

# 促生细菌挥发性物质通过 lncRNA 调控拟南芥的关键功能基因表达

李菲<sup>1,2</sup>, 石天龙<sup>1,2</sup>, 汤晓辛<sup>1</sup>, 龚记熠<sup>2</sup>, 徐小蓉<sup>2</sup>, 唐明<sup>1,2\*</sup>

(1. 西南喀斯特山地生物多样性保护国家林业局重点实验室, 贵州师范大学生命科学学院, 贵阳 550001;

2. 贵州省植物生理与发育调控重点实验室, 贵州师范大学生命科学学院, 贵阳 550001)

**摘要:** 促生细菌的挥发性有机物可以促进拟南芥生长, 提高拟南芥对多种生物和非生物胁迫的抗性。近年来的研究表明, 植物 lncRNA 参与生长、发育和胁迫反应等生理过程。通过去 rRNA 全转录组测序, 筛选出受到根际促生细菌 ACCC 01380 挥发性有机物处理影响的拟南芥 lncRNA, 并通过顺式分析获得可能受到这些 lncRNA 调控的蛋白编码基因; 通过 ceRNA 分析获得 lncRNA 影响蛋白编码基因表达的调控网络。结果发现, 拟南芥重要的生长发育调控和胁迫响应基因, 如细胞分裂素降解基因 AtCKX7、SOS2-like protein kinase PKS5, 都受到促生细菌的挥发性有机物通过 lncRNA 的调控。研究结果初步证实了根际促生细菌挥发性有机物在促生过程中 lncRNA 的作用, 并为进一步研究 lncRNA 功能以及根际促生细菌挥发性有机物的作用机理奠定了基础。

**关键词:** 枯草芽孢杆菌; 拟南芥; 挥发性有机物; lncRNA; ceRNA 网络

中图分类号: Q945.78

文献标识码: A

文章编号: 1672-352X (2020)03-0427-08

## The regulation of growth promoting bacterial volatile substances on key functional genes in *Arabidopsis thaliana* through lncRNA network

LI Fei<sup>1,2</sup>, SHI Tianlong<sup>1,2</sup>, TANG Xiaoxin<sup>1</sup>, GONG Jiyi<sup>2</sup>, XU Xiaorong<sup>2</sup>, TANG Ming<sup>1,2</sup>

(1. Key Laboratory of Biodiversity Conservation of Southwest China Karst Mountain Area, School of Life Sciences, Guizhou Normal University, Guiyang 550001; 2. Key Laboratory of Plant Physiology and Developmental Regulation, School of Life Sciences, Guizhou Normal University, Guiyang 550001)

**Abstract:** Volatile organic compounds produced by plant growth promoting bacteria can promote the growth of *Arabidopsis thaliana* and improve its resistance to a series of biotic and abiotic stresses. Recent studies have shown that plant lncRNAs are involved in physiological processes such as growth, development and stress response. In this study, *Arabidopsis* lncRNAs affected by rhizosphere-promoting bacteria ACCC 01380 volatile organic compounds were screened using full transcriptome sequencing, and protein-coding genes that may be regulated by these lncRNAs were studied by ceRNA analysis. As results, we found that key growth regulation and stress response genes in *Arabidopsis*, such as cytokinin degradation gene AtCKX7, SOS2-like protein kinase PKS5, were regulated by the lncRNAs responsive to volatile organic compounds of ACCC 01380. This research will significantly improve our understanding of the role of lncRNAs in the physiological effects of rhizosphere-promoting bacterial volatile organic compounds and lay a foundation for the further study of lncRNAs function and the mechanism of bacterial volatile organic compounds.

**Key words:** *Bacillus subtilis*; *Arabidopsis thaliana*; volatile organic compounds; lncRNA; ceRNA network

植物会把相当比例的光合作用产物分泌到根际土壤中, 供给其根际生存着的多种有益微生物。部

分植物根际细菌可产生多种挥发性化合物, 对宿主植物有广泛的生理影响, 包括促进植物生长, 提高

收稿日期: 2019-11-02

基金项目: 国家自然科学基金委与贵州师喀斯特研究中心 (U1812401) 资助。

作者简介: 李菲, 副教授。E-mail: lifei00987@outlook.com

\* 通信作者: 唐明, 副教授。E-mail: gzk1ppdr@gznu.edu.cn

植物对多种逆境的抗性, 诱导植物产生系统抗性等等。部分植物促生细菌已经作为生物菌剂的重要成分在农业上广泛应用。

对于挥发性有机物的生理功能, 研究比较透彻的是解淀粉芽孢杆菌 GB03, 一种可商业购买的根际促生细菌。GB03 的挥发性有机物可以促进拟南芥生长, 提高拟南芥对多种生物和非生物胁迫的抗性<sup>[1-2]</sup>。基因芯片分析表明, GB03 的挥发性物质影响与细胞壁修饰、初级和次级代谢、胁迫响应、激素调节等相关过程的约 600 个拟南芥基因的表达<sup>[3]</sup>。某些功能基因对于 GB03 挥发性有机物促进拟南芥对铁的吸收、提高拟南芥光合作用活性及对渗透胁迫和盐胁迫的抗性具有关键作用<sup>[1-6]</sup>。

长链非编码 RNA (Long noncoding RNA, lncRNA) 是一类长约 200 nt 至 100 kb 的非编码 RNA<sup>[7]</sup>。研究表明, lncRNAs 可以通过顺式或反式作用调控基因的表达<sup>[8]</sup>。基因上游的 lncRNA 可能与该基因的启动子或其他顺式作用元件相互作用, 在转录或转录后水平调节该基因的表达<sup>[9-10]</sup>。例如, lncRNA 可能与转录因子结合位点重叠; 因此, lncRNA 的转录可以阻止转录因子与该位点的结合<sup>[8-10]</sup>。此外, 转录因子结合位点附近的 lncRNA 可以增强转录因子与位点的结合<sup>[10]</sup>。lncRNA 可分为正义 (sense)、天然反义 (natural antisense)、双向 (bidirectional)、内含子 (intronic) 和基因间 (intergenic) lncRNA<sup>[8, 11]</sup>。不同类型的 lncRNA 具有不同的调节作用。

