高通量测序解析大青叶有效成分的生物合成途径

陈军峰¹, 李卿¹, 王芸¹, 张青磊², 张磊^{2*}, 陈万生^{1*}

1. 第二军医大学长征医院药学部, 上海 200003;

2. 第二军医大学药学院生药学教研室, 上海 200433

* 联系人, E-mail: zhanglei@smmu.edu.cn; chenwansheng@smmu.edu.cn

收稿日期: 2017-10-21; 接受日期: 2017-12-26; 网络版发表日期: 2018-02-22

国家自然科学基金(批准号: 81325024, 31100221, 31670292)资助

摘要 中药大青叶是菘蓝的干燥叶片, 主要活性物质为吲哚类化合物, 包括靛蓝、靛玉红、色胺酮等, 以上吲哚类化合物的生物合成起始于色氨酸途径, 目前在植物中的合成机制还未完全阐明. 本研究以成熟菘蓝叶片为研究对象, 对茉莉酸甲酯诱导的叶片进行了转录组分析, 对已知途径进行了转录组注释, 并对未解析途径进行了预测分析. 分别获得了色氨酸代谢途径中色氨酸、吲哚乙酸、吲哚苷和萆类吲哚生物碱合成几个代谢支路中16个催化步骤的38个编码基因. 通过共表达网络分析, 预测转录因子bHLH125可能对吲哚途径具有核心调控作用; CYP2A6-1和CYP735A2等蛋白可能催化生成靛蓝、靛玉红、色胺酮前体的羟基化反应, 对这两个蛋白与底物分子的结合进行分子对接建模, 均显示对吲哚分子具有较好的结合力. 本研究为后续开展菘蓝代谢调控和育种, 以及开展吲哚类物质合成生物学研究提供候选基因.

关键词 大青叶, 吲哚类成分, 转录组分析, 共表达网络

十字花科菘蓝属(*Isatis*)植物在亚洲和欧洲具有悠久的历史, 部分种的叶和根均具有药用价值, 同时叶片可提取蓝色染料. 在我国, 菘蓝作为药材被广泛种植和应用, 根入药为板蓝根, 叶为大青叶, 主要栽培种为菘蓝(*Isatis indigotica* Fort.)^[1]. 临床上对于流行性感冒、流行性乙型脑炎、肝炎、腮腺炎等疾病具有显著疗效^[2].

菘蓝的主要活性物质有两大类, 分别为木脂素及吲哚类化合物. 目前在板蓝根中已明确的活性物质主要为木脂素类化合物, 包括松柏醇、落叶松脂素^[3]及其糖苷化衍生物^[4,5], 这些物质对流感病毒引起的炎症

反应及甲、乙型流感病毒具有显著的抑制活性. 然而, 这些活性物质在板蓝根中含量很低(<2‰), 大幅提高药材中这些活性成分的含量, 对提高菘蓝材品质具有重要意义. 以此为目标, 本课题组整合高通量代谢组学、转录组学分析手段完成了菘蓝中木脂素合成途径的表征, 研究了合成途径催化酶的特征^[6,7], 构建转录调控网络, 预测其限速步骤和关键调控因子^[8], 精细描绘了菘蓝中木脂素合成途径的转录特征和转录调控机制, 为开展板蓝根的代谢调控及育种提供了详实的理论依据.

吲哚类物质在大青叶中大量富集, 是其主要化学

引用格式: 陈军峰, 李卿, 王芸, 等. 高通量测序解析大青叶有效成分的生物合成途径. 中国科学: 生命科学, 2018, 48: 412–422

Chen J F, Li Q, Wang Y, et al. Comprehensive transcriptomic profiling of *Isatis indigotica* leaves reveals biosynthesis of indole active ingredients (in Chinese). Sci Sin Vitae, 2018, 48: 412–422, doi: 10.1360/N052017-00195

