白桦 <i>BpAL4</i> 基因的克隆	72
4.2.3	白桦 <i>BpAL4</i> 基因生物信息学分析	74
4.2.4	候选蛋白互作强弱初步鉴定	74
4.2.5	三次重复元件 pCXGUS 植物表达载体构建	76
4.2.6	三亲杂交及农杆菌的遗传转化	77
4.2.7	烟草中 <i>GUS</i> 基因表达水平鉴定	78
4.2.8	非生物胁迫下白桦 <i>BpAL4</i> 基因表达水平鉴定	78
4.3	结果与分析	78
4.3.1	候选蛋白互作强弱鉴定	78
4.3.2	白桦 <i>BpAL4</i> 基因的克隆	79
4.3.3	白桦 <i>BpAL4</i> 基因生物信息学分析	79
4.3.4	白桦 <i>BpAL4</i> 基因植物表达载体的构建	82
4.3.5	三次重复元件 pCXGUS 植物表达载体构建	83
4.3.6	<i>BpAL4</i> 蛋白及 MYBPLANT 元件共转化分析	84
4.4	讨论	85
4.5	本章小结	87
5	白桦 <i>BpGT14</i> 基因表达载体及 RNA 干扰载体的构建及遗传转化	88
5.1	实验材料	88
5.2	实验方法	88
5.2.1	白桦总 RNA 的提取与 cDNA 的合成	88
5.2.2	白桦 <i>BpGT14</i> 基因 pMD18-T 载体的构建	88
5.2.3	白桦 <i>BpGT14</i> 基因干扰载体的构建	89
5.2.4	三亲杂交法转化根癌农杆菌	89
5.2.5	根癌农杆菌介导的白桦遗传转化	90
5.2.6	转基因植株的分子鉴定	91

5.2.7	转基因白桦组培苗的移栽	92
5.2.8	石蜡切片	92
5.2.9	转基因白桦木质素、纤维素、半纤维素和果胶的测定	93
5.3	结果与分析	96
5.3.1	白桦 <i>BpGT14</i> 基因 pMD18-T 载体的构建	96
5.3.2	白桦 <i>BpGT14</i> 基因干扰载体的构建	97
5.3.3	三亲杂交	97
5.3.4	转基因植株的获得	98
5.3.5	转基因植株的分子鉴定	98
5.3.6	RNAi 白桦茎段结构分析	99
5.3.7	<i>BpGT14</i> -RNAi 转基因白桦细胞壁成分分析	101
5.4	讨论	103
5.5	本章小结	106
6	白桦微繁过程中 <i>BpGT14</i> 表达及 DNA 甲基化的变异机制	108
6.1	实验材料	108
6.1.1	植物材料	108
6.1.2	实验器材	108
6.2	实验方法	108
6.2.1	培养基及培养条件	108
6.2.2	外植体选取与消毒	109
6.2.3	腋芽增殖的继代培养	109
6.2.4	愈伤组织的诱导及分化	109
6.2.5	再生植株生根及移栽	109
6.2.6	酶液制取	109
6.2.7	保护酶活性的测定	109
6.2.8	木质素测定	111
6.2.9	丙二醛含量测定	111
6.2.10	可溶性糖含量测定	111
6.2.11	可溶性蛋白含量测定	112
6.2.12	白桦 DNA 提取及亚硫酸盐处理	112
6.2.13	<i>BpGT14</i> 及甲基转移酶基因表达量分析	113
6.2.14	数据处理	115

6.3 结果与分析	115
6.3.1 愈伤组织的诱导和分化	115
6.3.2 甲基转移酶基因的表达	115
6.3.3 甲基转移酶的活性分析	116
6.3.4 5'-CCGG 的胞嘧啶甲基化水平	118
6.3.5 再生不同阶段细胞氧化还原水平	120
6.3.6 再生不同阶段变异的主成分分析	120
6.3.7 愈伤再生不同发育阶段 <i>BpGT14</i> 基因表达分析	124
6.3.8 白桦 <i>BpGT14</i> 基因内含子生物信息学分析	124
6.3.9 愈伤再生不同发育阶段 <i>BpGT14</i> 基因启动子及编码区 DNA 甲基化 变异	128
6.4 讨论	131
6.4.1 甲基转移酶基因在再生过程中的表达	131
6.4.2 白桦组织培养中 DNA 甲基化的变化	131
6.4.3 微繁殖过程中的生理变化	132
6.4.4 不同发育阶段的表观遗传参数和生理指标的变化	133
6.4.5 愈伤再生不同发育阶段 <i>BpGT14</i> 基因启动子及编码区 DNA 甲基化 变异	133
6.5 本章小结	134
7 展望	136
参考文献	137