植物 lncRNA 参与生长、发育和胁迫反应等生理过程。例如, 与野生型植物相比, 拟南芥中 lncRNA npc48 的过表达增加了莲座丛直径和叶片锯齿, 延迟了开花时间<sup>[12]</sup>。总共有 664 个 lncRNA 被鉴定为响应玉米叶片的干旱胁迫<sup>[13]</sup>。其中, 567 个 lncRNAs 上调, 而 97 个 lncRNAs 下调。通过深度测序已经在不同植物中报道了更多的 lncRNA, 包括拟南芥<sup>[14-16]</sup>, 水稻 *Oryza*<sup>[17-18]</sup>, 玉米 *Zea mays* L.<sup>[19]</sup>, 苜蓿 *Medicago sativa* L.<sup>[20]</sup>, 毛果杨<sup>[21]</sup>和小麦<sup>[22]</sup>。

研究表明, lncRNA 可以作为内源性靶标模拟物 (endogenous target mimics, eTMs), 并且涉及与竞争性内源性 RNA (ceRNA) 类似的几种生理活动。Deng 等<sup>[23]</sup>研究了 NaCl 胁迫下陆地棉的转录组, 并确定 lnc\_973 和 lnc\_253 可作为 ceRNA 调节 ghr-miR399 和 ghr-156e 对 NaCl 胁迫的响应。Xu 等<sup>[24]</sup>从磷缺乏的水稻中鉴定了 lncRNAs, 并在缺磷条件下构建了水稻根的 ceRNA 调控网络。总的来说, lncRNA 可能通过 lncRNA-miRNA-mRNA 调节网络参与各种生物过程。

研究已经证明 lncRNA 参与植物生长、发育调控和对逆境的响应等多个生理过程, 但根际促生细菌是否影响宿主植物 lncRNA 的表达以及受到影响的 lncRNA 的可能功能和作用机理尚鲜见报道。因此, 本研究通过去 rRNA 全转录组测序, 筛选出受到根际促生细菌 ACCC 01380 处理影响的拟南芥 lncRNA, 并通过顺式分析获得可能受到这些 lncRNA 调控的蛋白编码基因, 通过 ceRNA 分析获得 lncRNA 的调控网络。本文是首个根际促生细菌调控植物 lncRNA 表达的报道, 将有利于进一步揭示根际促生细菌调控植物生理过程的分子机理。

## 1 材料与方法

### 1.1 材料

枯草芽孢杆菌 (*Bacillus subtilis*) ACCC 01380, 由中国农业菌种保藏中心提供。挑取保存在 -80℃ 冰箱中的菌种, 置于 28℃ 恒温培养箱中 LB 固体培养 24 h; 挑取单菌落接种于装有 3 mL LB 液体培养基的离心管中, 30℃ 恒温摇床培养 12 h, 所得菌液 (以下简称菌株或菌液) 放于 4℃ 冰箱备用。

取大约 50 粒哥伦比亚 (Col) 生态型拟南芥种子放入灭过菌的 1.5 mL 的离心管中, 加入 800 μL 无菌水, 再加入 150 μL 的 84 消毒液, 放在 37℃ 摇床中摇 15 min; 然后将离心管中的液体吸干净, 往离心管中加入 1 mL 的灭菌水洗涤种子, 共洗涤 3 次。将消毒杀菌后的拟南芥种子均匀地播种在 0.6% MS 培养基上, 放置在 4℃ 冰箱内春化 5 d。随后放置在 23℃ 的温室中, 光照条件下生长 16 h, 黑暗条件下生长 8 h, 生长 6 d, 等待长出拟南芥幼苗。

本实验采用的是两侧分隔平皿。往分隔平板两侧各加入 5 mL 1% MS 培养基; 拟南芥苗种在培养皿的右侧 MS 培养基中。往左侧培养基中加入 12 μL 菌液, 作为菌处理组, 另一半平皿的左侧培养基中加入 12 μL 无菌水, 作为对照组。随后将平皿放置在 23℃ 的温室中, 光照条件下生长 16 h, 黑暗条件下生长 8 h, 生长 2 d (48 h)。采集两种处理下拟南芥的根尖, 进行后续全转录组测序和分析。

### 1.2 方法

**1.2.1 文库构建与全转录组测序分析** 采用上海延慕实业有限公司的总 RNA 提取试剂盒提取 2 种处理下的拟南芥根部样品总 RNA, 每种处理取 3 个样, 共 6 个拟南芥根部样品。使用 Agilent 2200 检测仪对提取的 12 个样品的总 RNA 进行质量检测。

取总 RNA 500 ng, 加入去核糖体探针, PCR 反应 25 min 左右, 然后依次加入 RNA 消化试剂 37℃

PCR 反应 30 min, DNA 消化试剂 37°C PCR 反应 30 min, 最后体系为 50  $\mu\text{L}$ , 用 Agencourt RNA Clean XP Beads 磁珠纯化, 取 6  $\mu\text{L}$  在 94°C 的 PCR 仪进行 7~8 min 打断。取上一步溶液加入一链反转试剂, 合成一链 cDNA, 再加入二链反转试剂, 合成二链 cDNA, 最后用 DNA Clean Beads 磁珠进行纯化。取上一步溶液加入末端修复/dA 尾(脱氧腺嘌呤)添加试剂, 进行 PCR 反应 60 min, 再加入连接酶及 adaptor, 20°C PCR 反应 15 min。用上一步 100  $\mu\text{L}$  连接产物进行 2 次片段分选实验, 用 AMPure XP Beads 磁珠纯化连接产物。取上一步纯化产物加入扩增试剂, 共 50  $\mu\text{L}$  进行 PCR 扩增反应, 15 个循环。最后用 50  $\mu\text{L}$  Agencourt AMPure XP Beads 磁珠进行纯化, 用 23  $\mu\text{L}$  0.1 $\times$  TE 混匀后, 室温静置 2 min, 磁力架上 5 min, 小心吸取 20  $\mu\text{L}$  上清至一新离心管中。用 D1000ScreenTape, 放到 Agilent 2200 检测仪检测。文库检测结果为 300 bp $\leq$ 库主峰长 $\leq$ 500 bp, 浓度 $\geq$ 2 ng $\cdot\mu\text{L}^{-1}$ , 峰型单一, 没有杂峰, 则该文库符合后续上机要求。样品库检合格后, 采用 Illumina HiSeq 平台进行 rRNA-free 双末端上机测序。

