高等植物阿拉伯半乳糖蛋白的功能研究

马浩力, 余礼, 梁荣洪, 赵洁*

武汉大学生命科学学院, 杂交水稻国家重点实验室, 武汉 430072

* 联系人, E-mail: jzhao@whu.edu.cn

收稿日期: 2014-10-12; 接受日期: 2014-11-17; 网络版发表日期: 2015-02-03

国家自然科学基金(批准号: 31170171)和高等学校博士学科点专项科研基金优先发展领域项目(批准号: 20130141130008)资助

doi: 10.1360/N052014-00295

摘要 阿拉伯半乳糖蛋白(AGPs)是广泛分布于植物体内的一类富含羟脯氨酸糖蛋白的总称, 前人研究结果表明, 该类蛋白参与了高等植物生长发育的诸多生物学过程. 本文对高等植物 AGPs 基因家族的分类、AGPs 糖基化的氨基酸密码和 AGPs 鉴定的方法进行了系统的阐述, 并对在生殖发育、营养器官发生、植物与环境间的互作等方面 AGPs 基因的功能研究现状进行了概述, 对 AGPs 研究的前景进行了探讨, 希望为今后的相关研究带来启迪.

关键词

阿拉伯半乳糖蛋白
高等植物
生物学功能

阿拉伯半乳糖蛋白(arabinogalactan proteins, AGPs)由蛋白质骨架和富含阿拉伯糖和半乳糖的糖基侧链组成, 大多数情况下具有 C 端 GPI 锚——糖基磷酸肌醇锚(glycosylphosphatidylinositol anchor), 其糖基化程度较高, 一般分子量为 60~300 kD, 糖基侧链部分约占整个糖蛋白分子量的 90%以上^[1,2]. 由于 AGPs 糖基侧链中各种单糖和寡糖组分的含量和排列方式、蛋白质骨架的含量、氨基酸的种类和组成比例、以及糖基化位点的分布之间均存在一定的差异, 因此, AGPs 是一类异质性程度非常高的糖蛋白^[3]. 前人利用与 AGPs 特异结合的苯基偶氮染色剂(β -GlcY)、识别 AGPs 糖基表位的单克隆抗体以及一系列分子遗传学手段(如基因沉默、突变体和过量表达等)对拟南芥(*Arabidopsis thaliana*)、烟草(*Nicotiana tabacum*)和番茄(*Lycopersicon esculentum*)等高等植物中 AGPs 的分布和生物学功能进行了较为深入的研究, 发现此类蛋白在细胞增殖与生长、细胞分化、细胞程序性死

亡、体细胞胚胎发生、花药发育与花粉管生长、合子分裂和胚胎发育等过程中发挥着重要的功能. 因此, 研究高等植物 AGPs 蛋白的功能具有重要的理论意义和潜在的应用价值.

1 AGPs 的分类与鉴定

1.1 AGPs 的分类

根据 AGPs 核心蛋白骨架的不同, 广义上可将 AGPs 分为经典 AGPs(classical AGPs)和非经典 AGPs(non-classical AGPs) 2 类. 经典 AGPs 的蛋白骨架一般由 3 部分组成: N 端信号肽、长度可变且富含 PAST(脯氨酸 Pro, 丙氨酸 Ala, 丝氨酸 Ser 和苏氨酸 Thr)的区域和 C 端 GPI 锚定信号^[4]. 非经典 AGPs 在富含 PAST 区域之外具有富含天冬酰胺(Asn)或半胱氨酸(Cys)的区域, 部分不具有 C 端 GPI 锚定信号^[5]. AGPs 蛋白前体加工成为成熟蛋白的过程中, N 端信

引用格式: 马浩力, 余礼, 梁荣洪, 等. 高等植物阿拉伯半乳糖蛋白的功能研究. 中国科学: 生命科学, 2015, 45: 113–123

Ma H L, Yu L, Liang R H, et al. Functional studies of arabinogalactan proteins in higher plants. SCIENTIA SINICA Vitae, 2015, 45: 113–123, doi: 10.1360/N052014-00295

号肽被肽酶切除, 而C端疏水结构域将会被GPI锚代替^[3,5]。一类经典AGPs的蛋白骨架中富含PAST的区域被一段富含赖氨酸(Lys)的区域所间隔, 通常被称为赖氨酸富集的AGPs(Lysine-rich AGPs)^[6-8]; 另一类经典AGPs因其成熟蛋白骨架仅具有10~15 aa而被称为阿拉伯半乳糖多肽(AG peptides)。含有功能保守结构域的非经典AGPs常被称为嵌合AGPs(chimeric AGPs), 根据其保守结构域的不同分为3类: 类成束蛋白AGPs(fasciclin-like AGPs, FLAs)含有1~2个具有黏附作用的成束蛋白结构域^[9,10]; 类木质形成素AGPs(xylogen-like AGPs, XLAs)含有1个存在8个保守半胱氨酸的脂转运蛋白结构域^[11,12]; 类植物蓝素AGPs(phytocyanin-like AGPs, PLAs)含有1个与二价铜离子结合的植物蓝素结构域, 植物蓝素参与了光合作用中的电子传递, 因其氧化态为蓝色和还原态无色而得名^[13,14]。AGPs蛋白结构简图见图1。

综上所述, 高等植物中AGPs通常可分为7个亚类: 经典AGPs、非经典AGPs、赖氨酸富集AGPs、AG多肽、类成束蛋白AGPs、类木质形成素AGPs和类植物蓝素AGPs。

1.2 AGPs糖基化的氨基酸密码

AGPs蛋白骨架的一个典型特点是核心区域富含PAST, 并且这些氨基酸呈现出规律分布, 如多个Ala-Pro, Ser-Pro, Thr-Pro重复出现^[2]。前人创造性的