1 绪 论

1.1 糖基转移酶简介

糖基转移酶 (glycosyltransferase, GT) 是一个重要的功能多样化酶家族, 与多糖和糖苷的合成密切相关。植物逐渐进化形成了 GT 大家族, 这个大家族在植物的生长和发育过程中催化糖基化的一系列反应(Lao et al., 2014)。根据 Carbohydrate-Active enZymes (CAZy) 数据库, 模式植物拟南芥基因组中编码了超过 450 个糖基转移酶, 而水稻基因组包含 600 个成员 (Coutinho et al., 2003)。

糖基转移酶 (GT; EC 2.4.x.y) 可以催化活性糖基从供体分子转移到受体分子, 如糖类、核酸、脂质、蛋白质及各式各样的有机化合物 (Vogt and Jones, 2000)。糖基转移酶一般在 C 端保守, C 端被普遍认为是识别糖基的区域, 而 N 端因其不保守性则被认为是底物的识别位点 (Lairson et al., 2008)。糖基化对于生物体是一个关键的反应, 到目前为止, CAZy 数据库定义了超过 97 个 GT 家族 (<http://www.cazy.org/>)。复杂的碳水化合物的合成在植物体内尤为重要, 这些物质控制植物体内各种分子功能、细胞与细胞的相互作用、信号及蛋白糖基化 (Yonekura-Sakakibara and Hanada, 2011)。寡糖和多糖的合成涉及几百种不同的糖基转移酶, 它们可以将糖基从活化的供体分子转移到受体, 糖基化特异性的目标分子 (Breton et al., 2006)。因此, 所有这些糖基转移酶家族在植物细胞的结构和功能方面均扮演着重要角色。

1.2 糖基转移酶功能

1.2.1 植物防御反应

植物小分子糖基化对植物耐受逆境有着重要意义。拟南芥中一种糖基转移酶 UGT73B2 被证明与植物氧化胁迫抗性有关 (Kim et al., 2010), 烟草中的一种糖基转移酶能够糖基化羟基香豆素、羟基甲氧基香豆素和羟基肉桂酸, 并且在超敏反应中, 水杨酸能够上调其表达 (Fraissinet-Tachet et al., 1998)。同时有研究发现, 在拟南芥中过表达 *UGT74E2* 糖基转移酶基因会使植物的抗旱及抗盐性增强 (Tognetti et al., 2010)。尽管已有研究表明了糖基转移酶基因与植物逆境胁迫相关, 但其中的具体机制目前尚不清楚。

1.2.2 植物次生代谢产物修饰

植物次生代谢产物对植物的生长发育具有重要的作用，而糖基化对于次生代谢产物的形成有着关键意义。大量的次生代谢产物具有多种糖型，这些糖型可以影响它们的物理和化学特性（周文灵等，2009）。目前，已经发现大量参与次生代谢物形成的糖基转移酶。拟南芥中，两个糖基转移酶基因 *UGT73C6* 和 *UGT78D1* 被发现与黄酮醇糖苷物的生物合成相关（Jones et al., 2003）。甜叶菊的叶子中有大量双萜化合物积累，而这些化合物不同位置的糖基化导致了其口感的不同（Richman et al., 2005）。同时有研究表明，部分糖基转移酶在植物苯丙烷途径中扮演着重要的角色，它们可以通过糖基化改变分子的水溶性等各种特征，起到调控作用（王会勇，2013）。植物木质素的合成前体为苯丙烷衍生物，这意味着糖基转移酶对植物木质素的合成具有重要调控作用。

