

铁皮石斛 *phytoeyanin* 基因家族全基因组分析

宋 争, 李潞滨, 梁立雄, 王 涛*

(林木遗传育种国家重点实验室, 国家林业局林木培育重点实验室, 中国林业科学研究院林业研究所, 北京 100091)

摘要: [目的] 探讨 *Phytoeyanins* (PCs) 在铁皮石斛发育过程和逆境胁迫环境下的潜在功能。 [方法] 利用拟南芥和水稻 *phytoeyanin* 基因家族蛋白序列 (AtPCs 和 OsPCs), 采用本地化软件 BLASTP 对铁皮石斛全基因组数据库进行搜索, 并采用 SMART 和 pfam 数据库验证质体蓝素样结构域, 获得铁皮石斛 *phytoeyanin* 基因家族编码序列 (DoPCs); 利用 Signal P4.1、Big-PI Plant Predictor、NetNGlyc1.0 Server 等在线软件对 DoPC 家族序列的结构特点进行分析; 利用软件 ClustalW 和 MEGA 对 DoPC 基因进行氨基酸序列的比对和系统进化分析; 利用转录组数据绘制 heatmap 图对 DoPCs 基因在与美孢胶膜菌共生萌发的种子中的表达情况进行分析。 [结果] 显示: 铁皮石斛全基因组中共预测到包含 PCLD 结构域的 38 个 DoPCs, 可分属 4 个亚家族, 2 个 Cys 残基在 DoPCs 家族中高度保守; 19 条 DoPCs 蛋白序列含有 4 个完整的保守铜结合配体 (His、Cys、His、Gln/Met), 有 32 条 DoPCs 骨架中含有 N-端信号肽, 22 条 DoPCs 含有糖基磷脂酰肌醇锚定位点, 有 27 条 DoPCs 含 N-糖基化位点; 16 个 DoPCs 基因在与美孢胶膜菌共生萌发的种子中检测到表达, 其中, *DoUCL2*, 4 和 *DoENODL14* 基因表达量最高, 7 个基因在共生萌发的种子中明显上调表达, 只有 *DoENODL6* 基因明显下调表达。 [结论] 在全基因组范围内分析铁皮石斛 DoPC 家族蛋白结构特点, 可以为探究兰科植物与微生物共生的相互作用及进一步研究兰科菌根形成的分子机制提供数据基础。

关键词: 铁皮石斛; *phytoeyanin* 基因家族; 全基因组分析; 共生萌发种子

中图分类号: S718.46

文献标识码: A

文章编号: 1001-1498(2018)02-0098-09

Genome-wide Analysis of *Phytoeyanin* Gene Family in *Dendrobium officinale*

SONG Zheng, LI Lu-bin, LIANG Li-xiong, WANG Tao

(State Key Laboratory of Tree Genetic and Breeding, Key Laboratory of Tree Breeding and Cultivation, State Forestry Administration, Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, China)

Abstract: [Objective] To study the sequence characteristics of *Dendrobium officinale* [Method] Local BLASTP was used to search the *phytoeyanin* gene of *D. officinale* (DoPCs) in the database of *Dendrobium* whole genome, and then the conserved Plastocyanin-like domain and structural characteristics of DoPCs were predicted. The sequence alignment and phylogenetic analysis of DoPCs were conducted by using Clustal W and MEGA. Finally, 38 DoPCs were expected to have conserved PCLD domains and two Cys residues, belonging to four subfamilies. [Result] Among these DoPC proteins, 19 DoPCs contained four copper ligands (His, Cys, His, and Gln/Met), 32 DoPCs were predicted having N-terminal signal peptides, 22 DoPCs had putative C-terminal glycosylphosphatidylinositol-anchor signals, and 27 DoPCs had putative arabinogalactan glycomodules. The results of DoPCs expression analysis in symbiotic with *Tulasnella calospora* and asymbiotic germinated seeds from *D. officinale* by using RNA-seq data showed that 16 DoPCs were expressed in symbiotic germinated seeds, and *DoUCL2*, 4 and *DoENODL14* had the highest expression with different levels. Seven DoPC genes were obviously up-regulated and only one was down-reg-

收稿日期: 2016-11-04

基金项目: 国家高技术研究发展计划 (863 计划) 子课题 (2013AA102607-3); ‘十二五’农村领域国家科技支撑计划子课题 (2012BAD01B0702)

作者简介: 宋 争 (1982—), 男, 博士研究生. 主要研究方向: 园林植物. E-mail: songzheng566@163.com

* 通讯作者: 王 涛, 博士, 助理研究员. 主要研究方向: 园林植物遗传育种. E-mail: wangtao5757@126.com

ulated in symbiotic germinated seeds comparing with that of asymbiotic. [**Conclusion**] The study of the protein structure characteristics of DoPCs family in *Dendrobium officinale* during the whole genome, would help in researching the interaction between Orchidaceae and microorganism, furthermore would serve as data foundation, in researching the molecular mechanism of Orchid mycorrhiza growth.

Keywords: *Dendrobium officinale*; *phytoeyanin* gene family; genome-wide analysis; symbiotic germinated seeds

质体蓝素 (Phytoeyanin, PC) 是一类存在于植物中的古老超蛋白家族, 该家族均有质体蓝素样结构域 (Plastocyanin-like domain, PCLD), 具有 2 个能形成二硫键的半胱氨酸 (Cys), 其中, 部分成员含有 4 个保守的铜离子结合位点: 2 个组氨酸 (His)、1 个半胱氨酸和 1 个甲硫氨酸 (Met) 或谷氨酰胺 (Gln)^[1-2]。根据铜离子结合位点的氨基酸残基类型、光谱和氧化还原性能、蛋白结构组分等特征, PCs 家族分为 4 个亚家族: 花青苷样蛋白 (Uclacyanin-like protein, UCL)、漆树兰蛋白 (Stellacyanin-like protein, SCL)、质体蓝素样蛋白 (Plantacyanin-like protein, PLCL) 和早期结瘤素样蛋白 (Early nodulin-like protein, ENODL)^[3]。UCL 和 PLCL 有相同的铜离子结合位点 (2 个 His, 1 个 Cys 和 1 个 Met), 但 UCL 属于嵌合型糖蛋白, 其蛋白骨架中含有糖蛋白样结构域, 而 PLCL 为非糖蛋白^[3]。SCL 的铜离子结合位点中 Gln 替代了 Met, SCL 和 UCL 家族不仅包含天冬酰胺 (Asn) 残基连接的 N-糖基化位点, 还具有通过丝氨酸 (Ser) 和羟脯氨酸 (Hyp) 残基连接的 O-糖基化位点^[4]。ENODL 与其它亚家族之间有高度的相似性, 但该亚家族成员的蛋白骨架缺完整的铜离子结合位点, 铜配位体中的 4 个氨基酸残基部分或全部被其它氨基酸残基取代^[5]。研究发现, 拟南芥 (*Arabidopsis thaliana*) 和水稻 (*Oryza sativa*) 中的大多数 ENODLs 是嵌合的阿拉伯半乳糖蛋白, 其蛋白骨架包含 N-端信号肽 (Signal peptide, SP) 和阿拉伯半乳糖蛋白样域 (Arabinogalactan protein-like region, ALR)^[5-6]。