利用人工合成的肽段来测试不同类型的氨基酸排列方式会出现怎样的糖基化模式, 结果显示, 脯氨酸非连续分布的Ala-Pro和Ser-Pro重复序列中脯氨酸均被羟基化和糖基化, 且糖基侧链富含阿拉伯糖和半乳糖; 而脯氨酸连续分布的Ser-Pro_{3.5}重复序列中脯氨酸均被羟基化和糖基化, 且糖基侧链成分主要是寡聚阿拉伯糖^[15-17]。根据此类实验证据, 得出了一个羟脯氨酸连续假说: 如果蛋白骨架中Hyp连续分布(如Ser-Hyp_{3.5})则进行阿拉伯寡聚糖的糖基化, 而不连续的Hyp分布(如多个Ala-Hyp和Ser-Hyp重复)则进行阿拉伯半乳糖多聚糖的糖基化。目前, 在国际上有多个研究小组证实了这种假说的正确性。例如, 纯化表达于烟草BY-2悬浮细胞系中的Le-AGP1蛋白, 通过生化分析的结果证实了Hyp连续性假说的正确性^[18]。

尽管许多已知的AGPs包含XP重复(X代表A, S和T), 仍然有一些AGPs并不具有非连续的XP重复。例如, 被AG糖基化修饰的SOS5/FLA4(salt overly sensitive 5/fasciclin-like AGP 4)包含2个[AT]PPP基序^[19], AtAGP30含有3个PPAKAPIKLP重复序列^[20]。在sporamin能否被阿拉伯半乳糖糖基化的序列分析中, 发现有效糖基化的氨基酸序列符合[not basic]-[not T]-[AVSG]-Pro-[AVST]-[GAVPSTC]-[APS]公式^[21]。依据羟脯氨酸连续假说, 具有GPI锚定信号的拟南芥蛋白中40%都能被AG糖基化修饰^[22]。如果这种

图1 AGPs蛋白结构简图

预测是正确的, Extensin 蛋白、类 COBRA 蛋白、类 SKU5(skewed5)蛋白以及受体类蛋白均有可能被 AG 糖基化修饰, 即含有 1 个分泌信号且具有 AG 糖单元的蛋白质均可被视作类似 AGPs 的蛋白^[1].

1.3 AGPs 的鉴定

AGPs 的鉴定一般有 2 种方法:

(1) 从蛋白质到基因. (i) 利用 β -GlcY 试剂与大部分 AGPs 糖基侧链特异结合并形成沉淀的原理获得沉淀物, 利用洗涤和透析将 β -GlcY 除去; (ii) 利用氟化氢(HF)法去掉 AGPs 上的糖基侧链; (iii) 对 AGPs 蛋白骨架进行测序获得氨基酸序列, 这样就可以利用氨基酸密码子的原则从基因组中得到编码 AGPs 的基因(图 2A). 这种方法所获得的是最为直接的蛋白质证据, 此外, 还可以获得 AGPs 蛋白骨架中脯氨酸被羟基化的信息. 不足之处在于只能分离丰度较高且能与 β -GlcY 结合的 AGPs, 对于低丰度以及不能与 β -GlcY 结合的 AGPs 则难以获得. 前人采取这种方法从拟南芥中分离到 19 个 AGPs, 包括经典类、

多肽类、成束蛋白类和植物蓝素类^[4,9]; 从水稻(*Oryza sativa*)糊粉层细胞中分离到 6 个可与 β -GlcY 发生反应的 AGPs, 分别是 OsAGP1, OsAGPEP1, OsAGPEP2, OsAGPEP3, OsENDOL1 和 OsLTPL1^[23].

(2) 从基因到蛋白质. 利用 AGPs 蛋白骨架对 PAST 这 4 种氨基酸的偏好性并基于羟脯氨酸连续性假说(hyp contiguity hypothesis)从基因组中鉴定到可能编码 AGPs 的基因(图 2B). 通过计算 PAST 百分比, 筛选到具有信号肽且 PAST 含量>50%和 PAST 含量>35%(长度在 50~75 aa)的蛋白, 利用这种方法可以鉴定到经典 AGPs、赖氨酸富集的 AGPs、AG 多肽和非经典的 AGPs^[10,24]. 然而, 已知嵌合 AGPs 的 PAST 百分比通常为 30%~45%, 因此, 利用氨基酸偏好性鉴定诸如 FLAs, XLAs 和 PLAs 是非常困难的. 利用这 3 类蛋白分别含有类成束蛋白结构域、类脂转运蛋白结构域和类植物蓝素结构域的特点, 筛选具有保守结构域、并依据羟脯氨酸连续性假说鉴定可能编码 FLAs, XLAs 和 PLAs 的基因. 这种方法的优点是高通量的蛋白序列筛选得到大量可能编码 AGPs 的基

图 2 鉴定 AGPs 的 2 种方法
1,2,3 示不同的 AGPs

因, 缺点则是这些基因是否编码 AGP 蛋白需要进一步的验证.

拟南芥、水稻和玉米(*Zea mays*)等基因组序列的公布以及高通量RNA测序方法的发展使人们利用生物信息学方法从这些物种中预测AGPs基因家族成为可能. 截至目前, 根据AGPs蛋白骨架的特征预测在拟南芥和水稻基因组中分别存在111和117个AGPs^[10,12-14,24,25]. 棉花(*Gossypium spp*)和小麦(*Triticum aestivum*)中分别得到19和34个FLAs^[26,27]; 研究报道棉花中还存在GhH6L(*Gossypium hirsutum* H6 like protein)(经典类), GhAGP31(非经典类)和GhPLA1(植物蓝素类)等AGP蛋白^[28-30], 而小麦中则存在1个富含组氨酸的AGPs^[31]. 烟草是研究AGPs所选用的最早的植物物种, 但目前仅得到8个AGPs^[7,32-37]. 此外, 1992年Kieliszewski等人^[38]从玉米中鉴定得到了1个同时具备AGPs和Extensins结构特征的非经典AGPs. 几种高等植物中已知AGPs的种类见表1.

2 AGPs 的生物学功能

利用遗传学和分子生物学相关技术对烟草、番茄和拟南芥等物种中 AGPs 的生物学功能进行了研究, 发现 AGPs 主要在生殖发育过程(包括花粉发育、花粉管生长、自交不亲和和大孢子母细胞形成)、营养器官发生与发育(包括根和茎等器官)、植物与环境间的互作等方面起着重要作用(表 2).