1.2.3 植物次生细胞壁合成

细胞壁的成分中，90%左右是多糖，主要为纤维素、半纤维素和果胶类，糖基转移酶参与细胞壁组分（纤维素、半纤维素和果胶）的合成，催化各种分子的糖基化，因此对植物细胞壁的发育有着重要影响（Doblin et al., 2010; Sado et al., 2009）。GT 家族在植物中的重要功能为参与植物细胞壁多糖的合成。细胞壁中最丰富的多糖是纤维素，一种线性聚合物，可以通过纤维素合酶 GT2 家族成员（CESA）在质膜中合成（Liu et al., 2012）。多糖是具有分支的结构体，在高尔基体通过编码合成大量 GT 家族成员后分泌到细胞壁形成弯曲交联结构。蛋白质的糖基化发生在内质网和高尔基体，这是翻译后修饰的一个最普遍体现（Yang et al., 2013a）。目前，对杨树中糖基转移酶的研究较为广泛，对其转录图谱进行分析发现，多个基因对植物次生细胞壁具有重要作用，同时分析发现，糖基转移酶在杨树木质部中参与了碳水化合物的合成和重构，进而直接影响木质部的发育（Williamson et al., 2002）。

1.2.4 植物激素平衡

植物体内的激素水平平衡对于植物的生长发育及对外界环境的响应具有重要的意义（王军和侯丙凯，2008）。目前在植物体内，众多激素的糖苷物均已被发现，这表明激素的糖基化对于体内激素的调节有着重要的意义（Woodward and Bartel, 2005）。目前认为，糖基化可以使激素活性降低或者消失（王军和侯丙凯，2008）。拟南芥中对于 *UGT84B1* 基因的过表达植株研究表明，该基因可导致植株生长素

缺失 (Jackson et al., 2002), 而赤豆中的一个脱落酸糖基转移酶体外研究表明, 该基因产物可以使反式脱落酸糖基化 (Xu et al., 2002)。目前, 在烟草中, 对水杨酸的糖基化研究较为全面, 同时拟南芥中也发现了水杨酸诱导的糖基转移酶 (王军和侯丙凯, 2008)。

1.3 糖基转移酶 14 研究进展

糖基转移酶 14 家族属于糖基转移酶超家族, 该家族在植物细胞壁合成中扮演着重要角色 (Richmond, 2000)。对于生物乙醇的生产, 细胞壁是最丰富的纤维素物质来源, 因此对糖基转移酶 14 调控细胞壁生长发育的研究, 对于生物能源的开发具有重要的研究意义 (Cao et al., 2008)。在拟南芥、水稻、杨树、高粱和葡萄中 (表 1-1), 共鉴定了 62 个 *GT14* 基因和 106 个 *DUF266* 基因 (目前有假说称, *DUF266* 蛋白是一个新的 GT 分类, 与 GT 相关, 命名为 *GT14-like*) (Yang et al., 2008)。对于这些 *GT14* 基因的系统发育分析将 *GT14* 和 *GT14-like* 基因分到了两个不同的类别, 但蛋白结构域、3D 构型和基因表达分析揭露出两个类别均属于一个家族, 因此, 将其命名为 *GT14/GT14-like* 家族, 包括两个亚家族 (Ye et al., 2011)。

表 1-1 不同物种中 *GT14* 和 *GT14-like* 基因数目

亚家族	拟南芥	水稻	杨树	高粱	葡萄	总计
<i>GT14</i>	11	12	17	12	10	62
<i>GT14-like</i>	22	19	27	22	16	106
合计	33	31	44	34	26	168

在拟南芥和杨树中, 一半的 *GT14/GT14-like* 基因在茎和木质部优先表达, 暗示其在细胞壁合成中的重要作用 (Yang et al., 2013a)。AtGLCAT14A 是拟南芥 *GT14* 家族的一个蛋白质, 拟南芥中有 11 个蛋白质属于这个家族, 敲除该基因的突变体幼苗的胚轴和根的生长率提高, 表明该基因可能与细胞伸长相关 (Liu et al., 2012)。