**1.2.2 全转录组测序数据分析** 进行测序数据质量控制, 排除序列中的低质量序列, 未测出序列、过短序列、含有双端接头的序列, 以保证测序数据的质量与准确性, 从而获得 Clean Data (过滤后数据)。利用 Hisat2 软件与参考基因组进行序列比对, 得到 Mapped Data 后进行文库质量评估。将比对到基因组的测序数据采用 Cufflinks 软件对于每一个基因的表达情况进行确认、统计及分析。将表达量统计的结果进行差异筛选, 使用 DEGseq2 进行差异分析, 计算实验组 VS 对照组的 lncRNA 的差异情况, 获得差异倍数、显著性 ( $P$ -value) 以及误判率 ( $FDR$ -value) 筛选显著性差异基因。根据 lncRNA 的转录水平, 使用 STEM 软件对差异表达的 lncRNA 进行趋势分析 (STC, Series test of cluster), 发现基因表达的趋势特征。将相同变化特征的基因聚集在一种变化趋势中, 从而找到实验变化过程中最具有代表性的基因群, 揭示生物样本在变化过程中所特有的规律。

**1.2.3 枯草芽孢杆菌调控 lncRNA 的功能预测** 枯草芽孢杆菌调控 lncRNA 的功能预测主要通过基因组位置分析、GO 富集分析等手段进行, 获得 lncRNA 的相关功能信息。基因组位置分析主要是通过筛选与目标 lncRNA 相邻的上下游蛋白编码基因, 并获得其功能信息; 模块 GO (Gene Ontology) 富集分析主要通过 R 软件的 Blast2GO 和 topGOv.2.12.0 包对每个模块进行 GO 富集分析。

**1.2.4 CeRNA 网络分析** 使用 miRanda 和 TargetScan 鉴定含有 microRNA 响应元件 (MRE) 的 mRNA。使用 psRNATarget 和 psRobot 预测 lncRNA 是否可能是 miRNA 靶标。基于对照和枯草芽孢杆菌 ACCC 01380 挥发性有机物处理拟南芥根部样品的 mRNA 和 lncRNA 的差异表达情况, 预测 miRNA、lncRNA 和 mRNA 的相互作用。使用 Cytoscape 3.4.0 构建可视化 ceRNA 网络。

## 2 结果与分析

### 2.1 解淀粉芽孢杆菌挥发性有机物促进拟南芥生长

枯草芽孢杆菌的挥发性有机物表现出非常稳定促进拟南芥生长的能力。把拟南芥放在中间有分隔的平皿右侧, 左侧滴加大肠杆菌 DH5 $\alpha$  或枯草芽孢杆菌菌液。这样, 对拟南芥生长的影响仅来自于菌株产生的挥发性物质。与大肠杆菌 DH5 $\alpha$  处理相比, 枯草芽孢杆菌的挥发性有机物处理下的拟南芥植株, 植株更加茁壮 (图 1A), 主根长度明显增加, 苗干重也有显著的增长 (图 1B 和 1C)。

### 2.2 枯草芽孢杆菌挥发性有机物处理下拟南芥根部全转录组测序

为了全面探索表达受到枯草芽孢杆菌挥发性有机物影响的拟南芥 lncRNA, 提取受到 ACCC 01380 挥发性有机物处理 48 h 拟南芥根的 RNA 用于全转录组去 rRNA 测序 (对照和 ACCC 01380 挥发性有机物处理分别进行 3 个生物学重复)。通过测序样品获得总共 511 504 276 个 raw reads, 质量控制后, 获得 474 107 712 clean reads。与拟南芥 TAIR10 参考基因组进行比对, 结果表明全转录组测序 reads 质量较高。6 个测序样本的全部 clean reads, 平均 98.57% 的 reads 可以 map 到基因组的 unique 位置。

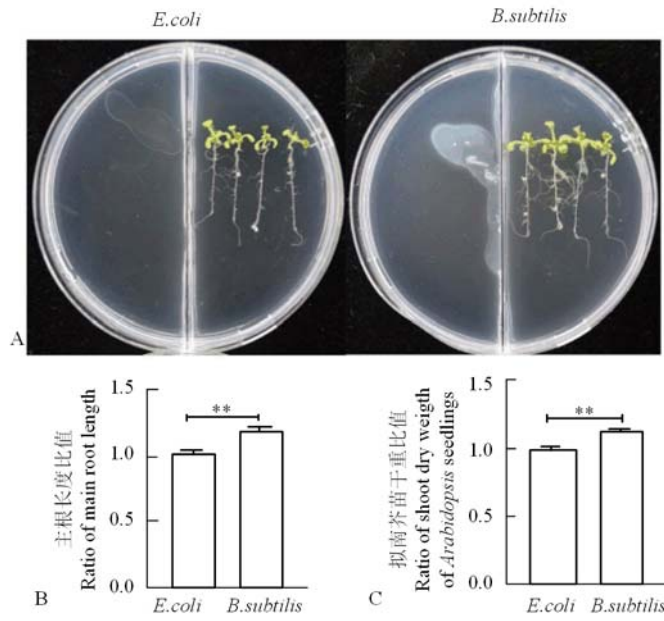
使用 StringTie 进行转录本重构和组装。从 6 个拟南芥根部全转录组测序文库中总共获得 4 083 个 lncRNA 用于后续分析。

### 2.3 枯草芽孢杆菌挥发性有机物处理引起拟南芥根部 lncRNA 的差异表达

分析了用大肠杆菌 DH5 $\alpha$  和枯草芽孢杆菌 ACCC 01380 VOC 处理的拟南芥幼苗根的转录组数据, 以鉴定差异表达的蛋白质编码基因和 lncRNA。ACCC 01380 VOCs 处理极大地改变了拟南芥蛋白编码基因和 lncRNA 的表达。共 962 个蛋白编码基因在对照和 ACCC 01380 VOC 处理之间显示出差异表达, 其中 779 个基因下调, 而 183 个基因上调 (图 2A)。在 4 083 个 TAIR10 lncRNA 中, 有 233 个在 DH5 $\alpha$  和 ACCC 01380 VOC 处理之间表现出显著