2.1 在生殖发育中的作用

自 2002 年以来, 本实验室通过多种技术手段对拟南芥和烟草 AGPs 的分布和生物学功能进行了系统的研究. 研究表明, AGPs 在花药发育与花粉管生长、

合子第一次分裂和胚胎发育等过程中发挥着重要的功能^[71-76]. 烟草合子体外培养能进行正常的不对称分裂形成顶细胞和基细胞, 在培养基中添加 β -GlcY 时产生了异常的对称分裂的两细胞原胚, 表明 AGPs 在合子的第一次极性分裂过程中发挥着作用^[72]; 进一步的研究发现, AGPs 通过影响细胞分裂时细胞板的位置和细胞壁成分的分布来控制合子的第一次分裂^[76]. 利用抑制差减杂交技术对烟草非对称分裂和对称分裂的合子的表达变化在转录水平上进行了研究, 成功地筛选到合子非对称和对称分裂过程中差异表达基因^[77]. 利用识别 AGPs 糖基侧链的抗体进行免疫组化实验显示, AGPs 在烟草花药、花粉、花粉管和胚胎中分布, 利用 β -GlcY处理体外培养的胚胎, 发现 AGPs 参与胚胎的萌发、子叶形成、细胞壁纤维素和果胶沉积, 通过控制胚胎细胞分裂从而影响胚胎发育^[73-75].

高等植物生殖发育是一个极其复杂而精确的基因调控过程, 其中, AGP 类基因参与了花器官和种子形成过程中一系列重要的生物学事件. 拟南芥 *AGP6* 和 *AGP11* 基因在雄蕊、花粉和花粉管中特异性表达^[78], 在 *agp6 agp11* 双突变体植株中花粉完全败育; 通过 siRNA(small interfering RNA)和 amiRNA (artificial microRNA)介导的 RNAi(RNA interference)技术干扰 *AtAGP6* 和 *AtAGP11* 基因的表达, 导致花粉败育和花粉管生长受阻^[39]. 进一步研究证实 *AGP6* 和 *AGP11* 基因可能通过调节水分的吸收控制花粉的萌发^[40,41]. 拟南芥 *FLA3* 基因通过调节花粉内壁纤维素等物质沉积而影响花粉的正常发育和成熟花粉的育性, 表明 *FLA3* 基因在小孢子发育过程中尤其是花粉细胞壁发育过程中起着重要的作用^[55]. 最近的研究发现, 大白菜(*Brassica rapa pekinensis*)MF8(male

表 1 几种高等植物中的 AGPs^{a)}

类型	拟南芥	烟草	棉花	水稻	小麦	玉米
经典类	22	1	1	11	—	—
赖氨酸富集类	3	1	—	2	—	—
多肽类	16	—	—	15	—	—
成束蛋白类	21	—	19	27	34	—
植物蓝素类	30	—	1	38	—	—
木质形成素类	13	—	—	21	—	—
非经典类	6	6	1	3	1	1
合计	111	8	22	117	35	1

a) 小麦中组氨酸富集的 AGPs 在本文中作为非经典类; 水稻中木质形成素类的 AGPs 由本实验室马腾飞博士生等人于 2014 年完成(未发表的数据)

表 2 几种高等植物 AGPs 的生物学功能及相关信息

物种	蛋白名称	类型	遗传分析	植株表型	生物学功能	参考文献
拟南芥	AtAGP6 AtAGP11	经典	突变体和 RNAi	<i>agp6 agp11</i> 双突变体植株育性降低; 双 RNAi 植株花粉减少, 花粉管变短	花粉发育和花粉管生长	[39~41]
拟南芥	AtAGP17	赖氨酸富集	突变体	农杆菌对 <i>AtAGP17</i> 的 <i>rat1</i> 突变体根的转化效率降低	微生物抗性	[42,43]
拟南芥	AtAGP18	赖氨酸富集	RNAi 和过量表达	RNAi 植株育性降低, 大孢子母细胞不能进行有丝分裂; 在大孢子母细胞中过量表达 <i>AtAGP18</i> 基因导致产生多个功能大孢子	雌配子体发生	[44~46]
拟南芥	AtAGP19	赖氨酸富集	突变体	<i>agp19</i> 植株莲座叶变得又圆又小, 叶绿素含量降低, 生长缓慢, 下胚轴和茎变短, 角果和种子数量减少	营养生长和生殖发育	[47]
拟南芥	AtAGP24	多肽	<i>IDA</i> 过量表达	<i>AtAGP24</i> 基因在 <i>IDA(inflorescence deficient in abscission)</i> 过量表达植株的脱落部位表达量升高	器官脱落	[48,49]
拟南芥	AtAGP30	非经典	突变体	在脱落酸处理下 <i>agp30</i> 突变体萌发较快; <i>agp30</i> 悬浮细胞系无法分化形成根	种子萌发和根的开发	[20]
拟南芥	AtAGP31	非经典	突变体	对茉莉酸甲酯负调控, 在 <i>COI1(coronatine-insensitive 1)</i> 敲除突变体中表达下降	参与维管束发育与逆境应答	[50~52]
拟南芥	AtAGP57	经典	突变体	细胞壁成分发生改变	细胞壁发育与结构维持	[53]
拟南芥	AtFLA1	类成束蛋白	突变体	<i>fla1</i> 突变体侧根数目增加, 根和叶片的再生能力变弱	根的发育和细胞分化	[54]
拟南芥	AtFLA3	类成束蛋白	RNAi 和过量表达	RNAi 和过量表达植株育性均降低; RNAi 植株花粉内壁发育异常	花粉发育	[55]
拟南芥	SOS5/AtFLA4	类成束蛋白	突变体	<i>sos5</i> 突变体植株在一定盐离子浓度下根表皮细胞异常膨胀, 根长变短	盐离子耐受性和细胞膨胀	[19]
拟南芥	AtFLA11 AtFLA12	类成束蛋白	突变体	<i>fla11 fla12</i> 双突变体植株茎的机械强度下降; 细胞壁中纤维素、阿拉伯糖和半乳糖含量减少	次生细胞壁发育	[56~59]
拟南芥	AtXYP1 AtXYP2	木质形成素	突变体	<i>xyp1 xyp2</i> 双突变体的叶脉不连续, 木质部管状分子出现不正常的连接方式	维管束发育	[11,60]
烟草	NtTTS	非经典	反义抑制	反义抑制植株育性下降, 花粉管生长速度减慢	花粉管导向和生长	[61,62]
烟草	Na120 kD	非经典	RNAi	RNAi 植株不能排斥自体花粉	花粉自交不亲和	[63,64]
棉花	GhAGP4	类成束蛋白	RNAi	RNAi 植株棉纤维变短	棉纤维的起始和生长	[65]
棉花	GhFLA1	类成束蛋白	RNAi 过量表达	RNAi 植株棉花纤维的发生与伸长受到抑制; 过量表达植株棉花纤维的长度增加	棉纤维的起始和生长	[66]
黄瓜	CsAGP1	赖氨酸富集	过量表达	过量表达 <i>CsAGP1</i> 基因的烟草植株的株高增加	茎的生长	[67]
番茄	LeAGP1	赖氨酸富集	过量表达	过量表达 <i>GFP-LeAGP1</i> 植株的株高减少, 分支增多, 种子数量及大小减少	营养生长和生殖发育	[68]
番茄	attAGP	经典	RNAi	RNAi 植株中 <i>attAGP</i> 基因的表达量降低; 菟丝子(<i>Cuscuta chinensis</i>)的黏附效率明显下降	寄生植物黏附	[69]
大白菜	BcMF8	经典	RNAi	RNAi 植株中花粉形态异常, 萌发率降低	花粉发育	[70]