1.3.1 *GT14/DUF266* (*GT14-like*) 蛋白的鉴定

为了验证 *GT14* 和 *DUF266* (*GT14-like*) 蛋白是属于同一个具有特殊支链蛋白结构域家族的假说, 作者用 HMMER-InterProScan 的方法在 5 个已测序的植物物种中获得了所有含有支链结构域的蛋白质序列 (Yang et al., 2008)。5 个已测序的植物物种包括一年生拟南芥 (双子叶植物)、水稻 (单子叶植物)、杨树 (多年生双子叶植物)、高粱 (单子叶植物) 和葡萄 (多年生双子叶植物), 该研究

共鉴定出 168 个含有支链结构域的非冗余的全长蛋白质序列，其中拟南芥中有 33 个、水稻中有 31 个、杨树中有 44 个、高粱中有 34 个、葡萄中有 26 个。蛋白质序列的长度为 75~651 个氨基酸，平均长度为 376 个氨基酸。用 HMMER-InterProScan 的方法所得到的蛋白质包含 CAZy 数据库中 11 个拟南芥 GT14 家族的蛋白质和水稻 GT 数据库中 12 个 GT14 家族的蛋白质，这表明基因筛选的方法是有效的。

1.3.2 *GT14* 和 *GT14-like* 基因系统发育的关系

对获得的 168 个含有支链蛋白结构域的全长蛋白质序列构建系统进化树，结果如图 1-1A 所示，这 168 个蛋白质主要被分成两个进化枝，分别被命名为 GT14 和 GT14-like。GT14-like 蛋白先前被称为 DUF266 蛋白，可被分成 5 组 (A1~A5)，GT14 蛋白可被分成 8 组 (B1~B8)。用 InterProScan 软件对这 168 个 GT14 和 GT14-like 的蛋白质序列进行功能分析 (Hunter et al., 2009)，共鉴定出 3 个蛋白结构域：核心-2/1-分支酶功能域 (IPR021141)，GT14 结构域 (IPR003406)，以及一个钙离子结合位点结构域 (IPR018247)。如图 1-1B 所示，这三种结构域组成两种类型的域结构，第一种类型是只含有支链结构域，第二种类型是既含有支链结构域又含有 GT14 结构域。GT14 蛋白的系统发育的进化枝只包含第二种类型的域结构，GT14-like 蛋白的进化枝只包含第一种类型的域结构 (图 1-1A)。域结构分类和系统发育分类之间的一致性验证了系统进化分析的结果。

1.3.3 *GT14* 和 *GT14-like* 基因的进化

从图 1-1A 中的系统进化树中可以看出，在 *GT14* 及 *GT14-like* 的进化枝中，双子叶植物中的基因数量是单子叶植物中的 1.5 倍。然而在一些系统进化组中双子叶植物与单子叶植物基因数量的比率是不同的，如 A1-1 组中，双子叶植物有 17 个基因，单子叶植物有 2 个，B2 组中只有双子叶植物，这表明 *GT14* 和 *GT14-like* 亚家族在双子叶植物 (A1 组和 B2 组) 和单子叶植物 (A2 组和 B1 组) 中均经历基因谱系特异性的扩增。此外，B8 组包含单子叶植物水稻和高粱的基因，以及仅有的双子叶植物杨树的基因，这可能表明其他两个双子叶植物 (拟南芥和葡萄) 基因组中谱系特异性基因丢失。

1.3.4 *GT14* 和 *GT14-like* 基因的表达模式分析

为了研究拟南芥和杨树中 *GT14* 及 *GT14-like* 基因的功能多样性，我们使用公共微阵列数据对其组织特异性表达模式进行了分析。基于其表达模式的相似性，

拟南芥中的 *GT14* 和 *GT14-like* 基因被分为 6 个共表达簇 (图 1-2A), 例如, 簇 1 在根和茎中表现出优势; 簇 2 在叶中显示低水平的基因表达 (图 1-2A)。杨树 *GT14* 和 *GT14-like* 基因被分为两个共表达簇: 一个在花中显示优先表达 (即雌性和雄性两性), 另一个在木质部中优先表达 (图 1-2B)。大多数表达簇均包含 *GT14* 和