成分, 结构具有多样性^[9-11]. 通过细胞膜色谱筛选, 证明靛玉红(indirubin)及色安酮(tryptanthrin)均具有良好的抗白血病活性, 是大青叶中主要的活性成分^[12]. 在植物中, 以上吲哚化合物的生物合成属于色氨酸代谢途径的分支, 目前对这个途径的研究还十分欠缺. 首先, 普遍认为靛玉红、靛蓝、色安酮等活性成分是由前体化合物吲哚酚(indoxyl)、靛红(isatin)自发氧化形成的二聚体^[13-16], 但是生成这些前体化合物的催化酶还未明确; 其次, 虽然植物中色氨酸代谢途径的转录调控机制研究已经比较深入^[17], 但是针对靛蓝、靛玉红合成这个分支却鲜有研究; 此外, 吲哚途径在菰蓝中的种属特异性、器官特异性也都需要进一步解析. 基于以上问题, 本研究开展了菰蓝的转录组注释及表达谱分析, 预测参与菰蓝叶中吲哚生物合成途径的催化酶基因以及潜在的调控基因, 帮助研究人员更加深入地了解吲哚类物质在菰蓝中合成的特异性, 从而为进一步利用代谢工程、合成生物学等手段提升菰蓝中吲哚活性物质含量而提供理论依据.

1 材料与方法

1.1 植物材料与诱导

菰蓝种植于上海浦东益大本草园, 于9月播种, 选用萌发后5个月的成熟植株开展实验. 选定叶片后, 用喷壶将茉莉酸甲酯(methyl jasmonate, MeJA) (10 mmol/L)均匀喷洒于菰蓝叶片表面, 收集未喷洒叶片为对照, 之后依次在处理1, 3, 6, 12, 24 h后采摘同株叶片, 每个样品分为两份, 一份用于化合物检测, 一份液氮冷冻后, 于-80℃保存用于RNA提取. 每个处理设置3个生物学重复.

1.2 化合物检测

干燥菰蓝叶片研磨成粉末, 取200 mg粉末用10 mL溶剂(甲醇:氯仿=1:1)超声提取3次. 取1 mL提取液挥发干燥, 100 μL甲醇复溶, 用于化学检测. 化合物定量使用三重四级杆质谱(Agilent 6410, Agilent, 美国)测定. 色谱柱使用Agilent Poroshell 120 EC-C18, column (3.0X 150 mm, 2.7 μm particle size). 流动相为乙腈:甲醇(50:50), 流速0.3 mL/min, 进样时间15 min. 目标化合物的定量采用多重反应模式(multiple reaction monitoring, MRM)检测, 正离子模式下产物荷比(m/z)如下: 靛蓝, 263.1→190; 靛玉红, 263.1→219; 吲哚, 249.1→130; 色安酮, 249.1→130.

1.3 转录组测序与分析

(1) 基因注释. 依据已报道菰蓝转录组数据组^[6], 以KEGG数据库吲哚合成途径途径和CYP数据库(<http://drnelson.uthsc.edu/Arabidopsis.Blast.file.html>)为参考基因, 使用Local Blast进行基因比对注释($E < 1.00 \times 10^{-5}$);

(2) 转录组测序和分析. 18个样本(6时间点×3重复) RNA采用Illumina TruSeq RNA Sample Prep Kit (Illumina, 美国)分别建库, 利用Illumina HiSeq2000进行双端100 bp测序, 每个样本生产3G数据量. 数据去杂后, 使用Trinity *de novo* assembler (<http://trinityrnaseq.sourceforge.net/>)进行组装, 重叠序列最小一致性90%. 转录本表达丰度以RPKM (reads per kilobase of exon model per million mapped reads)定义, 以0 h样本为对照, 比较每组转录本在样本间差异的显著性($FDR < 0.5$)找出差异表达基因;

(3) 表达谱分析. 采用MEV (Ver 4.9.0)软件进行基因表达热图和表达模式分析; 皮尔森相关系数计算依据表达谱, 采用如下公式计算:

$$r = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{S_X} \right) \left(\frac{Y_i - \bar{Y}}{S_Y} \right).$$

所得基因对根据关联性的相关系数(正相关>0.95, 负相关<-0.95)和显著性分析(<0.01)进行筛选, 筛选出来的基因对使用Cytoscape (Ver 3.4)进行可视化分析进行共表达网络构建.

1.4 qRT-PCR检测

菰蓝叶片总RNA提取采用RNAprep pure Plant Kit (天根, 北京). 定量后每个样品取1 μg总RNA反转录为cDNA. PCR试剂使用TransStart Green qPCR SuperMix (全式金, 北京), 基因扩增在TAKARA-TP800定量PCR仪上运行(TaKaRa, 日本). 每个检测样本设3个生物学重复, *LiPolyubiquitin1*为内参^[6], 用 $2^{-\Delta\Delta C_t}$ 法对基因表达量行相对定量分析.