PC 家族中能结合铜离子的 3 个亚家族 (UCL、SCL 和 PLCL) 为电子传递体。研究显示, PCs 在植物发育和抗逆过程扮演着重要角色^[5, 7-8], PCs 参与了植物对非生物胁迫的应激反应, 如干旱胁迫、盐胁迫、冷胁迫和金属离子等^[5, 9-11]。

全基因组分析是鉴定基因家族成员、阐明其生物学作用的首要步骤和有效途径。利用生物信息学分析方法对植物 PC 家族基因全基因组分析的报道不多, 仅在拟南芥、水稻和大白菜 (*Brassica rapa*) 基

因组中分别发现 38、62 和 84 条 PCs 基因^[5-7]。铁皮石斛 (*Dendrobium officinale*) 为兰科石斛属植物中具有观赏价值和医药价值的一种附生兰, 其基因组测序的完成为该物种 PC 基因家族 (DoPC) 的全基因组预测和生物信息学分析提供了宝贵资源, 有助于了解该基因家族的潜在功能。本研究可为石斛属 PC 基因家族功能验证提供数据基础。

1 材料与方法

1.1 数据库和分析软件

1.1.1 数据库 铁皮石斛基因组数据库: <http://202.203.187.112/herbalplant>; 拟南芥基因组数据库: <http://arabidopsis.org/>; 国家水稻数据中心数据库: <http://www.ricedata.cn/gene/>; SMART 数据库: <http://smart.embl-heidelberg.de>; pfam 数据库: <http://pfam.xfam.org/>。

1.1.2 分析软件 氨基酸序列分析软件: BLASTP, clustalW; N-端信号肽预测: SignalP 4.1 (<http://www.cbs.dtu.dk/services/SignalP/>); 糖基磷脂酰肌醇锚定信号 (C-terminal glycosylphosphatidylinositol signal, GAS) 预测: Big-PI Plant Predictor (http://mendel.imp.ac.at/gpi/plant_server.html) 和 PSORT (<http://psort.hgc.jp/form.html>); N-糖基化位点预测: NetNGlyc 1.0 Server (<http://www.cbs.dtu.dk/services/NetNGlyc/>); 系统进化分析软件: MEGA 5.0; 图像处理软件: PhotoShop。

1.2 方法

1.2.1 铁皮石斛 DoPC 基因家族的预测 下载铁皮石斛全基因组蛋白编码序列和拟南芥、水稻 PC 基因家族成员的蛋白序列 (AtPCs 和 OsPCs) 到本地。利用 BLASTP 工具进行比对搜索 (缺省参数设置), 手动去除错配序列。

1.2.2 DoPC 家族蛋白 PCLD 结构域验证 将 BLASTP 比对获得的潜在 DoPCs 蛋白序列文件保存为 FASTA 格式, 利用 SMART 和 pfam 数据库验证 DoPCs 蛋白序列的质体蓝素样结构域 PCLD (缺省参数设置)。

1.2.3 DoPC 家族的结构分析 将存在 PCLD 结构域的 DoPCs 蛋白序列文件保存为 FASTA 格式,利用 SignalP 4.1、Big-PI Plant Predictor 和 NetNGlyc 1.0 Server 软件分别预测其 SP、GAS 位点、N-糖基化位点(缺省参数设置)。依据 Mashifuchi 等^[6]研究拟南芥 AtPCs 蛋白序列潜在 ALR 的原则:以非邻接的脯氨酸(Pro)残基基序([Ala/Ser/Thr/Gly]-Pro-X(0,10)-[Ala/Ser/Thr/Gly]-Pro)和邻接的 Pro 残基基序([Ala/Ser/Thr/Gly]-Pro₃₋₄)为阿拉伯半乳聚糖(AG)糖模块;Ser-Pro₂₋₄为假定的伸展蛋白糖模块,手动分析 DoPCs 蛋白序列中潜在的 ALRs。

1.2.4 结构域的多重序列比对分析 利用软件 Clustal W 2.1 对 DoPCs 氨基酸序列 PCLD 结构域进行多重序列比对分析(缺省参数设置),手动标记结构域特征位点。

1.2.5 系统进化树的构建 首先,利用 Clustal W 软件对所有 DoPCs 的氨基酸序列进行比对,生成软件 MEGA 可读的比对文件(缺省参数设置);然后,利用 MEGA5.0 软件采用邻接法(Neighbor-joining Method)对上述文件进行进化树分析。Bootstrap 设置为 1 000。最后生成无根树。

1.2.6 基因表达分析 按照 Zhao 等^[12]发表的方法采集正常生长的铁皮石斛成熟种子(中国林业科学研究院科研温室),分为 2 组进行体外共生和非共生萌发试验。一组种子用来构建与美孢胶膜菌(*Tulasnella calospora*, ITS 序列 GenBank 登录号:GU166418.1)体外共生萌发体系(DoTcg),选取萌发到第三阶段(原分生组织出现)的种子利用 70% 酒精表面消毒后进行台盼蓝染色^[13],在种子内检测到共生真菌菌丝,表明共生体系建立^[13];另一组种子进行体外非共生萌发到相同阶段(Do),两组试验分别设置三组生物学重复。液氮取样、提取总 RNA、构建 RNA-seq 文库进行高通量测序^[12]。将测序获得的 DoPCs 家族基因 reads 数进行标准化处理获得基因表达的 RPKM 值(Reads per kilobase per million mapped reads)值^[12],再取 \log_2 对数输入软件 Heatmap v1.0 (Heatmap Illustrator)绘制 DoPCs 基因表达热图^[14]。设置 $|\log_2(\text{Fold-change})| > 1$ (p-value < 0.01)作为基因差异表达筛选条件。

2 结果与分析

2.1 DoPC 家族的鉴定和分类

利用 BLASTP 搜索比对工具从铁皮石斛蛋白数

据库中分别找出与拟南芥 PCs (AtPCs) 和水稻 (Os-PCs) 相似度较高的序列 (E-value < 0.001), 经过 SMART 预测和 Pfam 数据库比对两种方法验证, 在拟南芥和水稻中均有 38 条 PC 样蛋白序列 (DoPCs) 含有 1 个 PCLD 结构域。已知拟南芥基因组中含有 38 条 PC 基因^[6, 15], 在数量上与铁皮石斛的 PC 基因相同; 而水稻和大白菜基因组中分别含有 62 条和 84 条 PC 基因, 远多于铁皮石斛和拟南芥基因组中 PC 基因的数量。

对 DoPCs 家族成员的 PCLD 结构域进行多重序列比对分析, 结果 (图 1) 表明: 38 条 DoPCs 蛋白序列均含有 2 个 Cys 残基, Cys 在 DoPCs 家族中高度保守; 19 条 DoPCs 蛋白序列中含有 4 个完整的铜结合配体 (His、Cys、His、Met/Gln), 其中, 11 条为 His、Cys、His、Met 类型, 8 条为 His、Cys、His、Gln 类型。19 个可结合铜的 DoPCs 蛋白归类为 3 个亚家族中: 5 个花青苷样蛋白 (DoUCLs)、8 个漆树蓝样蛋白 (DoSCLs) 和 6 个质体蓝素样蛋白 (DoPLCLs); 其余 19 个不含铜结合位点的 DoPCs 为早期结瘤素样蛋白 (DoENODLs) (表 1)。多重序列比对发现在 DoENODLs 亚家族中, 与其它亚家族的铜结合位点所对应的位置上有一些较保守的基序, 这些氨基酸残基取代了 His、Cys、His、Met/Gln, 可能与结合铜有关。用亚家族名称命名后的 DoPCs 所对应的铁皮石斛基因 ID、结构域数量及铜结合位点等配体信息见表 1。