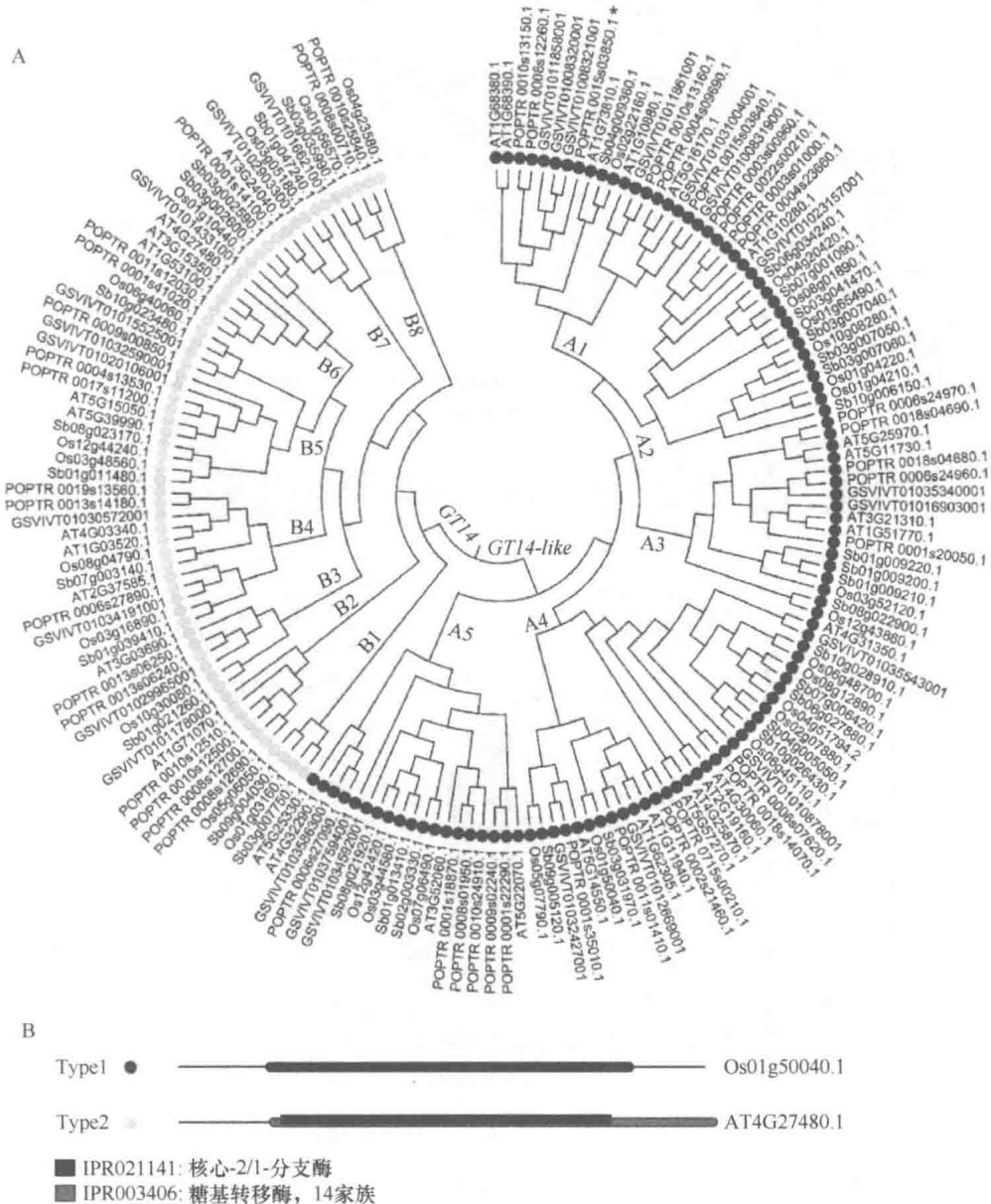


图 1-1 拟南芥、水稻、杨树、高粱、葡萄中的 *GT14* 和 *GT14-like* 基因的系统发育关系 (A) 及蛋白质结构域 (B) (彩图请扫封底二维码)