1.5 同源建模和分子对接

CYP蛋白结构以CYP2A6为模板(PDB ID: 3ebs),

使用MODELER 9v11进行同源建模. 分子对接用Site finder进行结合预测, 采样数100, 取top20进行分析.

2 结果与分析

2.1 菰蓝中吲哚生物合成途径

植物中吲哚类物质的次生代谢途径起始于色氨酸(tryptophan)代谢途径, 以KEGG数据库色氨酸途径(ko00380)为参考, 对菰蓝的吲哚合成途径进行基因注释, 共预测催化16个步骤的38条基因(表1). 植物中色氨酸途径包括4个支路, 分别为吲哚苷(indole glucosinolate, IG)、茚乙酸(indole-3-acetic acid, IAA)、萜类吲哚生物碱(terpene indole alkaloids, TIA), 以及生成靛蓝、靛玉红(图1). 吲哚(indole)是整个途径网络的主要节点, 由分支酸(chorismate)起始生成吲哚, 进而生成各类吲哚衍生物. 分支酸依次被ASA, AnPRT, PAI, IPS, TSA催化生成吲哚, 吲哚由TSB催化生成色氨酸. 通过注释共获得该路径上7个催化步骤的13条编码基因. 吲哚乙酸(indole-3-acetic acid, IAA)是植物中普遍存在的

重要生长激素, 其生物合成起始于色氨酸, 通过注释共获得该路径上5个催化步骤的17条编码基因. TDC具有2个unigenes, 催化色氨酸脱羧生成色胺(tryptamine), 与萜类代谢途径合并生成TIAs. 靛蓝、靛玉红等吲哚成分在植物中合成途径还未完全阐明, 无法直接通过参考基因进行注释, 需要进一步通过转录分析进行预测和注释.

2.2 吲哚类成分在菰蓝中的分布

大青叶和板蓝根中主要活性成分的累积特征差异显著^[18,19]. 大青叶的主要活性成分为靛蓝、靛玉红和色安酮, 对从大青叶提取的青黛进行含量分析(图2A), 靛蓝的比例占51.38%, 靛玉红比例为2.15%, 色安酮0.21%, 显示吲哚类物质是青黛中的主要成分; 同时说明这几个化合物虽然衍生自相同前体, 但后续途径中的催化水平差异较大. 此外, 这3种化合物在菰蓝根和叶中的累积水平差异巨大(图1A), 靛蓝、靛玉红和色安酮在菰蓝叶中的含量分别为331.08, 132.43和4.35 $\mu\text{g/g}$, 含量水平均远高于在根部的累积(靛蓝