差异 ( $\text{Log}_2FC > 0.585$  或  $< -0.585$ ,  $FDR < 0.05$  是显着的差异基因) (图 2B)。其中, ACCC 01380 VOCs

上调了 88 个 lncRNA, 下调了 145 个 lncRNA。



(A) : 大肠杆菌 DH5 $\alpha$  和枯草芽孢杆菌挥发性有机物对拟南芥的生长不同影响的代表性图片; (B) 两种处理下拟南芥根部主根长度比值; (C) 两种处理下拟南芥苗干重比值。将对照处理 (*E.coli* DH5 $\alpha$  挥发性有机物处理) 数值设定为 1。枯草芽孢杆菌挥发性有机物处理的展示结果为测量数值与对照处理数值的比值。图中结果为每种处理 (4 棵拟南芥苗) 进行 3 次测量所得平均值, 误差线表示标准差。星号表示 *t* 检验显示枯草芽孢杆菌处理跟大肠杆菌处理间的差异具有显著性 ( $P < 0.05$ )

(A). Representative pictures of different effects of *Escherichia coli* DH5 $\alpha$  and *Bacillus subtilis* ACCC 01380 volatile organic compounds on *Arabidopsis* growth. (B) The ratio of main root length of *Arabidopsis* seedlings treated with *Escherichia coli* DH5 $\alpha$  or *Bacillus subtilis* ACCC 01380 VOCs. (C) The ratio of shoot dry weight of *Arabidopsis* seedlings treated with *Escherichia coli* DH5 $\alpha$  or *Bacillus subtilis* ACCC 01380 VOCs. The value of the control treatment (*E.coli* DH5 $\alpha$  volatile organic compounds treatment) is set to 1. The displayed number is the ratio of *B. subtilis* VOCs treatment to the control treatment. The results in the figure are the average of three measurements for each treatment (four *Arabidopsis* seedlings), and the error bars represent the standard deviation. Asterisks indicate a significant difference between *Bacillus subtilis* treatment and *E.coli* treatment ( $P < 0.05$ ) with *t*-test

图 1 枯草芽孢杆菌的挥发性有机物促进拟南芥生长

Figure 1 The promotion of *Bacillus subtilis* ACCC 01380 VOCs to the growth of *Arabidopsis*

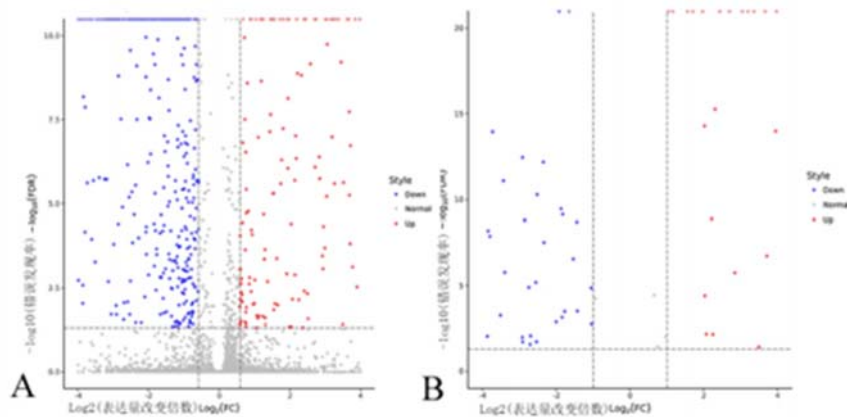


图 2 大肠杆菌或枯草芽孢杆菌 ACCC 01380 VOCs 处理 48 h 的拟南芥根之间差异表达的 mRNA (A) 和 lncRNAs (B)  
Figure 2 Differentially expressed protein coding genes(A) and lncRNAs(B) between *Arabidopsis* roots treated with DH5 $\alpha$  or ACCC 01380 VOCs for 48 hours

### 2.4 差异表达 lncRNA 的顺式作用靶基因及其功能注释

Antisense lncRNAs 和 *cis*-lncRNAs (sense 链上

临近位置) 可以作为顺式调节元件影响临近的蛋白质编码基因的表达<sup>[25]</sup>, 进而影响植物各种生理过程。我们预测了差异表达 lncRNA 的顺式作用靶基

因。受到解淀粉芽孢杆菌挥发性有机物调控的 233 个 TAIR 10 lncRNA 中, 有 192 个 lncRNA, 其上下游附近 (10 kb) 存在蛋白编码基因。位于这 192 个 lncRNA 的附近 (上下游 10 kb 范围) 的蛋白质编码基因中,