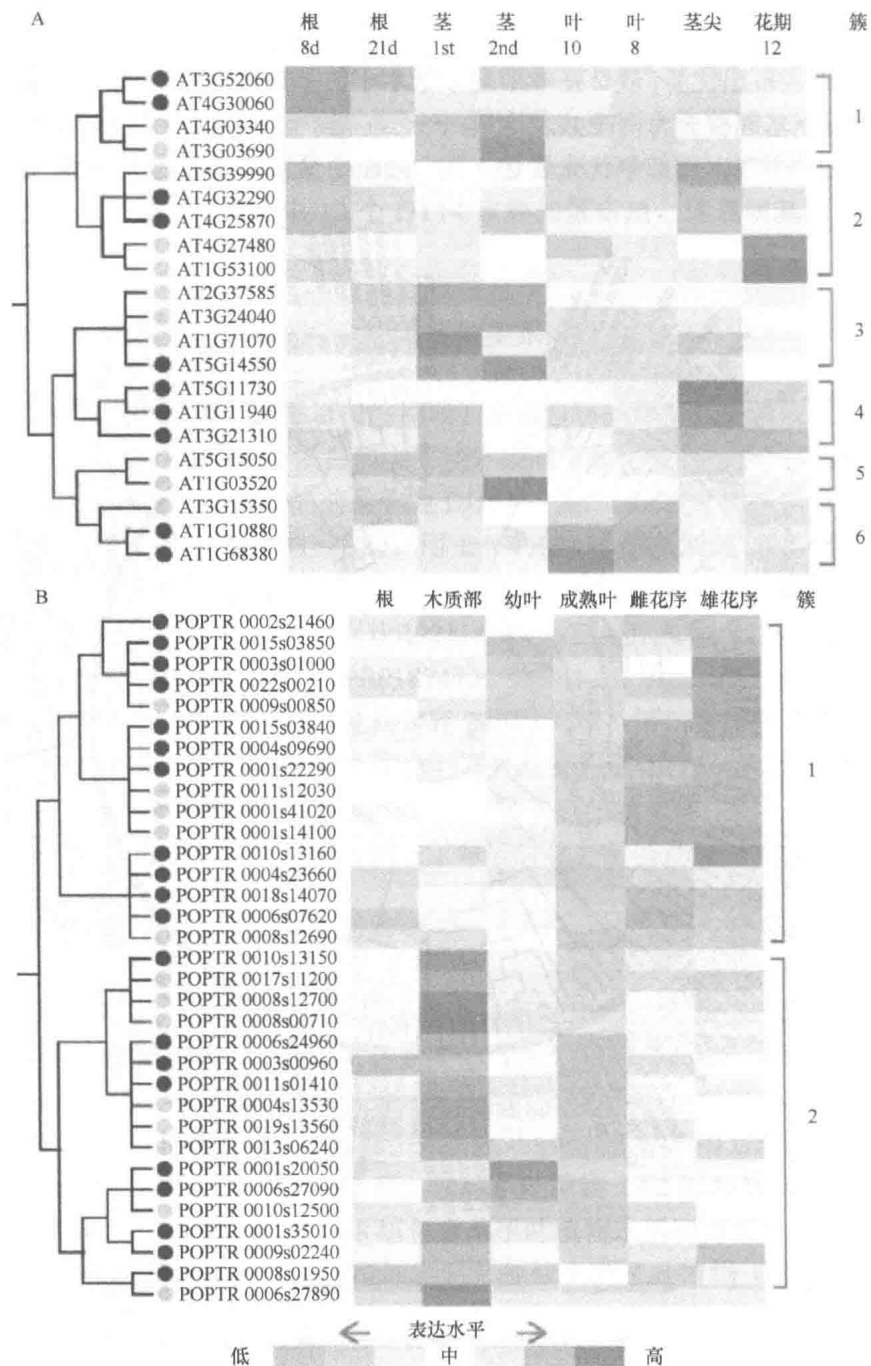


图 1-2 拟南芥 (A) 和杨树 (B) 中的 *GT14* (由黄色圆圈标记) 和 *GT14-like* (由红色圆圈标记) 基因的聚类 (彩图请扫封底二维码)

根 8d 表示第 8 天的根; 根 21d 表示第 21 天的根; 茎 1st 表示第一段茎; 茎 2nd 表示第二段茎;
叶 10 表示第 10 片叶子; 叶 8 表示第 8 片叶子; 花期 12 表示第 12 个花期