fertility 8)通过控制花粉内壁和萌发孔的形成来影响花粉的发育,当抑制该基因表达后,花粉形态异常,萌发率降低^[70].

在大叶烟草花柱引导组织中特异表达的蛋白(transmitting tissue specific proteins, TTS)是一种花粉

管生长促进蛋白,为花粉管提供营养物质和表面黏附的作用;在引导组织中 TTS 蛋白的糖基化梯度可能与花粉管的定向生长有关^[61,62]. 研究人员发现,自交不亲和现象与一类存在于花柱引导组织中的 S-RNase 糖蛋白相关, S-RNase 进入到不亲和花粉中

产生毒性作用. 烟草 120 kD 糖蛋白兼有 AGPs 和伸展蛋白的属性并且能与 S-RNase 结合, 该蛋白基因的 RNAi 株系中自交不亲和和受到了抑制, 证实 AGPs 参与了花粉识别过程^[63,64].

在拟南芥孢子体向配子体转化过程以及早期胚胎中 *AtAGP18* 基因特异表达, 其 RNAi 植株中功能大孢子不能进行正常的有丝分裂, 证实 *AtAGP18* 在雌配子发生过程中起着关键性的作用^[44]. 过量表达 *AtAGP18* 基因导致拟南芥植株变矮, 形态异常, 角果变短^[45]; 在大孢子母细胞中增强 *AtAGP18* 基因的表达导致 1 个胚珠中能够产生 2~4 个功能大孢子, 进而导致植株败育^[46]. 运用 *GUS*(β -glucuronidase) 和 *GFP*(green fluorescent protein) 报告基因、以及荧光原位杂交技术(fluorescence *in situ* hybridization, FISH) 对拟南芥中多个 AGP 基因的定位进行研究, 发现其主要在生殖器官中表达, 表明 AGP 在生殖发育过程中发挥着重要作用^[79].

2.2 在器官发生中的作用

拟南芥 *APAPI*(*AtAGP57C*), 是一个能与细胞壁成分果胶和阿拉伯木聚糖共价连接的经典 AGP, 敲除该基因后, 细胞壁成分发生改变, 证明该基因在细胞壁的形成与形态维持上有重要功能^[53]. 在拟南芥厚壁组织细胞中 *AtFLA11* 和 *AtFLA12* 基因表达^[56~58]; *fla11* 突变体中木质部存在微弱缺陷, 而 *fla11 fla12* 双突变体中茎的弹性系数和机械强度都下降, 这与细胞壁的结构和成分发生了变化相关, 即纤维素微纤丝的角度增大且纤维素含量下降^[59]. *AtFLA1* 基因在愈伤组织和幼苗中表达, *fla1* 突变体侧根数目增加, 根和叶片的再生能力减弱, 表明它在根的发育和细胞分化过程起重要作用^[54]. 在棉花纤维中 *GhFLA1* 基因大量表达, 且过量表达该基因导致棉花纤维的长度增加; 在 RNAi 植株中棉花纤维的发生与伸长受到抑制并且细胞壁的糖成分发生了变化, 表明 *GhFLA1* 参与了棉纤维的发生与伸长^[66]. *GhAGP4* 是一个 FLA 类基因, 该基因的 RNAi 植株中棉纤维变短, 质量下降, 纤维素积累受到影响, 推测 *GhAGP4* 也可能参与棉花纤维的发生与伸长^[65].

前人研究证明, AGPs 在拟南芥 *IDA*(inflorescence deficient in abscission) 基因控制的器官脱落过程中发挥作用^[48,49]. *ida* 突变体花器官脱落延迟, 而过表达该基因则导致花器官提前脱落, 并在脱落区产生大量