2.2 DoPCs 家族结构分析

为进一步明确 DoPCs 家族成员序列结构特征, 本文采用生物信息学分析方法预测了该家族成员蛋白骨架的 N-端 SP、ALR 和 GAS 以及 N-糖基化位点等。预测分析结果显示: 有 31 条 DoPCs 骨架含有 N-端 SP, 负责引导该链合成的蛋白质进入分泌通路^[3]。采用 big-PI Plant Predictor 和 PSORT 在线软件预测结果显示: 22 条 DoPCs 含有 GAS 位点, 而糖基磷脂酰肌醇锚定蛋白是一种膜蛋白, 它可通过糖基磷脂酰肌醇插入内质网膜和质膜, 主要负责携带蛋白定位至质膜上, 与膜上的脂筏连接^[22-23]。此外, 有 27 条 DoPCs 含 N-糖基化位点, 可连接糖链^[6]。

已知非邻接重复出现的 Pro 可以被羟基化作为阿拉伯半乳聚糖糖基化位点, 而连续重复出现的 Pro 被羟基化为阿拉伯糖糖基化位点^[5, 16]。当 Ala、Ser、Thr、Gly 和 Pro 在核心蛋白中为主要氨基酸组分, 且 Pro 以基序为: [Ala/Ser/Thr/Gly]-Pro-X(0,10)-

[Ala/Ser/Thr/Gly]-Pro 的非邻接重复方式出现,即为假定 AG 糖模块^[17-19]。有研究发现,当 Pro 以基序为 Ser-Pro2-4 连续重复出现时,为假定的伸展蛋白糖模块;然而, Ser-Pro2-4 基序中的 Pro 不完全羟基化时,可形成非邻接型的 Pro 模块^[20]。在拟南芥的相关研究中把高度相似的非邻接型 Ala/Thr/Gly-Pro 基序和 Ser-Pro 基序均作为 AG 糖模块的代表。Mashifuchi 等将 [Ala/Ser/Thr/Gly]-Pro3-4 基序也作为假定 AG 糖模块^[6]。本文结合前人的研究成果,将 [Ala/Ser/Thr/Gly]-Pro-X (0, 10)-[Ala/Ser/

Thr/Gly]-Pro 和 [Ala/Ser/Thr/Gly]-Pro3-4 基序作为假定 AG 糖模块的预测规则。由于 N-端 SP 作为分泌信号,是阿拉伯半乳糖蛋白 (Arabinogalactan protein, AGP) 的先决条件,因此,本文只对含有 SP 的 31 条 DoPCs 是否存在 AG 糖模块进行预测。经统计,共有 26 条 DoPCs 存在 AG 糖模块,预测为嵌合的 AGPs,包括 13 条 DoENODLs、1 条 DoUCLs、6 条 DoPLCLs 和 6 条 DoSCLs,其中,10 条 DoPCs 预测包含伸展蛋白糖模块,有 8 条 DoPCs 既可以是伸展蛋白又可以是 AGPs (表 2)。