有 94 个功能基因的表达, 受到了 ACCC 01380 挥发性有机物的影响, 其中包括对拟南芥生长和抗逆具有关键影响的功能基因。

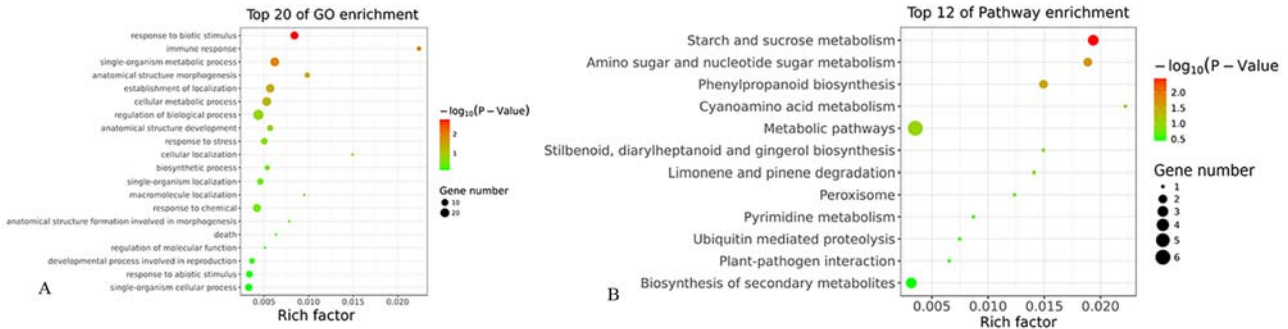
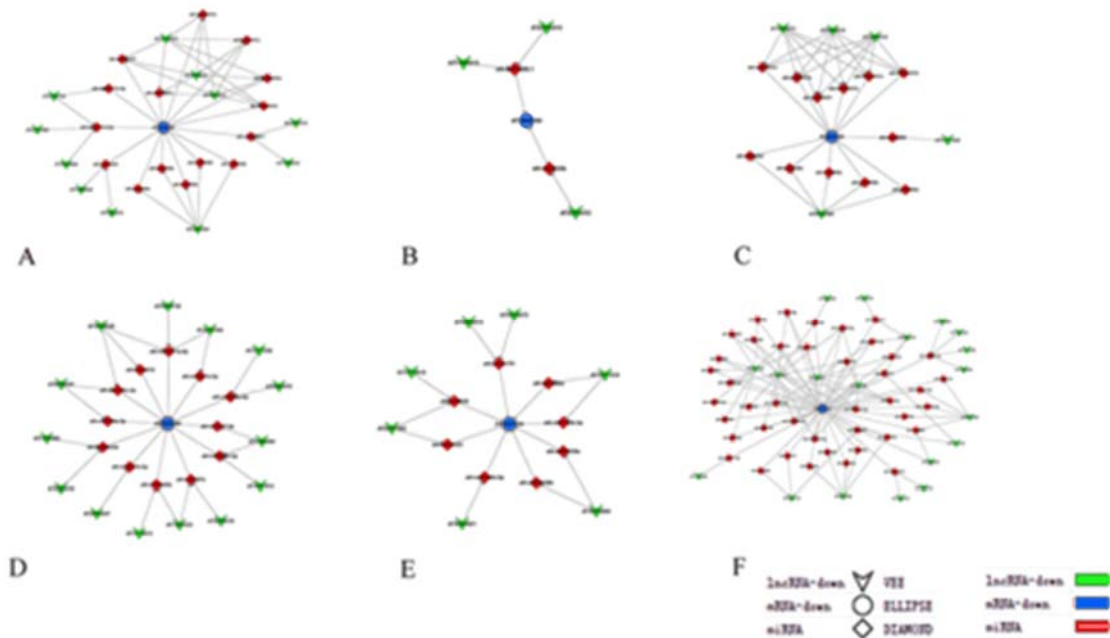


图 3 ACCC 01380 挥发性有机物处理下发生差异表达的 lncRNA 附近位置的 95 个蛋白编码基因的 GO (A) 和 KEGG (B) 富集分析

Figure 3 GO(A) and KEGG(B) enrichment analysis of 95 protein-coding genes *cis* regulated by the lncRNAs treated with ACCC 01380 VOCs



(A) AT1G06000, (B) AT2G30360, (C) AT3G24520, (D) AT3G63440, (E) AT4G37430, (F) AT5G41610. 蓝色, 红色和橙色节点分别表示 mRNA、miRNA 和 lncRNA。每个灰色边缘表示 miRNA 和 mRNA 或 lncRNA 之间的靶标关系  
(A)AT1G06000, (B) AT2G30360, (C) AT3G24520, (D) AT3G63440, (E) AT4G37430, (F)AT5G41610. The sky-blue, red, and orange nodes denote protein coding genes, miRNAs, and lncRNAs, respectively. Each gray edge denotes a target relationship between a miRNA and a protein coding gene or lncRNA

图 4 差异表达 lncRNAs 的 ceRNA 调控网络  
Figure 4 ceRNA regulatory networks of differentially expressed lncRNAs

根据 Ryu 等<sup>[26]</sup>的结果, 细胞分裂素在 GB03 挥发性有机物促进拟南芥叶片生长过程中, 具有关键作用。GB03 的挥发性有机物不能促进细胞分裂素受体基因的突变体 *cre1* 植株的叶片生长。本研究发现, 编码催化细胞分裂素降解的基因, CKX6 (AT3G63440), 受到 ACCC 01380 挥发性有机物下

调表达 (Log<sub>2</sub>FC 为 -14.58, FDR 为 0)。而该基因序列上存在一个 Antisense lncRNA, AT3G63445, 也受到 ACCC 01380 挥发性有机物下调表达 (Log<sub>2</sub>FC 为 -1.82, FDR 为 7.28E-10)。

GB03 挥发性有机物可以提高拟南芥耐盐胁迫能力。拟南芥 *HKT1* 基因在这一过程中发挥重要作用<sup>[1]</sup>。



全转录组分析表明,拟南芥参与离子转运的基因, *ATCHX18* (*AT5G41610*) 在根部被 ACCC 01380 的挥发性有机物抑制 ( $\text{Log}_2\text{FC}$  为  $-15.70$ ,  $\text{FDR}$  为  $0$ )。该基因序列中存在的 Antisense lncRNA, *AT5G41612*, 也被 ACCC 01380 的挥发性有机物抑制表达 ( $\text{Log}_2\text{FC}$  为  $-14.19$ ,  $\text{FDR}$  为  $0$ )。

芥子油苷(Glucosinolate)是一种生物活性化合物,在胁迫响应、抗氧化活性和抗病菌侵入等方面具有重要的调控作用。脂肪芥子油苷(Aliphatic glucosinolate)的生物合成途径相关基因的表达被 GB03 挥发性有机物上调,导致 Glucosinolate 含量增加,提高了拟南芥抗虫能力<sup>[27]</sup>。细胞色素 P450 家族基因 *CYP81* (*AT4G37430*) 是 indolic glucosinolate 生物合成途径的关键酶。全转录组分析表明,该基因受到 ACCC 01380 挥发性物质下调表达 ( $\text{Log}_2\text{FC}$  为  $-16.23$ ,  $\text{FDR}$  为  $0$ )。该基因 DNA 编码序列中存在一个反义 lncRNA, *AT4G37432*, 也被 ACCC 01380 挥发性有机物下调 ( $\text{Log}_2\text{FC}$  为  $-14.24$ ,  $\text{FDR}$  为  $0$ )。