能与 β -GlcY 反应的分泌物. 在 *ida* 突变体的脱落区中, AGPs 基因的表达水平低于野生型, *IDA* 可能是器官脱落和诱导脱落区分泌 AGPs 的正调控因子. 其中, *AtAGP24* 基因在突变体和超表达植株中的表达水平发生了变化, 表明该基因可能参与了器官的脱落过程. 番茄 *LeAGP-1* 是一个具有 GPI 锚的经典 AGPs, 其基因的过量表达植株较野生型矮小, 分支多, 花蕾增多, 但许多花蕾不能成熟而导致产量下降, 种子较小, 暗示过量表达该基因破坏了生长素与细胞分裂素之间的平衡. 然而, 过表达缺失 GPI 锚的 *LeAGP-1* 植株没有产生特殊表型, 表明 GPI 锚对于 *LeAGP-1* 的定位和功能至关重要^[68]. *LeAGP1* 在其他物种中的同源基因同样参与植株的生长与发育, 烟草中过量表达黄瓜(*Cucumis sativus*) *CsAGP1* 基因, 促进了茎的生长^[67]. 拟南芥 *AtAGP19* 基因的 T-DNA 插入突变体 *atagp19* 与野生型相比, 莲座叶较小、较圆、较扁平, 叶绿素含量减少, 植株发育延迟, 下胚轴和主枝较短, 角果量较少且育性降低, 证实 *AtAGP19* 基因在细胞分裂与生长、叶的发育以及生殖发育过程中起着重要的作用^[47]. 在拟南芥根中, *AtAGP30* 基因特异表达, 过量表达 *AtAGP30* 不影响根的正常发育, 但是完全扰乱了芽的发育. 对 *atagp30* 突变体分析发现, 该基因对根的再生以及种子的萌发起着重要的调控作用, 抑制了由脱落酸(abscisic acid, ABA)引起的种子萌发的延迟, 并影响了与 ABA 功能相关基因的表达, 推测 *AtAGP30* 基因可能参与了 ABA 的应答^[20]. 非经典 AGP 基因 *AtAGP31* 优势表达于维管束组织, 而且被茉莉酸甲酯负调控, 推测该基因可能与细胞壁组分相互作用, 参与逆境应答与维管束的发育^[50~52]. AGPs 可能作为受细胞分裂因子调控下游靶基因, 也有可能通过调节受体的敏感性或作为促使调控因子与受体结合的共受体来发挥作用, 然而, 这些推测还需进一步证实.

木质形成素是一个嵌合的 AGPs, 包含类脂转运蛋白的区域、类 AGPs 的区域和 GPI 锚定信号. 它的分子量为 50~100 kD, 远大于预测值 16 kD. 此外, 木质形成素能够与 β -GlcY 反应并被 JIM13 抗体识别也证明它是一个 AGP 蛋白. 拟南芥中存在 2 个与木质形成素同源的基因 *AtXYP1* 和 *AtXYP2*, 其双突变体植株维管组织发育出现异常, 叶脉不连续, 导管细胞之间的连接混乱, 表明该类 AGPs 与植物器官形成和发育密切相关^[11,60].

2.3 在植物与环境间的互作中的作用

在 T-DNA 插入突变体 *rat1*(resistant to *Agrobacterium transformation* 1) 中 *AtAGP17* 基因表达量下降; *rat1* 的根对农杆菌转化具有抗性; 利用 β -GlcY 试剂对野生型植株的根进行处理也同样产生抗性反应, 推测 *AtAGP17* 可能通过影响农杆菌与根的结合活力从而影响农杆菌侵染时根部的信号转导^[42,43]. 在 *rat1* 突变体中, 依赖于水杨酸的获得性免疫标记基因的表达量均升高, 表明该基因的突变激活了植物体内水杨酸依赖的抗逆信号通路^[43]. 拟南芥 *SOS5/FLA4* (Salt Overly Sensitive 5) 基因编码一个与细胞表面黏附作用相关的 FLA 蛋白, 在维持根细胞形态过程中起着重要作用. 在盐离子的胁迫下, 点突变的 *sos5* 突变体根尖膨大且根的伸长受到影响, 细胞壁变薄^[19]; 并且 *AtFLA4* 能与 ABA 协同作用参与 ABA 信号通路, 影响根的发育^[80]. 当寄生植物菟丝子黏附在番茄植株上时, 诱导番茄植株受伤位点产生一种新的 AGP(attAGP); 利用 RNAi 技术发现, 如果 *attAGP* 基因表达量降低, 菟丝子的黏附效率明显下降. 因此, 推测番茄 *attAGP* 基因的表达促进了寄生植物与宿主的黏附^[69].

3 展望

近 50 年来, 高等植物 AGPs 的研究使人们对这类糖蛋白的重要性有了进一步的认识. 但是, AGPs

在植物体内是通过支链糖基还是蛋白核心发挥调节或控制作用目前还不清楚, 其分子机理一直是植物发育研究中的未解之谜. AGPs 的作用机理仍需要深入的研究, 未来对于高等植物 AGPs 的研究应该着重于以下 4 个方面: (i) 基因组时代的来临使得研究者们使用生物信息学手段鉴定编码 AGPs 的基因成为了可能, 随着越来越多的物种全基因组序列的公布, 鉴定除拟南芥和水稻之外的已测序物种(如玉米、棉花等)中编码 AGPs 的基因家族将是一个重要的方向; 继续丰富对羟脯氨酸连续假说的验证, 更深入地了解 AG 糖基化的氨基酸密码; 在蛋白质组学和生物化学层面寻找预测的 AGPs 基因存在的蛋白和糖生物学证据; (ii) 利用遗传学和分子生物学手段鉴定水稻、玉米、小麦等重要农作物中 AGPs 基因的生物学功能. 在这些物种中 AGPs 的研究尚集中在基因的鉴定阶段, 其功能研究的空白留待研究者们填补; (iii) AGPs 功能的实现多依赖于 AG 糖基侧链的存在, 解析 AG 糖基侧链的分子结构和糖基侧链形成的酶动力学等机制将是主要的挑战, 研究 AGPs 相关的糖基转移酶将成为解决这一问题的关键; (iv) 大多数的 AGPs 由于存在 GPI 锚定信号而定位于细胞膜上, 研究证据显示, 具有 GPI 锚的 AGPs 可能参与了细胞间的通讯和信号转导, 未来研究的主要挑战是寻找与 AGPs 相互作用的蛋白或其他因子, 从而揭示这类糖蛋白的功能和作用机制.