图1 根据转录组注释预测的菰蓝吲哚类化合物生物合成途径

虚线箭头表示多步催化反应, 蓝色箭头表示靛蓝、靛玉红合成途径; 基因名后数字为同源异构体数目

表 1 菰蓝吲哚生物合成途径基因的转录组注释

编码蛋白	同源异构体数	注释号
色氨酸合成途径		
甲酸酯合成酶A (anthranilate synthase A, ASA)	—	comp29842_c1_seq1
ASB	—	comp14924_c0_seq1
邻氨基苯甲酸磷酸核糖基转移酶(anthranilate phosphoribosyltransferase, AnPRT)	—	comp29673_c0_seq2
磷酸核糖氨基苯甲酸异构酶(phosphoribosylanthranilate isomerase, PAI)	—	comp34848_c0_seq1
吲哚甘油磷酸合成酶(indoleglycerol phosphate synthetase, IPS)	3	IPS1(comp23515_c0_seq1); IPS2(comp23515_c0_seq2); IPS3(comp28963_c0_seq1)
色氨酸合成酶A (trptophan synthase α subunit, TSA)	3	TSA1(comp17779_c0_seq1); TSA2(comp23941_c0_seq2); TSA3(comp23941_c0_seq4)
TSB	3	TSB1(comp33572_c0_seq1); TSB2(comp35312_c0_seq1); TSB3(comp23941_c0_seq3)
萘类吲哚生物碱合成途径		
色氨酸脱羧酶(tryptophan decarboxylase, TDC)	2	TDC1(comp23341_c0_seq1); TDC2(comp23625_c0_seq2)
吲哚苷合成途径		
细胞色素酶(CYP79B2/B3)	—	comp15803_c0_seq1
芥子油苷裂解酶(SUR)	3	SUR1(comp31255_c0_seq1); SUR2(comp31255_c0_seq4); SUR3(comp35622_c0_seq1)
S-糖基转移酶(UDP glucosyl transferase 74B1, UGT74B1)	1	comp34352_c0_seq1
硫酸基转移酶(ST)	1	comp34528_c0_seq1
吲哚乙酸合成途径		
色氨酸氨基转移酶(tryptophan aminotransferase in <i>Arabidopsis</i> , TAA)	4	TAA1(comp29256_c0_seq3); TAA2(comp29256_c0_seq1); TAA3(comp19746_c0_seq1); TAA4(comp24399_c0_seq1)
色氨酸氨基转移酶(TAR)	—	comp24399_c0_seq1
黄素单加氧酶(YUCCA)	6	YUC6(comp12786_c0_seq1); YUC3 (comp18376_c0_seq1); YUC4(comp2702_c0_seq1); YUC1(comp31232_c0_seq1); YUC1(comp31232_c0_seq2); YUC9(comp31266_c0_seq1)
脱落酸醛氧化酶(ABA-aldehydeoxidase 1, AAO1)	3	AAO3(comp31556_c0_seq1); AAO3(comp31556_c0_seq2); AAO1(comp31556_c1_seq1)
酰胺酶(AMI1)	3	AMI1(comp29196_c0_seq1); AMI2(comp29196_c0_seq2); AMI3(comp31139_c0_seq1)

6.10 $\mu\text{g/g}$ 、靛玉红1.68 $\mu\text{g/g}$ 、色安酮0.76 $\mu\text{g/g}$), 该结果显示, 菰蓝中的吲哚类成分主要在叶片中合成和累积, 提示菰蓝叶片是适合研究吲哚合成途径的特异器官.

MeJA可以诱导菰蓝毛状根中吲哚途径基因发生显著的转录变化^[6]. 以靛蓝、靛玉红以及共同前体化合物吲哚为检测对象, 利用LC-MS/MS检测了MeJA诱导后4个化合物在菰蓝叶中的含量变化. 如图2C所示,

图2 菰蓝中吲哚类活性成分的含量变化

A: 青黛中活性成分的组成; B: 吲哚类成分在菰蓝不同组织中的含量; C: MeJA诱导后菰蓝叶中吲哚类成分的含量变化, 误差线表示3个生物重复的标准偏差, 差异显著性分析由 t -test方差分析显示与, *: $P < 0.05$, 具有显著差异; **: $P < 0.01$, 差异极显著

吲哚累积量在诱导24 h后提高了2.18倍, 靛蓝提高了1.40倍, 靛玉红提高1.61倍, 而色胺酮含量无显著变化。

2.3 MeJA诱导菰蓝叶中吲哚合成途径的转录组分析

MeJA可以通过调控合成途径的转录水平而引起吲哚类化合物累积的显著变化。对MeJA诱导后24 h内的菰蓝叶片进行转录组测序, 依据基因注释对吲哚途径上的36个unigenes的表达谱进行转录变化分析(图3A)。结果显示, 整个合成途径中基因的转录变化具有多样性。依据表达模式HLC聚类分析, 36个基因的表达主要呈7个共表达模式(图3B~H)。其中集合B~D中主要由上调表达基因富集(图3B~D), 提示这些基因可能与吲哚物质合成水平提升有关。值得注意的是, 色氨酸合成酶的两类亚基(TSA和TSB)通常形成 $\alpha\beta\alpha$ 形式四聚体以更有效地行使催化功能^[20-22]。根据共表达聚类TSA2与TSB3具有较高共表达关系(图3D), 提示在TSA和TSB的各自3个unigenes中, TSA2和TSB3更有可能形成聚合体。