此外,还有一些重要的功能基因,如 *SIP4* (*AT2G30360*, 蛋白激酶,调控质膜  $\text{H}^+$ -ATPase 的活性)、*HSFC1* (*AT3G24520*, 热激转录因子)、*UGT89C1* (*AT1G06000*, 调节生长素稳态的关键酶) 及 *ALMT3* (*AT1G18420*, 磷匮乏及铝胁迫响应基因) 等,均位于受到 ACCC 01380 挥发性有机物调控的 lncRNA 的顺式位置,且在 ACCC 01380 挥发性物质处理前后,表达发生变化。我们对这 94 个表达受到 ACCC 01380 挥发性有机物的影响且位于 lncRNA 附近的基因进行 GO 和 KEGG 富集分析。基于 GO 富集分析,我们发现可能受到 lncRNA 调控, ACCC 01380 挥发性有机物处理后表达发生变化的 94 个基因参与的生物过程,最显著富集的两个是 response to biotic stimulus 和 immune response (图 3A)。ACCC 01380 挥发性有机物处理,显然是一个生物刺激。这个生物刺激引起了拟南芥一定的免疫响应。我们的分析结果说明, lncRNA 在拟南芥对 ACCC 01380 挥发性有机物的免疫响应中,很可能起到重要的调控作用。此外,单生物代谢过程、解剖结构形态发生及定位建立 3 个 GO 条目也在 94 个蛋白编码基因中显著富集。这些是基础代谢的重要通路。

KEGG 富集分析的结果(图 3B)显示,差异表达 lncRNA 顺式作用的靶基因主要富集在基础的合成和代谢途径中,如淀粉和蔗糖代谢、氨基糖和核苷酸糖代谢以及苯丙烷生物合成。淀粉和蔗糖是植

物生长的物质基础。苯丙烷是 SA 合成的重要前体物质,在植物生长、繁殖和防御过程中起重要作用。这说明枯草芽孢杆菌的挥发性有机物可调控拟南芥特定 lncRNA,对拟南芥生长至关重要。同时,我们观察到差异表达 lncRNA 顺式作用靶基因在植物生物刺激响应途径中高度富集。这些数据表明 lncRNAs 可能在 ACCC 01380 的挥发性有机物对拟南芥的生理影响中起重要作用。

## 2.5 ceRNA 网络分析

lncRNA 可与 miRNA 结合以竞争性的抑制其与 mRNA 的结合,进而在功能基因表达调控中发挥重要作用,这一机制是 lncRNA 发挥生理功能的重要途径<sup>[23-24]</sup>。本研究中,我们选择表达量受到 ACCC 01380 挥发性有机物影响的重要功能基因及其相关 lncRNA,预测其序列中的 miRNA 靶标,并构建 lncRNA-miRNA-mRNA 的调节网络(图 4)。ceRNA 分析表明, lncRNA 通过与共有 miRNA 结合,影响这些 miRNA 对 mRNA 的调节作用,最终导致拟南芥许多具有重要生理影响的功能基因的表达改变。*At2g30360* (Other names: CIPK11, PKS5, SIP4), 编码一个 SOS2-like protein kinase, 可通过磷酸化调控拟南芥根部质膜  $\text{H}^+$ -ATPase 和 ABA 信号传导途径重要基因 *ABI5* 的活性<sup>[28]</sup>。拟南芥 *pks5* 突变体根际酸化能力增强,具有更强的耐碱胁迫能力<sup>[29-31]</sup>。ACCC 01380 挥发性有机物处理引起该基因表达下调( $\log_2\text{FC} = -17.197\ 385\ 44$ ,  $\text{FDR}=0$ )。

ceRNA 分析显示, lncRNA *AT2G46572/AT2G45315/AT1G48315* 可通过 miRNA *ath-miR833b/ath-miR1886.1* 调控 *At2g30360* 的表达(图 4B)。与 *At2g30360* 相同, ACCC 01380 挥发性有机物处理同样引起 lncRNA *AT2G46572* ( $\log_2\text{FC} = -12.075\ 151\ 43$ ,  $\text{FDR}=0.001\ 285\ 163$ ) / *AT2G45315* ( $\log_2\text{FC} = -13.358\ 531\ 56$ ,  $\text{FDR}=0$ ) / *AT1G48315* ( $\log_2\text{FC} = -1.429221458$ ,  $\text{FDR}=0.000\ 307\ 315$ ) 下调表达。以往的研究显示, GB03 挥发性有机物处理下调拟南芥 ABA 响应基因<sup>[2]</sup>, 提高拟南芥根际酸化能力<sup>[5]</sup>, 但具体分子机理还不清楚。lncRNA *AT2G46572/AT2G45315/AT1G48315* 可能在这一过程中起重要作用。

*At1g06000*(*UGT89C1*), 编码参与鼠李糖基化黄酮醇(rhamnosylated flavonols)合成过程中的黄酮-7-O-鼠李糖基转移酶(flavonol-7-O-rhamnosyltransferase)<sup>[32]</sup>。黄酮鼠李糖苷(Flavonol rhamnosides)是一大类植物次生代谢物质,可影响多个关键细胞过程。例如, flavonol 7-O-rhamnosides 在植物适应

紫外线长期照射和生长素稳态调节中起重要作用。ACCC 01380 挥发性有机物处理下, *At1g06000* 表达量下调 ( $\log_2FC = -17.693\ 907\ 21$ ,  $FDR=0$ )。ceRNA 分析表明, 多个 lncRNA 通过调控影响该基因的表达 (图 4A)。

细胞分裂素对于 GB03 挥发性有机物促进拟南芥生长的生理功能至关重要<sup>[26]</sup>。GB03 挥发性有机物不再能够促进细胞分裂素受体突变体 *cre* 的生长。AT3G63440(*AtCKX7*) 催化细胞分裂素的降解, 被 ACCC 01380 挥发性有机物下调表达 ( $\log_2FC = -14.578\ 265\ 57$ ,  $FDR=0$ )。ceRNA 分析表明, AT3G63440(*AtCKX7*) 受到 lncRNA 和 miRNA 复杂的网络调控 (图 4D)。

### 3 讨论与结论

相关研究表明, lncRNA 可以通过表观修饰或其他转录后调控机制影响关键功能基因的表达<sup>[33-34]</sup>。植物中, lncRNAs 在多个重要生物过程中起着关键作用<sup>[35]</sup>。根际促生芽孢杆菌的挥发性有机物对拟南芥生理过程有广泛的影响, 可以促进拟南芥植株生长, 提高拟南芥植株对盐、旱、病、虫等多个逆境的抗性。基因芯片研究表明, GB03 的挥发性有机物影响拟南芥根部 600 多个基因的表达, 其中包括生长素、脱落酸等激素信号传导关键基因, 以及铁、硫等营养元素吸收关键基因。但根际促生芽孢杆菌的挥发性有机物是否影响拟南芥 lncRNA 的表达? 根际促生芽孢杆菌的挥发性有机物对 lncRNA 表达的影响是否具有生理方面的影响?