DoUCL1	-LIHID---TKESS-SFVSQALGS-FVFNVVK---	GKINVLKV-TGPAFRSCKN-TVSSPLTTGND---	KIELTSAGRKWIYA-IGDICEKG-M
DoUCL2	WKVGVN---YTEWA-EDKEFRVGRDIFVFNVVK---	GKINVLKV-TGPAFRSCKN-TVSSPLTTGND---	KIELTSAGRKWIYA-IGDICEKG-M
DoUCL3	WDFD---VNNWP-NCKTFHAGNILEFKYQK---	GNITVVKV-DANGYANCKT-DNGL-TFDSGDD---	HIKLVRCRTH-YFICG-IADICKVRNM
DoUCL4	WGFN---VKNWP-NDKTFHAGDILEFKYQK---	GNITVVKV-DANGYANCKS-MNAISTFDSGYD---	NITLARGTH-YFICG-IAGCWFQGM
DoUCL5	WDFD---VNNWP-NCKTFHAGDILEFKYQK---	GNITVVKV-DANGYANCKT-DNGL-TFDSGDD---	HIKLVRCRTH-YFICG-IADICKVRNM
DoPLCL1	WDTGVD---YSSWA-SGKTFAGGDSLTFVAYGS---	GATVNEV-SANDYQCSSTSNYL-STDSGSGT---	TI TLKTACTHYFVCG-VPGHCIGG-M
DoPLCL2	WDTGVD---YSSWA-SGKTFAGGDSLTFVAYGS---	GATVNEV-SANDYQCSSTSNYL-STDSGSGT---	TI TLKTACTHYFVCG-VPGHCIGG-M
DoPLCL3	WDLSTD---LQSWA-SSQTFAPGDTLTFVFCGS---	FISVWEV-PKADYEAETI-NPIKAYTSGSA---	VVKLSSPGKRYFVCG-TPHCOCQG-M
DoPLCL4	WTAGVD---FSSWT-NCKLFVWGDLSFNYPAP---	ATITVNEV-KADYKSCCTASNYL-STDSGSGT---	TIRLKCACTHYFVCG-TPHCOCQG-M
DoPLCL5	WTFN---PAAWT-KGKRFRAGDVLIFKYS---	SVINVAV-NAAGYNACT-PKGSKIYKSGND---	RITLVKGPV-YFICG-PPACVCSG-M
DoPLCL6	WTSQVD---YSTWA-SGKVFVWGDLSLFTYS---	CLSAEVE-SKSDYSCSTSNAL-QTHDQNT---	KISLTPGSRFYFICG-TACHCAGG-M
DoSCL1	WTIAAN---YSSWA-SGKTFRIGDTIAFIYNK---	SFINVLEV-SKADYRSCNA-AAPIATHATGND---	SISITRYGHRFFIICG-PPHCACAGG-M
DoSCL2	WTVGNL---N-YTQWA-ASKTIRVGDITLFFKYNN---	QFINVLEV-KEPDFNACKN-SSAIKEYKTKGD---	SIFPITSTGHYFICG-PPHCOCQG-M
DoSCL3	WTVEVN---YSAWA-ATKIFNLGDTIVFTYNYK---	SSIDVVEL-EEAEFSSCNV-TAPIASYSTGND---	SIVIVRYGHSFFAICG-PPHCVAGG-M
DoSCL4	WTVEVN---YSAWA-ATKIFNLGDTIVFTYNYK---	SSIDVVEL-EEAEFSSCNV-TAPIASYSTGND---	SIVIVRYGHSFFAICG-PPHCVAGG-M
DoSCL5	WTVEVN---YSAWA-ATKIFNLGDTIVFTYNYK---	SSIDVVEL-EEAEFSSCNV-TAPIASYSTGND---	SIVIVRYGHSFFAICG-PPHCVAGG-M
DoSCL6	WTIFPNS---SFYFNWA-SQNKLAVGDITLVFNFT---	GATVNEV-PKSSYDACSTKNQVATLSGCPA---	TVTINSACKHYFICG-VTYCALN-M
DoSCL7	WTLILNSP---N-YTAWA-ACKPFRBGDTVYFYKYNK---	QFINVLEV-SKADYNSCKS-SSPASYSTGND---	SITIKTACHHYFICG-PPHCOCQG-M
DoSCL8	WTLNDKP---K-YIAWA-KSKTFYVGDDELLFQYDK---	QLINVLQV-NKQAYHCNT-EAPIAVFNTGYD---	SITIRSKHELYFICG-PPHCACAGG-M
DoENODL1	---NMFLYRT---	---SAENVLEVSQAQYEDCCG-GDIIIEIYRGPSS---	TVQLSTPGPHFFYFICG-IGLICEAGG-M
DoENODL2	WAIPISSKDPDFYKQWA-AKNRFRVDDTINFK---	YKNDVSLVVTSEYDKCRS-MHPIFFSNWCKT---	EFKFERACGLFYFICG-VNHCERCE-M
DoENODL3	W-KIPSSNSDSLQWA-EANRFQVGDYLVVD---	YDAKDSVLQVRKKDYHSCNT-SSPIAVHQDQNT---	RVRLRRSGAHYFICG-VBCECKGE-M
DoENODL4	WRKFTGFESETYNQWA-GNRFRHFGDSLIFYK---	YEDKSLLVVDKGYHBCNT-MNFLSRFTDQNT---	TFRFRFCYFYFICG-QSCICRAGG-M
DoENODL5	WDTAD---LTSWP-VNKLFSVGDVLFNQVD---	KNGYDTCNS-SNALLSRSDGNT---	SISLSSSGDIYFICG-KLPCFCGG-M
DoENODL6	WTVPTDPNSMSYNQWA-EKRRFRIDDSLLFVQ---	YPPDKDSVLQDKLAYNMCST-SSYIAKFDGNT---	VFTFNMSGLFYFICG-VEKNCRINE-M
DoENODL7	WNFC---V-N-YTLVWVKNRPFHVDLIFVQSS---	CMADIVQV-DETAIDKDA-SHPISNYKGRS---	YAFELNQTTRYFICG-SFGYCYQG-M
DoENODL8	WGVFPPLPNLYASWV-LNHFFHINDSLCKSTIPLHQLLIQIYVVTILCFIADFLYPPSODSVQVTEKAFNVCIIV-TDPIKLMEGNS---	YKKGEDSVLVKKEEDYESCNT-KNPIKFFFDGKS---	IFKFTSPGTFYFICG-VADICEKR-M
DoENODL9	WYDNLNVFKVYQKWA-ACKNIFSLGDFLIFNTDK---	NISVIQTNYQTYNRCDSDDADVDITWESAGEPFTDAVDVTVFPVLLRVGSNFFS3GNYDGCRCRGG-M	
DoENODL10	W-KIPSSNSDSLQWA-EANRFQVGDYLVVD---	YDAKDSVLQVRKKDYHSCNT-SSPIVHQDQNT---	RVRLRRSGAHYFICG-VBCECKGE-M
DoENODL11	WVLKP---AEKYNSWA-ERMRFQVQDKLIFN---	YKKGEDSVLVKKEEDYESCNT-KNPIKFFFDGKS---	VFIFDKSGPFFFIICG-ADGICAGG-M
DoENODL12	---MESKMEYSMSKCSVFEMFD---	TGLYDVKLVSRREFEINCSA-ENAFRTFEIGPA---	TVSLIEBGMRYFICG-SFGNYCFGL-M
DoENODL13	WGVFPPLPNLYASWV-LNHFFHINDSLCKSTIPLHQLLIQIYVVTILCFIADFLYPPSODSVQVTEKAFNVCIIV-TDPIKLMEGNS---	YKKGEDSVLVKKEEDYESCNT-KNPIKFFFDGKS---	IFKFTSPGTFYFICG-VADICEKR-M
DoENODL14	WVVKP---HESYDLWA-VNMRFPQNDKLVFK---	YKKGEDSVLVKKEEDYESCNT-KNPIKFFFDGKS---	VFLLDRSGPFFFIICG-EPCICAGG-M
DoENODL15	WRKFTGFESETYSQWA-GNRFRHFGDSLIFYK---	YEDKSLLVVDKGYHBCNT-MNFLSRFTDQNT---	TFRFRFCYFYFICG-QSCICRAGG-M
DoENODL16	WSNG---VSYWP-NCKTFYAGDVLFTFYLYK---	NSRNVLELTDVSDYEAETA-RSYTNAVYSGDD---	YVELRSGNT-YFIDG-IPHCNSGA-M
DoENODL17	WAFD---VNNWP-TGKTFHADDILEFKYQK---	DYHNVVKV-DSNGYANCKT-DNAISTFDSGDD---	QIKLASGTH-YFICG-IADH-M
DoENODL18	WDTAD---LTSWP-VNKLFSVGDVLFNQVD---	---KNGYDTCNS-SNALLSRSDGNT---	SISLSSSGDIYFICG-KLPCFCGG-M
DoENODL19	WRVFDNNTDLYTNWA-SKLRFGQVDSIVFE---	YKNDVSLVIRVEKRGYHCNE-SGNAVFKDGR---	VFLPKFYFVYFVSS-NFPCCKGG-M

注:蓝色背景代表可能参与结合铜离子的保守氨基酸残基 (His, Cys, His, Gln/Met);黄色背景代表参与二硫键形成的半胱氨酸残基 Cys;粉色背景代表 ENODLs 亚家族中保守氨基酸残基。The conserved amino acid residues (His, Cys, His, Met/Gln) involved in copper binding are indicated with a blue background. Cys residues involved in disulfide linkages are indicated with a yellow background. Amino acid residues conserved in the ENODLs are indicated with a pink background.

图 1 DoPCs 家族 PCLD 保守结构域序列多重比对

Fig. 1 Multiple sequence alignment of the PCLDs in DoPCs

根据 DoPCs 和 AtPCs 蛋白骨架中是否含有 N-端 SP、PCLD 结构域、ALR 和 C-端 GAS 这几个部分,将其分为七类 (图 2)^[5]。类群 I 的 DoPCs 成员有 16 个,其蛋白骨架包含所有以上 4 个结构域和基序;类群 II 和 III 在结构上与类群 I 存在一定的相似,但分别缺少 GAS 和 ALR,其中,归为类群 II 的 DoPCs 和

有 10 个,归为类群 III 的 DoPC 有 0 个;类群 IV 同时缺失 ALR 和 C-端 GAS 两个结构域,有 4 个成员被归为该族群;归为类群 V 的 6 个成员和归为类群 VI 的 2 个成员都不存在 N-端 SP 和 ALR 结构,但前者具有 C-端 GAS,而后者只具有 PCLD 这一个结构域;类群 VII 同时缺失 ALR 和 C-端 GAS,但存在 2 个

PCLD, 归为该类群的 DoPc 成员有 0 个。由于类群 I 和 II 中存在 N-端 SP 以及足够参与糖基化反应的 AG 糖模块, 可认为这两类中的基因家族成员是嵌合型的 AGPs。类群 III 和 VII 不含有铁皮石斛 PC 家