参考文献

- Seifert G J, Roberts K. The biology of arabinogalactan proteins. *Annu Rev Plant Biol*, 2007, 58: 137–161
- Ellis M, Egelund J, Schultz C J, et al. Arabinogalactan-proteins: key regulators at the cell surface? *Plant Physiol*, 2010, 153: 403–419
- Gaspar Y, Johnson K L, McKenna J A, et al. The complex structures of arabinogalactan-proteins and the journey towards understanding function. *Plant Mol Biol*, 2001, 47: 161–176
- Schultz C J, Johnson K L, Currie G, et al. The classical arabinogalactan protein gene family of *Arabidopsis*. *Plant Cell*, 2000, 12: 1751–1768
- Du H, Clarke A E, Bacic A. Arabinogalactan-proteins: a class of extracellular matrix proteoglycans involved in plant growth and development. *Trends Cell Biol*, 1996, 6: 411–414
- Li S X, Showalter A M. Cloning and developmental stress-regulated expression of a gene encoding a tomato arabinogalactan-proteins. *Plant Mol Biol*, 1996, 32: 641–652
- Gilson P, Gaspar Y M, Oxley D, et al. NaAGP4 is an arabinogalactan protein whose expression is suppressed by wounding and fungal infection in *Nicotiana glauca*. *Protoplasma*, 2001, 215: 128–139
- Sun W, Xu J, Yang J, et al. The lysine-rich arabinogalactan-protein subfamily in *Arabidopsis*: gene expression, glycoprotein purification and biochemical characterization. *Plant Cell Physiol*, 2005, 46: 975–984
- Johnson K L, Jones B J, Bacic A, et al. The fasciclin-like arabinogalactan proteins of *Arabidopsis*. A multigene family of putative cell

- adhesion molecules. *Plant Physiol*, 2003, 133: 1911–1925
- 10 Ma H, Zhao J. Genome-wide identification, classification, and expression analysis of the arabinogalactan protein gene family in rice (*Oryza sativa* L.). *J Exp Bot*, 2010, 61: 2647–2668
- 11 Motose H, Sugiyama M, Fukuda H. A proteoglycan mediates inductive interaction during plant vascular development. *Nature*, 2004, 429: 873–878
- 12 Kobayashi Y, Motose H, Iwamoto K, et al. Expression and genome-wide analysis of the xylogen-type gene family. *Plant Cell Physiol*, 2011, 52: 1095–1106
- 13 Mashiguchi K, Asami T, Suzuki Y. Genome-wide identification, structure and expression studies, and mutant collection of 22 early nodulin-like protein genes in *Arabidopsis*. *Biosci Biotechnol Biochem*, 2009, 73: 2452–2459
- 14 Ma H, Zhao H, Liu Z, et al. The phytoeyanin gene family in rice (*Oryza sativa* L.): genome-wide identification, classification and transcriptional analysis. *PLoS One*, 2011, 6: e25184
- 15 Kieliszewski M J, Lampion D T. Extensin: repetitive motifs, functional sites, post-translational codes and phylogeny. *Plant J*, 1994, 5: 157–172
- 16 Shpak E, Leykam J F, Kieliszewski M J. Synthetic genes for glycoprotein design and the elucidation of hydroxyproline-O-glycosylation codes. *Proc Natl Acad Sci USA*, 1999, 96: 14736–14741
- 17 Shpak E, Barbar E, Leykam J F, et al. Contiguous hydroxyproline residues direct hydroxyproline arabinosylation in *Nicotiana tabacum*. *J Biol Chem*, 2001, 276: 11272–11278
- 18 Zhao Z D, Tan L, Showalter A M, et al. Tomato LeAGP-1 arabinogalactan-protein purified from transgenic tobacco corroborates the Hyp contiguity hypothesis. *Plant J*, 2002, 31: 431–444
- 19 Shi H, Kim Y, Guo Y, et al. The *Arabidopsis* *SOS5* locus encodes a putative cell surface adhesion protein and is required for normal cell expansion. *Plant Cell*, 2003, 15: 19–32
- 20 van Hengel A J, Roberts K. AtAGP30, an arabinogalactan-protein in the cell walls of the primary root, plays a role in root regeneration and seed germination. *Plant J*, 2003, 36: 256–270
- 21 Shimizu M, Igasaki T, Yamada M, et al. Experimental determination of proline hydroxylation and hydroxyproline arabinogalactosylation motifs in secretory proteins. *Plant J*, 2005, 42: 877–889
- 22 Borner G H, Sherrier D J, Stevens T J, et al. Prediction of glycosylphosphatidylinositol-anchored proteins in *Arabidopsis*. A genomic analysis. *Plant Physiol*, 2002, 129: 486–499
- 23 Mashiguchi K, Yamaguchi I, Suzuki Y. Isolation and identification of glycosylphosphatidylinositol-anchored arabinogalactan proteins and novel β -glucosyl Yariv-reactive proteins from seeds of rice (*Oryza sativa*). *Plant Cell Physiol*, 2004, 45: 1817–1829
- 24 Schultz C J, Rumsewicz M P, Johnson K L, et al. Using genomic resources to guide research directions. The arabinogalactan protein gene family as a test case. *Plant Physiol*, 2002, 129: 1448–1463
- 25 Showalter A M, Keppler B, Lichtenberg J, et al. A bioinformatics approach to the identification, classification, and analysis of hydroxyproline-rich glycoproteins. *Plant Physiol*, 2010, 153: 485–513
- 26 Huang G Q, Xu W L, Gong S Y, et al. Characterization of 19 novel cotton *FLA* genes and their expression profiling in fiber development and inresponse to phytohormones and salt stress. *Physiol Plant*, 2008, 134: 348–359
- 27 Faik A, Abouzouhair J, Sarhan F. Putative fasciclin-like arabinogalactan-proteins (FLA) in wheat (*Triticum aestivum*) and rice (*Oryza sativa*): identification and bioinformatic analyses. *Mol Genet Genomics*, 2006, 276: 478–494
- 28 Wu Y, Xu W, Huang G, et al. Expression and localization of GhH6L, a putative classical arabinogalactan protein in cotton (*Gossypium hirsutum*). *Acta Biochim Biophys Sin*, 2009, 41: 495–503
- 29 Gong S Y, Huang G Q, Sun X, et al. GhAGP31, a cotton non-classical arabinogalactan protein, is involved in response to cold stress during early seedling development. *Plant Biol*, 2012, 14: 447–457
- 30 Poon S, Heath R L, Clarke A E. A chimeric arabinogalactan protein promotes somatic embryogenesis in cotton cell culture. *Plant Physiol*, 2012, 160: 684–695
- 31 Aizat W M, Preuss J M, Johnson A A, et al. Investigation of a His-rich arabinogalactan-protein for micronutrient biofortification of cereal grain. *Physiol Plant*, 2011, 143: 271–286
- 32 Youl J J, Bacic A, Oxley D. Arabinogalactan-proteins from *Nicotiana glauca* and *Pyrus communis* contain glycosylphosphatidylinositol membrane anchors. *Proc Natl Acad Sci USA*, 1998, 95: 7921–7926
- 33 Mau S L, Chen C G, Pu Z Y, et al. Molecular cloning of cDNAs encoding the protein backbones of arabinogalactan-proteins from the filtrate of suspension-cultured cells of *Pyrus communis* and *Nicotiana glauca*. *Plant J*, 1995, 8: 269–281
- 34 Du H, Simpson R J, Clarke A E, et al. Molecular characterization of a stigma-specific gene encoding an arabinogalactan-protein (AGP) from *Nicotiana glauca*. *Plant J*, 1996, 9: 313–323