本研究发现拟南芥根部大量 lncRNA 的表达受到 ACCC 01380 的挥发性有机物影响, 88 个 lncRNA 表达上调, 145 个 lncRNA 表达下调。顺式分析 (Cisanalysis) 表明, 这 233 个差异表达 lncRNA 中, 有 192 lncRNA, 其染色体位置上下游 10 kb 范围内存在蛋白编码基因; 筛选范围缩小到上下游 2 kb 范围内, 有 179 个 lncRNA, 该范围内存在蛋白编码基因。GO 和 KEGG 分析, 表明表达受到 ACCC 01380 的挥发性有机物影响的 lncRNA 调控拟南芥对生物刺激的相应和免疫响应过程, 以及淀粉和蔗糖代谢、Phenylpropanoid 生物合成通路。我们推测 lncRNAs 在枯草芽孢杆菌 ACCC 01380 挥发性有机物在对拟南芥的生理影响过程中起着重要作用。

一些研究表明, lncRNA 可以作为一种抑制 miRNA 功能的 ceRNA, 与 mRNA 竞争 miRNA 的结合<sup>[36-38]</sup>。本研究基于 ceRNA 作用原理构建了 lncRNA-miRNA-mRNA 相关网络。ceRNA 分析表

明, 受到 ACCC 01380 的挥发性有机物调控的 lncRNA, 可以通过 miRNA 影响重要功能基因的表达, 包括细胞分裂素降解基因 *AtCKX7* 和 SOS2-like protein kinase 等。这表明 lncRNA 可能在 ACCC 01380 的挥发性有机物影响拟南芥生长和抗性的过程中起重要作用。例如, lncRNA AT2G46572/AT2G45315/AT1G48315 可以作为 miRNA ath-miR833b/ath-miR1886.1 的 eTM (endogenous microRNA target mimic, eTM), 调节 mRNA At2g30360 (Other names: CIPK11, PKS5, SIP4) 的表达。这表明 lncRNA 可能通过形成 ceRNA 参与 ACCC 01380 挥发性有机物对拟南芥的生理影响。为了揭示其潜在机制, 需要充分利用拟南芥的突变体资源库, 进一步研究筛选的 lncRNA 作为 ceRNA 在 ACCC 01380 挥发性有机物调控拟南芥生理影响的作用机理。

在本研究中, 我们分析了 ACCC 01380 挥发性有机物对拟南芥根部 lncRNA 的表达影响, 并筛选和鉴定了一系列表达受到 ACCC 01380 挥发性有机物调控的 lncRNA。通过分析差异 lncRNA 顺式靶基因的 GO 和 KEGG 数据, 我们发现拟南芥根部 lncRNAs 可能在 ACCC 01380 挥发性有机物的生理作用中起重要作用。同时, 我们还发现 lncRNA 可能通过形成 ceRNA 影响 ACCC 01380 挥发性有机物的生理作用。本研究结果显著提高了我们对根际促生芽孢杆菌挥发性有机物的生理影响中 lncRNA 作用的理解。该研究为进一步研究 lncRNA 功能以及根际促生芽孢杆菌挥发性有机物作用机理奠定了基础。

### 参考文献:

- [1] ZHANG H M, KIM M, SUN Y, et al. Soil bacteria confer plant salt tolerance by tissue-specific regulation of the sodium transporter *HKT1*[J]. *MPMI*, 2008, 21(6): 737-744.
- [2] ZHANG H M, XIE X T, KIM M, et al. Soil bacteria augment Arabidopsis photosynthesis by decreasing glucose sensing and abscisic acid levels in planta[J]. *Plant J*, 2008, 56(2): 264-273.
- [3] ZHANG H M, KIM M, KRISHNAMACHARI V, et al. Rhizobacterial volatile emissions regulate auxin homeostasis and cell expansion in Arabidopsis[J]. *Planta*, 2007, 226(4): 839-851.
- [4] ZHANG H M, MURZELLO C, SUN Y, et al. Choline and osmotic-stress tolerance induced in Arabidopsis by the soil microbe *Bacillus subtilis*(GB03)[J]. *MPMI*, 2010, 23(8): 1097-1104.
- [5] ZHANG H M, SUN Y, XIE X T, et al. A soil bacterium regulates plant acquisition of iron via deficiency-inducible mechanisms[J]. *Plant J*, 2009, 58(4): 568-577.
- [6] RUDRAPPA T, BIEDRZYCKI M L, KUNJETI S G, et al. The rhizobacterial elicitor acetoin induces systemic resistance in *Arabidopsis thaliana*[J]. *Commun Integr Biol*, 2010, 3(2): 130-138.