族成员, 在拟南芥、水稻和杨树中也有类似的现象, 这有可能是由于基因在进化过程中失去功能, 发生消亡导致^[15]。

表 1 DoPC 家族成员及其结构信息

Table 1 The putative DoPCs named information

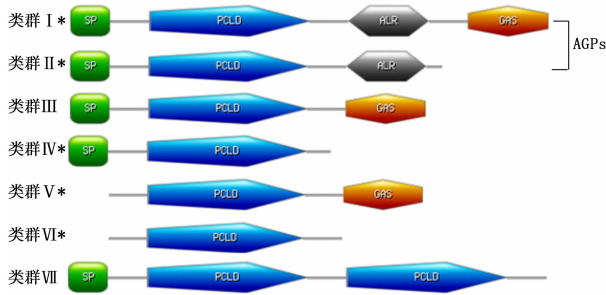
名称 Name	基因 ID Gene ID	AA	信号肽 Signal peptide	PCLD	铜结合位点 Copper-binding sites	N-糖基化位点 N-glycosylation sites	ALR	GAS
DoUCL1	Dendrobium_GLEAN_10023378	148	—	1	H - C - H - M	+	NC	+
DoUCL2	Dendrobium_GLEAN_10044449	256	+	1	H - C - H - M	+	+	—
DoUCL3	Dendrobium_GLEAN_10077806	128	+	1	H - C - H - M	+	—	—
DoUCL4	Dendrobium_GLEAN_10077807	129	+	1	H - C - H - M	+	—	—
DoUCL5	Dendrobium_GLEAN_10129350	128	+	1	H - C - H - M	+	—	—
DoPLCL1	Dendrobium_GLEAN_10014563	170	+	1	H - C - H - M	—	+	+
DoPLCL2	Dendrobium_GLEAN_10014564	170	+	1	H - C - H - M	—	+	+
DoPLCL3	Dendrobium_GLEAN_10021763	220	+	1	H - C - H - M	—	+	+
DoPLCL4	Dendrobium_GLEAN_10070311	191	+	1	H - C - H - M	—	+	+
DoPLCL5	Dendrobium_GLEAN_10077794	128	+	1	H - C - H - M	—	+	—
DoPLCL6	Dendrobium_GLEAN_10092500	191	+	1	H - C - H - M	—	+	+
DoSCL1	Dendrobium_GLEAN_10027914	207	+	1	H - C - H - Q	+	+	+
DoSCL2	Dendrobium_GLEAN_10030280	175	+	1	H - C - H - Q	+	+	+
DoSCL3	Dendrobium_GLEAN_10066345	169	+	1	H - C - H - Q	+	+	—
DoSCL4	Dendrobium_GLEAN_10066346	169	+	1	H - C - H - Q	+	+	—
DoSCL5	Dendrobium_GLEAN_10066347	170	+	1	H - C - H - Q	+	+	—
DoSCL6	Dendrobium_GLEAN_10088776	200	—	1	H - C - H - Q	+	NC	+
DoSCL7	Dendrobium_GLEAN_10114547	189	+	1	H - C - H - Q	+	+	+
DoSCL8	Dendrobium_GLEAN_10136317	199	—	1	H - C - H - Q	—	NC	—
DoENODL1	Dendrobium_GLEAN_10008487	133	—	1	—	—	NC	+
DoENODL2	Dendrobium_GLEAN_10009350	175	+	1	—	—	+	+
DoENODL3	Dendrobium_GLEAN_10011247	187	+	1	—	+	+	+
DoENODL4	Dendrobium_GLEAN_10015510	193	+	1	—	+	+	+
DoENODL5	Dendrobium_GLEAN_10017023	179	+	1	—	+	+	—
DoENODL6	Dendrobium_GLEAN_10028538	204	+	1	—	+	+	—
DoENODL7	Dendrobium_GLEAN_10028670	158	+	1	—	+	+	+
DoENODL8	Dendrobium_GLEAN_10035898	220	—	1	—	+	NC	+
DoENODL9	Dendrobium_GLEAN_10037071	224	+	1	—	+	+	+
DoENODL10	Dendrobium_GLEAN_10041716	187	+	1	—	+	+	+
DoENODL11	Dendrobium_GLEAN_10045771	358	+	1	—	+	+	+
DoENODL12	Dendrobium_GLEAN_10047036	139	—	1	—	+	NC	—
DoENODL13	Dendrobium_GLEAN_10059911	220	—	1	—	—	NC	+
DoENODL14	Dendrobium_GLEAN_10061798	174	+	1	—	+	NC	+
DoENODL15	Dendrobium_GLEAN_10064039	254	+	1	—	+	+	—
DoENODL16	Dendrobium_GLEAN_10077812	131	+	1	—	—	—	—
DoENODL17	Dendrobium_GLEAN_10077815	187	+	1	—	+	+	—
DoENODL18	Dendrobium_GLEAN_10080274	179	+	1	—	+	+	—
DoENODL19	Dendrobium_GLEAN_10129392	173	+	1	—	+	+	+

AA: 氨基酸序列的长度; Length of amino acid sequences; PCLD: (plastocyanin-like domain) PCLD 结构域数量; ALR: (Arabinogalactan protein-like region) 阿拉伯半乳聚糖样域; GAS: (C-terminal glycosylphosphatidylinositol signal) 糖基磷脂酰肌醇锚定信号; + 预测到; - 未预测到; + Predicted; - not predicted; NC: 由于蛋白序列前端未预测到信号肽而未做阿拉伯半乳聚糖样域的预测; There was no prediction of arabinogalactan protein-like region.

2.3 DoPCs 家族的系统发育分析

利用已知的 38 条 AtPCs 蛋白序列和预测得到

的 38 条 DoPCs 蛋白序列全长, 来构建系统发育树以分析其进化关系。系统发育树结果(图 3)表明: 包



注:PCs 类群结构示意图依据 MyDomains (<http://www.expasy.ch/cgi-bin/prosite/mydomains/>) 在线软件绘制。SP: N-端信号肽; PCLD: Plastocyanin 样域; ALR: 阿拉伯半乳糖样域; GAS: 糖基磷脂酰肌醇锚定信号; AGP: 阿拉伯半乳糖基模块; 星号(*)表示有 DoPC 成员。The diagram of features of DoPC domains was generated with online MyDomains software (<http://prosite.expasy.org/cgi-bin/prosite/mydomains/>). SP-signal peptide, PCLD-plastocyanin-like domain, ALR-AGP-like region, GAS-GPI-anchor signal, AGP-arabinogalactan protein. The stars indicated that there was DoPCs. The figure is not drawn to scale.