吲哚类物质在菰蓝叶中的高积累提示其在叶中具有特异的转录调控机制, 依据相关基因的共表达相关性, 建立了共表达转录调控网络(图4A)(相关系数 >0.95 或 <-0.95)以演示其调控机制。该网络包括144个基因之间662对显著相关性, 其中464对为正相关, 198对为负相关。为使共表达网络更简洁易于解读, 提高皮尔森相关系数筛选阈值(>0.98 或 <-0.98)构建共表达网络(图4B)。其中转录因子bHLH125-2分别与ASA和TSB1显现较高相关性, 提示该转录因子可能通过对色

氨酸途径的上游的调控影响靛蓝和靛玉红等物质的生物合成。该网络的其他转录因子均显示对色氨酸代谢网络的其他支路基因具有较高关联。

目前靛蓝、靛玉红和色胺酮几个物质在植物中的催化合成机制还未完全阐明, 吲哚经不同羟基化酶催化分别在2-位或3-位发生羟基化反应, 然后再糖基化或氧化生成吲哚苷(indican)和靛红(isatin)。这两个前体分别自发发生二聚化反应生成靛蓝和靛玉红^[15,16]。前体吲哚苷和靛红的合成水平直接决定大青叶中活性成分的含量。目前, 植物中生成吲哚苷和靛红的催化酶还没有被发现。但是一些异源蛋白被报道对吲哚具有催化能力, 为在菰蓝中解析这些催化步骤提供了依据。例如, 哺乳动物中分离的CYP2A6^[13], 微生物中的萘双加氧酶(naphthalene 1,2-dioxygenase)^[23]和酚羟化酶(phenol hydroxylase)^[24]均可催化吲哚的羟化反应, 提示菰蓝中的同源蛋白可能也具有相同功能。经同源比对, 菰蓝中没有基因可注释为萘双加氧酶和萘双加氧酶, 而4条unigenes (comp33180_c0_seq5, comp33180_c0_seq6, comp33180_c0_seq1, comp33180_c0_seq2)与CYP2A6具有显著的同源性, 提示菰蓝中这一反应更可能为CYP蛋白催化完成。此外, 通过构建共表达网络也显示出一些CYP可能参与到吲哚途径中。图4C为129个菰蓝CYP基因与吲哚合成基因的共表达聚类分析。结果显示, 7个基因集中17个吲哚途径基因与21个CYP基因具有共表达模式。以这些基因间的皮尔森相关系数建立共表达网络(图4D), 其中CYP735A2, 79B3, 71B6, 707A1, 71B21与TSA1, TSB1等吲哚合成基因具有极高相关性, 提示这些CYP可能与吲哚合成途径相

图3 MeJA诱导后菰蓝叶中吲哚合成途径的转录变化

A: 1, 3, 6, 12, 24 h表示诱导后不同检测时间点样本中目标基因的转录水平分别与对照(0 h)相比较, HLC聚类后7个主要表达模式在热图右侧标出, 图B~H分别显示了7个表达模式中基因的变化规律, 灰色折线代表基因变化规律, 红色代表基因集内的平均转录变化

关, 其中已知*CYP79B3*属于吲哚分支^[25].

使用qRT-PCR对这些候选CYP基因进行器官表达特征研究, *CYP735A2*和*CYP2A6-1*在叶中的转录水平显著高于根中(图5A). 对这两个基因对MeJA的响应模式进行检测, 结果显示两个基因都在MeJA的诱导下发生了转录上调(图5B和C), 特别是*CYP735A2*在诱导1 h检测到极显著的转录上调. 以上结果显示, 两个候选基因*CYP735A2*与*CYP2A6-1*的转录变化特征与菰蓝叶中吲哚类成分累积的变化特征具有一定的共表达模式,

可能与其合成相关. 进一步利用分子对接对这两个CYP蛋白与吲哚分子间的催化特性进行预测分析, 结果显示*CYP735A2*可与吲哚分子以疏水键和氢键结合, 结合能为 -5.8 kcal/mol, 活性口袋由10个氨基酸残基构成, 其中与Lys490形成氢键(图6A和B); *CYP2A6-1*可通过形成由9个氨基酸残基组成的活性口袋与吲哚分子结合, 作用力为疏水作用, 结合能 -5.5 kcal/mol(图6C和D). 以上预测结果说明*CYP735A2*与*CYP2A6-1*均极有可能对吲哚分子具有催化活性.