- [7] KIM E, SUNG S. Long noncoding RNA: unveiling hidden layer of gene regulatory networks[J]. Trends Plant Sci, 2012, 17(1): 16-21.
- [8] PONTING C P, OLIVER P L, REIK W. Evolution and functions of long noncoding RNAs[J]. Cell, 2009, 136(4): 629-641.
- [9] CHEN L L. Linking long noncoding RNA localization and function[J]. Trends Biochem Sci, 2016, 41(9): 761-772.
- [10] SIGNAL B, GLOSS B S, DINGER M E. Computational approaches for functional prediction and characterisation of long noncoding RNAs[J]. Trends Genet, 2016, 32(10): 620-637.
- [11] PACHNIS V, BRANNAN C I, TILGHMAN S M. The structure and expression of a novel gene activated in early mouse embryogenesis[J]. EMBO J, 1988, 7(3): 673-681.
- [12] BEN AMOR B, WIRTH S, MERCHAN F, et al. Novel long non-protein coding RNAs involved in Arabidopsis differentiation and stress responses[J]. Genome Res, 2008, 19(1): 57-69.
- [13] ZHANG W, HAN Z, GUO Q, et al. Identification of maize long non-coding RNAs responsive to drought stress[J]. PLoS One, 2014, 9(6): e98958.
- [14] DI C, YUAN J P, WU Y, et al. Characterization of stress-responsive lncRNAs in *Arabidopsis thaliana* by integrating expression, epigenetic and structural features[J]. Plant J, 2014, 80(5): 848-861.
- [15] WANG H, CHUNG P J, LIU J, et al. Genome-wide identification of long noncoding natural antisense transcripts and their responses to light in Arabidopsis[J]. Genome Res, 2014, 24(3): 444-453.
- [16] WANG Y Q, WANG X C, DENG W, et al. Genomic features and regulatory roles of intermediate-sized non-coding RNAs in Arabidopsis[J]. Mol Plant, 2014, 7(3): 514-527.
- [17] ZHANG Y C, LIAO J Y, LI Z Y, et al. Genome-wide screening and functional analysis identify a large number of long noncoding RNAs involved in the sexual reproduction of rice[J]. Genome Biol, 2014, 15(12): 512.
- [18] LU T T, ZHU C R, LU G J, et al. Strand-specific RNA-seq reveals widespread occurrence of novel cis-natural antisense transcripts in rice[J]. BMC Genom, 2012, 13(1): 721.
- [19] LI L, EICHTEN S R, SHIMIZU R, et al. Genome-wide discovery and characterization of maize long non-coding RNAs[J]. Genome Biol, 2014, 15(2): R40.
- [20] WEN J, PARKER B J, WEILLER G F. In silico identification and characterization of mRNA-like noncoding transcripts in *Medicago truncatula*[J]. Silico Biol, 2007, 7(5): 485-505.
- [21] SHUAI P, LIANG D, TANG S, et al. Genome-wide identification and functional prediction of novel and drought-responsive lincRNAs in *Populus trichocarpa*[J]. J Exp Bot, 2014, 65(17): 4975-4983.
- [22] XIN M M, WANG Y, YAO Y Y, et al. Identification and characterization of wheat long non-protein coding RNAs responsive to powdery mildew infection and heat stress by using microarray analysis and SBS sequencing[J]. BMC Plant Biol, 2011, 11(1): 61.
- [23] DENG F N, ZHANG X P, WANG W, et al. Identification of *Gossypium hirsutum* long non-coding RNAs (lncRNAs) under salt stress[J]. BMC Plant Biol, 2018, 18: 23.
- [24] XU X W, ZHOU X H, WANG R R, et al. Functional analysis of long intergenic non-coding RNAs in phosphate-starved rice using competing endogenous RNA network[J]. Sci Rep, 2016, 6: 20715.
- [25] SWIEZEWSKI S, LIU F Q, MAGUSIN A, et al. Cold-induced silencing by long antisense transcripts of an Arabidopsis polycomb target[J]. Nature 2009, 462(7274): 799-802.
- [26] RYU C M, FARAG M A, HU C H, et al. Bacterial volatiles promote growth in Arabidopsis[J]. Proc Natl Acad Sci U S A, 2003, 100(8): 4927-4932.
- [27] AZIZ M, NADIPALLI R K, XIE X T, et al. Augmenting sulfur metabolism and herbivore defense in Arabidopsis by bacterial volatile signaling[J]. Front Plant Sci, 2016, 7: 458.
- [28] ZHOU X N, HAO H M, ZHANG Y G, et al. SOS2-LIKE PROTEIN KINASE5, an SNF1-RELATED PROTEIN KINASE3-type protein kinase, is important for abscisic acid responses in Arabidopsis through phosphorylation of ABCISIC ACID-INSENSITIVE5[J]. Plant Physiol, 2015, 168(2): 659-676.
- [29] YANG Y Q, QIN Y X, XIE C G, et al. The Arabidopsis chaperone J3 regulates the plasma membrane H<sup>+</sup>-ATPase through interaction with the PKS5 kinase[J]. Plant Cell, 2010, 22(4): 1313-1332.
- [30] LIU J T, GUO Y. The alkaline tolerance in Arabidopsis requires stabilizing microfilament partially through inactivation of PKS5 kinase[J]. J Genet Genomics, 2011, 38(7): 307-313.
- [31] XU W, JIA L, SHI W, et al. The tomato 14-3-3 protein TFT4 modulates H<sup>+</sup> efflux, basipetal auxin transport, and the PKS5-J3 pathway in the root growth response to alkaline stress[J]. Plant Physiol, 2013, 163(4): 1817-1828.
- [32] ZONG G N, FEI S, LIU X, et al. Crystal structures of rhamnosyltransferase UGT 89C1 from *Arabidopsis thaliana* reveal the molecular basis of sugar donor specificity for UDP-β-l-rhamnose and rhamnosylation mechanism[J]. Plant J, 2019, 99(2): 257-269.
- [33] LIU X, HAO L, LI D, et al. Long non-coding RNAs and their biological roles in plants[J]. Genom Proteom Bioinform, 2015, 13(3): 137-147.
- [34] BAZIN J, BAILEY-SERRES J. Emerging roles of long non-coding RNA in root developmental plasticity and regulation of phosphate homeostasis[J]. Front Plant Sci, 2015, 6: 400.
- [35] ZHANG Y C, CHEN Y Q. Long noncoding RNAs: new regulators in plant development[J]. Biochem Biophys Res Commun, 2013, 436(2): 111-114.
- [36] XU X W, ZHOU X H, WANG R R, et al. Functional analysis of long intergenic non-coding RNAs in phosphate-starved rice using competing endogenous RNA network[J]. Sci Rep, 2016, 6: 20715.
- [37] SALMENA L, POLISENO L, TAY Y, et al. A ceRNA hypothesis: the rosetta stone of a hidden RNA language?[J]. Cell, 2011, 146(3): 353-358.
- [38] RUBIO-SOMOZA I, WEIGEL D, FRANCO-ZORILLA J M, et al. CeRNAs: miRNA target mimic mimics[J]. Cell, 2011, 147(7): 1431-1432.