图2 DoPCs 和 AtPCs 家族7个类群结构示意图

Fig.2 Schematic representation of seven groups of DoPCs and AtPCs

含 SP、PCLD 和 GAS 的类群 III、包含 SP 和 PCLD 的类群 IV 及包含 SP 和 2 个 PCLD 的类群 VII 的成员较集中,聚集在一起,但每个亚家族的多数成员较为分散地聚集在一起。DoPLCL1/2、DoUCL1/2、DoUCL3/5、DoSCL3/4/5、DoENODL5/18、DoENODL3/10、DoENODL8/13、DoENODL4/15 等 8 对旁系同源蛋白序列相似性较高,并且处于较近的分支上,表明它们之间很可能有功能冗余的现象。

2.4 DoPCs 家族基因在与共生萌发种子中的表达分析

为确定 *DoPC* 基因在铁皮石斛种子与菌根菌共生萌发过程中的表达水平,本文利用铁皮石斛体外非共生萌发种子(Do)和共生萌发种子(DoTcg)转录组数据进行 *DoPC* 基因的表达分析。结果(图4)表明:16 个 *DoPCs* 基因在与菌根真菌共生萌发的铁皮石斛种子中检测到表达,其中,DoUCL2、4、DoSCL6、7 和 DoENOD 3、7、14、17 基因表达量较高,而 DoUCL5 和 DoENODL1、17 基因在非共生萌发的种子中未检测到表达。对比 *DoPCs* 基因在共生萌发和非共生萌发种子中的表达水平,除去 3 个在非共生萌发种子中未表达的基因,差异表达基因中大部分基因在共生萌发的种子中明显上调表达,只有 DoENODL6 基因下调表达。这些差异表达基因可能参与铁皮石斛种子与美孢胶膜菌共生萌发生物学调控。

表2 DoPCs 家族成员阿拉伯半乳糖基模块预测

Table 2 Putative glycomodules in the putative DoPCs

Name	糖基模块 Number of putative glycomodules			
	[A/S/T]-P	([A/S/T]-P-X (0,10)-[A/S/ T]-P)X	[A/S/ T]-P-P	[A/S/T]- P-P-P
DoUCL1	-	-	-	-
DoUCL2	1	1	0	0
DoUCL3	0	0	0	0
DoUCL4	0	0	0	0
DoUCL5	0	0	0	0
DoPLCL1	0	2	0	0
DoPLCL2	0	2	0	0
DoPLCL3	2	8	0	1
DoPLCL4	1	3	1	0
DoPLCL5	2	0	0	0
DoPLCL6	2	2	2	0
DoSCL1	3	2	0	0
DoSCL2	0	2	0	1
DoSCL3	1	0	0	0
DoSCL4	1	0	0	0
DoSCL5	1	0	0	0
DoSCL6	-	-	-	-
DoSCL7	2	5	0	0
DoSCL8	-	-	-	-
DoENODL1	-	-	-	-
DoENODL2	0	0	1	0
DoENODL3	1	4	0	0
DoENODL4	0	3	0	0
DoENODL5	1	0	0	0
DoENODL6	0	5	0	1
DoENODL7	0	0	0	1
DoENODL8	-	-	-	-
DoENODL9	0	4	0	0
DoENODL10	1	4	0	0
DoENODL11	0	22	1	2
DoENODL12	-	-	-	-
DoENODL13	-	-	-	-
DoENODL14	-	-	-	-
DoENODL15	1	4	0	0
DoENODL16	0	0	0	0
DoENODL17	1	0	0	0
DoENODL18	1	0	0	0
DoENODL19	0	2	0	0

3 讨论

通过生物信息学分析,本文对铁皮石斛 *PC* 基因家族进行全基因组范围的预测分析,获得 38 个 *DoPCs* 基因。研究报道的拟南芥 *AtPC*、水稻 *OsPC*、大白菜 *BrPC* 分别有 38、62 和 84 个 *PC* 基因^[5-7],铁皮石斛基因组 *DoPC* 基因数量远少于水稻和大白菜的 *PC* 基因,与拟南芥的 *PC* 数量相同。对比分析