图4 菰蓝叶中吲哚合成途径的共表达转录分析

A: 具有显著相关性的基因间的共表达网络; B: 相关系数 >0.98 的基因间的共表达网络, 蓝色圆点表示途径基因, 灰色圆点表示转录因子; C: 吲哚合成途径与CYP家族基因的共表达分析, 箭头标明的基因集既包含了具有共表达模式的吲哚途径基因和CYP基因; D: 具有显著相关性的吲哚合成途径与CYP基因的共表达网络, 蓝色圆点表示吲哚途径基因, 橙色圆点表示转录因子。共表达网络中虚线表示两个基因负相关, 实线为正相关; 线的粗度表示相关系数值的大小

3 讨论

色氨酸代谢途径在植物中普遍存在, 而靛蓝和靛玉红等化合物却仅在少数植物中合成^[15], 探明这些具有物种特异性的合成和调控机制对了解植物次生代谢, 探明物种演化均具有重要的理论意义。同时, 菰蓝即具药用价值又可作为染料原料, 其市场需求依旧十分巨大。因此探明大青叶中吲哚类活性成分的生物合成机制对于提升大青叶的应用价值、找寻替代资源等方面都具有重要的价值。本研究初步探讨了大青叶中吲哚合成途径的合成特征, 预测了相关的催化酶基因和调控因子, 为完全阐明这一合成途径提供了

理论依据。

3.1 菰蓝叶片中吲哚途径的特征

依据注释, 参考KEGG途径信息, 对菰蓝叶片中的吲哚途径进行了基因注释。以吲哚为分支点, IAA, TIA, IG合成途径的催化酶基因也被注释到, 说明在菰蓝叶片中存在这些途径, 这一结果也与菰蓝中化合物的分布特征相吻合^[6,9-11]。同时也发现菰蓝叶片的吲哚途径存在特殊性, 例如, 在微生物的色氨酸途径中, 色氨酸可在色氨酸酶(tryptophanase)的催化下逆向生成吲哚^[26], 而菰蓝中则没有注释到该基因, 说明这个催化步骤在菰蓝或植物中不存在, 或需要特定的条件才

图 5 CYP 基因的表达特征

A: CYP 基因在菘蓝根、叶中的转录特征; B 和 C: CYP735A2 与 CYP2A6-1 响应 MeJA 诱导的转录变化。误差线表示 3 个生物重复的标准偏差, 差异显著性分析由 *t*-test 方差分析显示与, *: $P < 0.05$, 具有显著差异; **: $P < 0.01$, 差异极显著

会发生。同时色氨酸酶这个外源蛋白提供了一个理想的代谢调控元件来有效控制整个色氨酸途径的代谢流。

3.2 吲哚活性成分特异催化酶的预测

本研究通过两种策略来对未知催化途径进行预测。利用共表达网络, 发现 CYP735A2 最有可能与吲哚合成相关, 其拟南芥 (*Arabidopsis thaliana*) 中的同源基因可催化激动素的羟基化反应^[27], 吲哚和激动素的化合物骨架就有一定相似性(均为六元并五元杂环), 说明 CYP735A2 极有可能对吲哚分子具有结合催化能力。另一方面, 一些外源的单加氧酶或双加氧酶被发现可催化吲哚的羟基化反应, 包括细胞色素 P450^[13], 萘双加氧酶 (naphthalene 1,2-dioxygenase)^[23] 和酚羟化酶 (phenol hydroxylase)^[24]。菘蓝中没有注释到萘双加氧酶和酚羟化酶, 而多条菘蓝 CYP 基因 (comp33180_c0_seq5, comp33180_c0_seq6, comp33180_c0_seq1, comp33180_c0_seq2) 与 CYP2A6 具有显著同源性。以上信息为后续开展功能筛选有效缩小了候选基因的范围。

此外, 外源 CYP2A6 可以在烟草^[13] 和微生物中^[28,29] 催化生成吲哚酚, 提示 CYP2A6 可以直接作为菘蓝中吲哚活性成分代谢工程的遗传操作靶基因。