- 35 Schultz C J, Hauser K, Lind J L, et al. Molecular characterisation of a cDNA sequence encoding the backbone of a style-specific 120 kD glycoprotein which has features of both extensins and arabinogalactan proteins. *Plant Mol Biol*, 1997, 35: 833–845
- 36 Bosch M, Knudsen J S, Derksen J, et al. Class III pistil-specific extensin-like proteins from tobacco have characteristics of arabinogalactan proteins. *Plant Physiol*, 2001, 125: 2180–2188
- 37 Cheung A Y, May B, Kawata E E, et al. Characterization of cDNAs for stylar transmitting tissue-specific proline-rich proteins in tobacco. *Plant J*, 1993, 3: 151–160
- 38 Kieliszewski M J, Kamyab A, Leykam J F, et al. A histidine-rich extensin from *Zea mays* is an arabinogalactan protein. *Plant Physiol*, 1992, 99: 538–547
- 39 Levitin B, Richter D, Markovich I, et al. Arabinogalactan proteins 6 and 11 are required for stamen and pollen function in *Arabidopsis*. *Plant J*, 2008, 56: 351–363
- 40 Coimbra S, Costa M, Jones B, et al. Pollen grain development is compromised in *Arabidopsis agp6 agp11* null mutants. *J Exp Bot*, 2009, 60: 3133–3142
- 41 Coimbra S, Costa M, Mendes M A, et al. Early germination of *Arabidopsis* pollen in a double null mutant for the arabinogalactan protein genes *AGP6* and *AGP11*. *Sex Plant Reprod*, 2010, 23: 199–205
- 42 Nam J, Mysore K S, Zheng C, et al. Identification of T-DNA tagged *Arabidopsis* mutants that are resistant to transformation by agrobacterium. *Mol Gen Genet*, 1999, 261: 429–438
- 43 Gaspar Y, Nam J, Schultz C J, et al. Characterization of the *Arabidopsis* lysine-rich arabinogalactan-protein AtAGP17 mutant (*rat1*) that results in a decreased efficiency of agrobacterium transformation. *Plant Physiol*, 2004, 135: 2162–2171
- 44 Acosta-Garcia G, Vielle-Calzada J P. A classical arabinogalactan protein is essential for the initiation of female gametogenesis in *Arabidopsis*. *Plant Cell*, 2004, 16: 2614–2628
- 45 Zhang Y, Yang J, Showalter A M. AtAGP18 is localized at the plasma membrane and functions in plant growth and development. *Planta*, 2011, 233: 675–683
- 46 Demesa-Arevalo E, Vielle-Calzada J P. The classical arabinogalactan protein AGP18 mediates megaspore selection in *Arabidopsis*. *Plant Cell*, 2013, 25: 1274–1287
- 47 Yang J, Sardar H S, McGovern K R, et al. A lysine-rich arabinogalactan protein in *Arabidopsis* is essential for plant growth and development, including cell division and expansion. *Plant J*, 2007, 49: 629–640
- 48 Butenko M A, Patterson S E, Grini P E, et al. Inflorescence deficient in abscission controls floral organ abscission in *Arabidopsis* and identifies a novel family of putative ligands in plants. *Plant Cell*, 2003, 15: 2296–2307
- 49 Stenvik G E, Butenko M A, Urbanowicz B R, et al. Overexpression of *INFLORESCENCE DEFICIENT IN ABSCISSION* activates cell separation in vestigial abscission zones in *Arabidopsis*. *Plant Cell*, 2006, 18: 1467–1476
- 50 Liu C, Mehdy M C. A nonclassical arabinogalactan protein gene highly expressed in vascular tissues, *AGP31*, is transcriptionally repressed by methyl jasmonic acid in *Arabidopsis*. *Plant Physiol*, 2007, 145: 863–874
- 51 Hijazi M, Durand J, Pichereaux C, et al. Characterization of the arabinogalactan protein 31 (AGP31) of *Arabidopsis thaliana*: new advances on the Hyp-O-glycosylation of the Pro-rich domain. *J Biol Chem*, 2012, 287: 9623–9632
- 52 Hijazi M, Roujol D, Nguyen-Kim H, et al. Arabinogalactan protein 31 (AGP31), a putative network-forming protein in *Arabidopsis thaliana* cell walls? *Ann Bot*, 2014, 114: 1087–1097
- 53 Tan L, Eberhard S, Pattathil S, et al. An *Arabidopsis* cell wall proteoglycan consists of pectin and arabinoxylan covalently linked to an arabinogalactan protein. *Plant Cell*, 2013, 25: 270–287
- 54 Johnson K L, Kibble N A, Bacic A, et al. A fasciclin-like arabinogalactan-protein (FLA) mutant of *Arabidopsis thaliana*, *fla1*, shows defects in shoot regeneration. *PLoS One*, 2011, 6: e25154
- 55 Li J, Yu M, Geng L L, et al. The fasciclin-like arabinogalactan protein gene, *FLA3*, is involved in microspore development of *Arabidopsis*. *Plant J*, 2010, 64: 482–497
- 56 Brown D M, Zeef L A, Ellis J, et al. Identification of novel genes in *Arabidopsis* involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell*, 2005, 17: 2281–2295
- 57 Ito S, Suzuki Y, Miyamoto K, et al. AtFLA11, a fasciclin-like arabinogalactan-protein, specifically localized in sclerenchyma cells. *Biosci Biotechnol Biochem*, 2005, 69: 1963–1969
- 58 Persson S, Wei H, Milne J, et al. Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proc Natl Acad Sci USA*, 2005, 102: 8633–8638
- 59 MacMillan C P, Mansfield S D, Stachurski Z H, et al. Fasciclin-like arabinogalactan proteins: specialization for stem biomechanics and cell wall architecture in *Arabidopsis* and *Eucalyptus*. *Plant J*, 2010, 62: 689–703
- 60 Motose H, Sugiyama M, Fukuda H. An arabinogalactan protein(s) is a key component of a fraction that mediates local intercellular