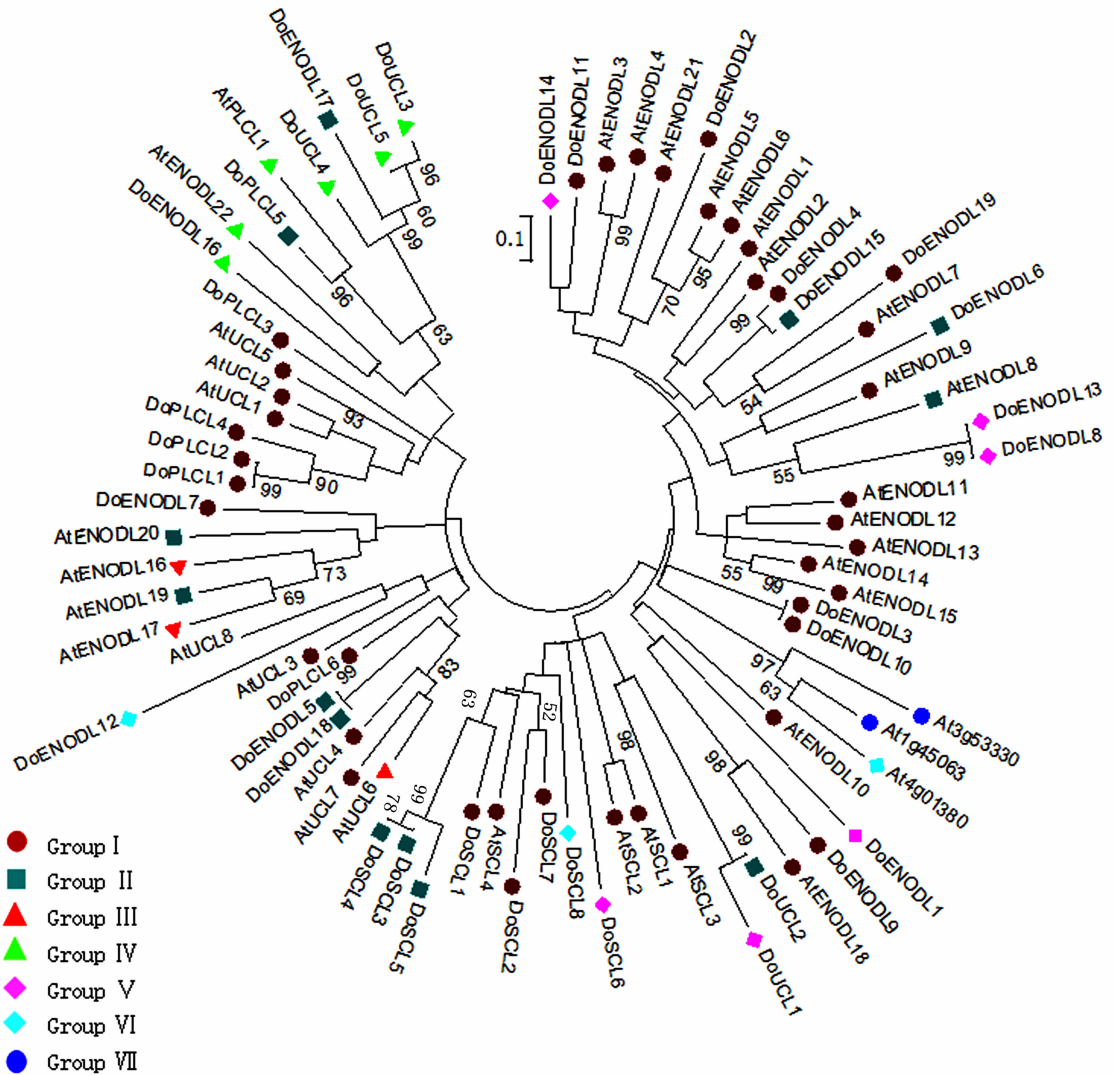


图3 DoPCs 和 AtPCs 家族系统发育分析

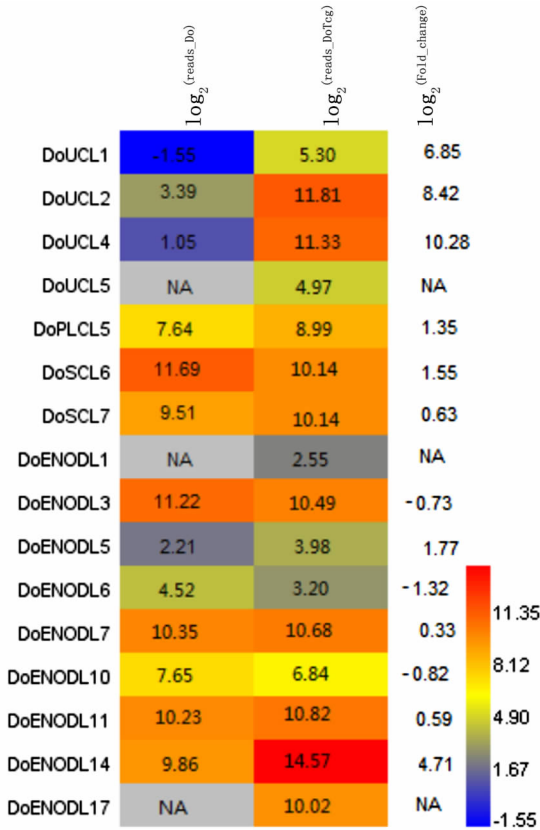
Fig. 3 Phylogenetic analysis of DoPCs and AtPCs

NCBI (<http://www.ncbi.nlm.nih.gov>) 上的基因组数据, 拟南芥、水稻和大白菜分别有 33 583、30 534 和 41 147 个蛋白编码基因, 铁皮石斛基因组 35 567 个蛋白编码基因, 推测该基因家族的数量可能与基因组蛋白编码基因的数量成正比。研究发现, 在拟南芥、水稻和大白菜中, 片段复制和串联复制可能是导致 PC 基因家族扩张的主要原因^[5, 7, 15]。

分析该家族的蛋白骨架结构特点发现, 2 个 Cys 在 DoPCs 家族中高度保守, 在 AtPCs、OsPCs 和 BrPCs 中也寻找到类似的现象, 研究发现, 2 个 Cys 形成的二硫键有助于维持 PCLD 结构的稳定, 且 Cys 的存在对于植物 PCs 家族基因调控功能的发挥有重要意义^[5-8]。在铁皮石斛、拟南芥、大白菜等植物中均发现, ENODLs 亚家族在对应的铜结合位点上 His、Cys、His、Met/Gln 已知配体的保守氨基酸残基

序列被其它氨基酸残基取代的现象, 这些保守氨基酸残基配体很可能与结合铜的功能相关^[6-7]。38 个 DoPCs 分为 4 个亚家族, 又根据 PC 蛋白骨架结构组成特征划分的 7 个类群, 在铁皮石斛 PC 蛋白序列中未预测到具有类群 III 和 VII 特征成员, 在拟南芥、水稻和杨树中也有类似的现象, 这有可能是由于基因在进化过程中失去功能, 发生消亡导致^[15]。系统发育分析显示, 与拟南芥、水稻、大白菜 PCs 的相关研究结果一样, 各亚家族成员分散聚集在不同分支中, 说明所有亚家族成员在单子叶植物和双子叶植物分化发生之前有着共同的祖先, 即使在铁皮石斛和拟南芥逐渐分化之后, 各亚家族之间也仍在发展进化^[5-7]。

阿拉伯半乳聚糖蛋白、伸展蛋白和富脯氨酸蛋白被认为是一大类富羟脯氨酸糖蛋白 (Hydroxypro-



注:右侧比色刻度尺中蓝色表示最低表达量,红色表示最高表达量,灰色表示表达量为“0”(NA)。图中数字分别为各基因在DoTeg和Do样品中检测到的reads数进行标准化处理、取 \log_2 对数后的表达水平(\log_2 (reads_DoTeg)和 \log_2 (reads_Do)),及在DoTeg/Do样品中的表达差异情况(\log_2 (Fold-change))。The color scale represents RPKM-normalized \log_2 transformed counts (\log_2 (reads_DoTeg) and \log_2 (reads_Do)). Blue indicates low expression, and red indicates high expression. Gray indicates no expression in the tissue.

图4 部分DoPCs家族成员在共生和非共生萌发的铁皮石斛种子中的表达情况

Fig. 4 Heatmap showing the expression levels of DoPC genes in symbiotic and asymbiotic germinated seeds of *D. officinale*

line-rich glycoproteins, HRGPs)。研究表明,HRGPs参与了细胞生长和发育的多个方面,从细胞壁结构的形成到细胞增殖、细胞间识别等^[21-22]。铁皮石斛中PC样阿拉伯半乳糖蛋白(PLAs)的研究发现,近一半的DoPCs也是嵌合的AGPs,与拟南芥、水稻相比,铁皮石斛中的PC-AGPs数量相对较少^[5,7]。在经典AGPs的核心蛋白骨架中,氨基酸残基Ala、Ser、Thr、Gly和Hyp是主要构成成分,该区域与C-端的疏水性GAS区域一起,可以使AGP定位到质膜上的脂筏结构域^[19,23]。根据是否包含SP、ALR、GAS和PCLD结构域,共有16个DoPC-PLAs(即类

群I成员)被预测为经典的AGPs,而另外10条DoPC-AGPs预测为非典型蛋白。同时,一些DoPC-AGPs预测到含有伸展蛋白糖模块,作为阿拉伯低聚糖的连接位点。AGPs可能参与细胞分裂和模式形成、幼苗生长、花粉管伸长、植物与微生物互作、植物有性生殖过程等生命活动的调控,其功能的发挥也取决于这些糖蛋白的糖基化情况^[24-26]。

PC家族ENODL亚家族基因首先是在豆科植物的根瘤中发现^[27],后来发现一些ENODLs基因在植物中受到丛枝菌根真菌侵染的激发而活跃表达^[26]。本研究发现,一些DoUCL、DoSCL和DoENODL基因在与菌根菌共生萌发的铁皮石斛种子中高表达,且DoUCL5和DoENODL1、17基因在非共生萌发的铁皮石斛种子中未检测到表达。与之前研究报道的在兰科植物与菌根真菌的共生体中也发现PC家族基因差异化高表达的现象相类似^[28],说明不仅ENODL基因在植物与微生物的共生关系可能发挥重要作用,UCL和SCL两个亚家族基因在兰科植物与菌根菌共生关系建立过程中也发挥重要作用。对比DoPCs基因在共生萌发和非共生萌发种子中的表达水平,发现差异表达的PC家族基因中大部分基因在共生萌发的种子中明显上调表达,只有DoENODL6基因明显下调表达,这些在共生萌发种子中差异表达明显的DoPCs可能以不同方式参与铁皮石斛种子与菌根菌共生萌发的生物学调控。

4 结论

本研究在铁皮石斛全基因组中共预测到38个DoPCs,可分属4个亚家族,2个Cys残基在DoPCs家族中高度保守;19条DoPCs蛋白序列含有4个完整的保守铜结合配体(His、Cys、His、Gln/Met),有32条DoPCs骨架中含有N-端信号肽,22条DoPCs含有糖基磷脂酰肌醇锚定位点,有27条DoPCs含N-糖基化位点;16个DoPCs基因在与美孢胶膜菌共生萌发的种子中检测到表达,其中,DoUCL2、4和DoENODL14基因表达量最高,7个基因在共生萌发的种子中明显上调表达,只有DoENODL6基因明显下调表达。本研究在全基因组范围内分析铁皮石斛DoPC家族蛋白结构特点,可以为探究兰科植物与微生物共生的相互作用及进一步解析兰科菌根形成的分子机制提供数据基础。

参考文献:

- [1] Ruan X, Luo F, Li D, *et al.* Cotton BCP genes encoding putative blue copper-binding proteins are functionally expressed in fiber development and involved in response to high-salinity and heavy metal stresses [J]. *Physiologia Plantarum*, 2011, 141(1): 71–83.
- [2] Hart P J, Eisenberg D, Nersissian A M, *et al.* A missing link in copper-proteins-crystal structure of cucumber stellacyanin at 1.6 resolution [J]. *Protein Science*, 1996, 5(11): 2175–2183.
- [3] Nersissian A M, Immoos C, Hill M G, *et al.* Uclacyanins, stellacyanins, and plantacyanins are distinct subfamilies of phytoacyanins Plant-specific mononuclear blue copper proteins [J]. *Protein Science*, 1998, 7(9): 1915–1929.
- [4] Van Driessche G, Van Beeumen J, Dennison C, *et al.* Heterogeneity of the covalent structure of the blue copper protein umecyanin from horse-radish roots [J]. *Protein Science*, 1995, 4(2): 209–227.
- [5] Ma H, Zhao H, Liu Z, *et al.* The phytoacyanin gene family in rice (*Oryza sativa* L.): genome-wide identification, classification and transcriptional analysis [J]. *PLoS ONE*, 2011, 10: e25184.
- [6] 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* [J]. *Bioscience Biotechnology and Biochemistry*, 2009, 73(11): 2452–2459.
- [7] Li J, Gao G, Zhang T, *et al.* The putative phytoacyanin genes in Chinese cabbage (*Brassica rapa* L.): genome-wide identification, classification and expression analysis [J]. *Molecular Genetics and Genomics*. 2013, 288(1–2): 1–20.
- [8] Dong J, Kim S T, Lord E M. Plantacyanin plays a role in reproduction in *Arabidopsis* [J]. *Plant Physiology*, 2005, 138(2): 778–789.
- [9] Diab A A, Teulat-Merah B A, This D, *et al.* Identification of drought-inducible genes and differentially expressed sequence tags in barley [J]. *Theoretical and Applied Genetics*, 2004, 109(7): 1417–1425.
- [10] Ezaki B, Sasaki K, Matsumoto H, *et al.* Functions of two genes in aluminium (Al) stress resistance: repression of oxidative damage by the *AtBCB* gene and promotion of efflux of Al ions by the *NtGD1* gene [J]. *Journal of Experimental Botany*, 2005, 56: 2661–2671.
- [11] Wu H, Shen Y, Hu Y, *et al.* A phytoacyanin-related early nodulin-like gene, *BcBCP1*, cloned from *Boea crassifolia* enhances osmotic tolerance in transgenic tobacco [J]. *Journal of Plant Physiology*, 2011, 168(9): 935–943.
- [12] Zhao M, Zhang G, Zhang D, Hsiao Y, Guo S. ESTs Analysis Reveals Putative Genes Involved in Symbiotic Seed Germination in *Dendrobium officinale* [J]. *PLoS ONE*, 2013, 8: e72705.
- [13] Zhao X, Yang J, Liu S, *et al.* The colonization patterns of different fungi on roots of *Cymbidium hybridum* plantlets and their respective inoculation effects on growth and nutrient uptake of orchid plantlet [J]. *World Journal of Microbiology and Biotechnology*, 2014, 30(7): 1993–2003.
- [14] Deng W, Wang Y, Liu Z, *et al.* HemI: A Toolkit for Illustrating Heatmaps [J]. *PLoS One*, 2014, 9: e119888.
- [15] Cao J, Li X, Lv Y, *et al.* Comparative analysis of the phytoacyanin gene family in 10 plant species; a focus on *Zea mays* [J]. *Frontiers in Plant Science*, 2015, 6: 515.
- [16] Tan L, Leykam J F, Kieliszewski M J. Glycosylation motifs that direct arabinogalactan addition to arabinogalactan-proteins [J]. *Plant Physiology*, 2003, 132(3): 1362–1369.
- [17] 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 [J]. *Plant Physiology*, 2002, 129(4): 1448–1463.
- [18] Schultz C J, Ferguson K L, Lahnstein J, *et al.* Post-translational modifications of arabinogalactan-peptides of *Arabidopsis thaliana* [J]. *The Journal of Biological Chemistry*, 2004, 279(44): 45503–45511.
- [19] Estevez J M, Kieliszewski M J, Khitrov N, *et al.* Characterization of synthetic hydroxyproline-rich proteoglycans with arabinogalactan protein and extensin motifs in *Arabidopsis* [J]. *Plant Physiology*, 2006, 142(2): 458–470.
- [20] Shpak E, Barbar E, Leykam J F, *et al.* Contiguous hydroxyproline residues direct hydroxyproline arabinosylation in *Nicotiana tabacum* [J]. *The Journal of Biological Chemistry*, 2001, 276(14): 11272–11278.
- [21] Majewska-Sawka A, Nothnagel E. The multiple roles of arabinogalactan proteins in plant development [J]. *Plant Physiology*, 2000, 122(1): 3–9.
- [22] Showalter A M. Arabinogalactan-proteins: structure, expression and function [J]. *Cellular and Molecular Life Sciences*, 2001, 58(10): 1399–1417.
- [23] Borner G H H, Lilley K S, Stevens T J, *et al.* Identification of glycosylphosphatidylinositol-anchored proteins in *Arabidopsis*. A proteomic and genomic analysis [J]. *Plant Physiology*, 2003, 132(2): 568–577.
- [24] De Vries S C. A relationship between seed development, arabinogalactan-proteins (AGPs) and the AGP mediated promotion of somatic embryogenesis [J]. *Physiologia Plantarum*, 2002, 114(4): 637–644.
- [25] Tan L, Showalter A M, Egelund J, *et al.* Arabinogalactan-proteins and the research challenges for these enigmatic plant cell surface proteoglycans [J]. *Frontiers In Plant Science*, 2012, 3(140): 1–10.
- [26] Harrison M J. Development of the arbuscular mycorrhizal symbiosis [J]. *Current Opinion in Plant Biology*, 1998, 1(4): 360–365.
- [27] Nap J P, Bisseling T. Nodulating function and nodulin gene regulation in root nodule development [M] // Gresshoff P M. *The Molecular Biology of Symbiotic Nitrogen Fixation*. Boca Raton: CRC Press Inc, 1980: 181–229.
- [28] Perotto S, Rodda M, Benetti A, *et al.* Gene expression in mycorrhizal orchid protocorms suggests a friendly plant-fungus relationship [J]. *Planta*, 2014, 239(6): 1337–1349.

(责任编辑:张 研)