- communication involved in tracheary element differentiation of *Zinnia* mesophyll cells. *Plant Cell Physiol*, 2001, 42: 129–137
- 61 Cheung A Y, Wang H, Wu H M. A floral transmitting tissue-specific glycoprotein attracts pollen tubes and stimulates their growth. *Cell*, 1995, 82: 383–393
- 62 Wu H M, Wang H, Cheung A Y. A pollen tube growth stimulatory glycoprotein is deglycosylated by pollen tubes and displays a glycosylation gradient in the flower. *Cell*, 1995, 82: 395–403
- 63 Hancock C N, Kent L, McClure B A. The stylar 120 kD glycoprotein is required for S-specific pollen rejection in *Nicotiana*. *Plant J*, 2005, 43: 716–723
- 64 Lee C B, Swatek K N, McClure B. Pollen proteins bind to the C-terminal domain of *Nicotiana alata* pistil arabinogalactan proteins. *J Biol Chem*, 2008, 283: 26965–26973
- 65 Li Y, Liu D, Tu L, et al. Suppression of *GhAGP4* gene expression repressed the initiation and elongation of cotton fiber. *Plant Cell Rep*, 2010, 29: 193–202
- 66 Huang G Q, Gong S Y, Xu W L, et al. A fasciclin-like arabinogalactan protein, GhFLA1, is involved in fiber initiation and elongation of cotton. *Plant Physiol*, 2013, 161: 1278–1290
- 67 Park M H, Suzuki Y, Chono M, et al. *CsAGPI*, a gibberellin responsive gene from cucumber hypocotyls, encodes a classical arabinogalactan protein and is involved in stem elongation. *Plant Physiol*, 2003, 131: 1450–1459
- 68 Sun W, Kieliszewski M J, Showalter A M. Overexpression of tomato LeAGP-1 arabinogalactan-protein promotes lateral branching and hampers reproductive development. *Plant J*, 2004, 40: 870–881
- 69 Albert M, Belastegui-Macadam X, Kaldenhoff R. An attack of the plant parasite *Cuscuta reflexa* induces the expression of attAGP, an attachment protein of the host tomato. *Plant J*, 2006, 48: 548–556
- 70 Lin S, Dong H, Zhang F, et al. *BcMF8*, a putative arabinogalactan protein-encoding gene, contributes to pollen wall development, aperture formation and pollen tube growth in *Brassica campestris*. *Ann Bot*, 2014, 113: 777–788
- 71 Hu Y, Qin Y, Zhao J. Localization of arabinogalactan protein epitope and the effects of Yariv phenylglycoside during zygotic embryo development of *Arabidopsis thaliana*. *Protoplasma*, 2006, 229: 21–31
- 72 Qin Y, Zhao J. Localization of arabinogalactan proteins in egg cells, zygotes, and two-celled proembryos and effects of β -D-glucosyl Yariv reagent on egg cell fertilization and zygote division in *Nicotiana tabacum* L. *J Exp Bot*, 2006, 57: 2061–2074
- 73 Qin Y, Chen D, Zhao J. Localization of arabinogalactan proteins in anther, pollen, and pollen tube of *Nicotiana tabacum* L. *Protoplasma*, 2007, 231: 43–53
- 74 Qin Y, Zhao J. Localization of arabinogalactan-proteins in different stages of embryos and their role in cotyledon formation of *Nicotiana tabacum* L. *Sex Plant Reprod*, 2007, 20: 213–224
- 75 Zhong J, Ren Y, Yu M, et al. Roles of arabinogalactan proteins in cotyledon formation and cell wall deposition during embryo development of *Arabidopsis*. *Protoplasma*, 2011, 248: 551–563
- 76 Yu M, Zhao J. The cytological changes of tobacco zygote and proembryo cells induced by β -glucosyl Yariv reagent suggest the involvement of arabinogalactan proteins in cell division and cell plate formation. *BMC Plant Biol*, 2012, 12: 126
- 77 Hu T X, Yu M, Zhao J. Comparative transcriptional profiling analysis of the two daughter cells from tobacco zygote reveals the transcriptome differences in the apical and basal cells. *BMC Plant Biol*, 2010, 10: 167
- 78 Pereira L G, Costa M, Coimbra S. Localization of arabinogalactan protein 6 fused with Sirius ultramarine fluorescent protein in *Arabidopsis* pollen and pollen tubes. *Plant Signal Behav*, 2013, 8: e25998
- 79 Pereira A M, Masiero S, Nobre M S, et al. Differential expression patterns of arabinogalactan proteins in *Arabidopsis thaliana* reproductive tissues. *J Exp Bot*, 2014, 65: 5459–5471
- 80 Seifert G J, Xue H, Acet T. The *Arabidopsis thaliana* *FASCICLIN LIKE ARABINOGALACTAN PROTEIN 4* gene acts synergistically with abscisic acid signalling to control root growth. *Ann Bot*, 2014, 114: 125–133

Functional Studies of Arabinogalactan Proteins in Higher Plants

MA HaoLi, YU Li, LIANG RongHong & ZHAO Jie

State Key Laboratory of Hybrid Rice, College of Life Sciences, Wuhan University, Wuhan 430072, China

Arabinogalactan proteins (AGPs) are a kind of hydroxyproline-rich glycoproteins in plants and implicated in various processes of growth and development. In this study, we firstly conducted a comprehensive elaboration on the identification, classification and glycosylation (amino acid code) of AGPs. Moreover, we summarized the research advances on gene functions of AGPs in reproductive development, vegetative organogenesis and interactions between plant and environment. The prospect of research on AGPs is also discussed in this paper and would give new insights toward the related research in the future.

arabinogalactan proteins, higher plants, biological function

doi: 10.1360/N052014-00295