3.3 菘蓝吲哚活性成分的代谢调控策略

通过本研究发现了一系列可能与菘蓝吲哚活性成分合成相关的基因。根据这些基因在途径中的作用可以制定不同的代谢调控策略。例如, TSA2 与 TSB3 处于色氨酸合成的上游途径, 并且预测与靛蓝、靛玉红等成分合成相关, 对其进行基因操作, 可以有效改变目标支路的代谢水平, 也可能引起整个色氨酸途径网络的代谢流变化; 预测 CYP 基因对靛蓝、靛玉红可能具有更特异性的影响; bHLH-215 作为转录因子, 也可以对整个代谢途径产生调控; 此外, 除了这些内源基因, 利用对吲哚有催化活性的异源基因也可以作为菘蓝代谢工程的操作基因, CYP2A6 在烟草中催生成吲哚的研究就是很好的范例^[13]。

值得注意的是, 不同于其他植物次生代谢途径, 吲哚途径与植物激素吲哚乙酸 (IAA) 同属于色氨酸代谢

图6 分子对接预测候选CYP对吲哚分子的催化活性

A: CYP735A2蛋白与吲哚分子的结合预测模型; B: CYP735A2与吲哚分子结合的活性口袋模型; C: CYP2A6-1蛋白与吲哚分子的结合预测模型; D: CYP2A6-1与吲哚分子结合的活性口袋模型。绿色氨基酸残基代表具有疏水作用的残基, 粉色氨基酸残基代表具有酸性(红色边框)和碱性作用(蓝色边框)。蓝色标注的原子朝向口袋外侧

网络, 代谢流的改变, 可能引起IAA合成的变化^[30,31], 而IAA作为重要的植物激素, 其含量水平的变化则可引起植物生长发育的失常。因此, 在关注目标化合物合

成水平的同时, 菰蓝吲哚途径的代谢工程还需注意在代谢流调控和植物正常的生理发育寻找一个平衡点, 才能在产量稳定的前提下实现大青叶品质的提升。

参考文献

- 1 乔传卓. 菰蓝和欧洲菰蓝的鉴别研究. 第二军医大学学报, 1985, 22: 237-242
- 2 Qin H Z, Shi B, Li S Y, et al. Comparing research in Antiviral action of *Rhizoma et Radix Baphicacanthis Cusiae* and *Radix Isatidis* against a virus. *Chin Archiv Trad Chin Med*, 2009, 27: 168-169
- 3 李彬. 板蓝根活性成分及品质评价. 博士学位论文. 上海: 第二军医大学, 2003
- 4 Li J, Zhou B, Li C, et al. Lariciresinol-4-*O*- β -D-glucopyranoside from the root of *Isatis indigotica* inhibits influenza A virus-induced pro-inflammatory response. *J Ethnopharmacol*, 2015, 174: 379-386
- 5 Yang Z, Wang Y, Zheng Z, et al. Antiviral activity of *Isatis indigotica* root-derived clemastanin B against human and avian influenza A and B viruses *in vitro*. *Int J Mol Med*, 2013, 31: 867-873
- 6 Chen J, Dong X, Li Q, et al. Biosynthesis of the active compounds of *Isatis indigotica* based on transcriptome sequencing and metabolites profiling. *BMC Genomics*, 2013, 14: 857
- 7 Zhang L, Chen J, Zhou X, et al. Dynamic metabolic and transcriptomic profiling of methyl jasmonate-treated hairy roots reveals synthetic characters and regulators of lignan biosynthesis in *Isatis indigotica* Fort. *Plant Biotechnol J*, 2016, 14: 2217-2227
- 8 Xiao Y, Ji Q, Gao S, et al. Combined transcriptome and metabolite profiling reveals that *li* PLR1 plays an important role in lariciresinol

- accumulation in *Isatis indigotica*. *J Exp Bot*, 2015, 66: 6259–6271
- 9 孙东东, 严世海, 陈建伟, 等. 板蓝根有效组分的抗病毒活性研究. *南京中医药大学学报*, 2013, 29: 53–55
 - 10 Mohn T, Plitzko I, Hamburger M. A comprehensive metabolite profiling of *Isatis tinctoria* leaf extracts. *Phytochemistry*, 2008, 70: 924–934
 - 11 Liu Y F, Chen M H, Lin S, et al. Indole alkaloid glucosides from the roots of *Isatis indigotica*. *J Asian Nat Products Res*, 2016, 18: 1–12
 - 12 Wu X, Chen X, Jia D, et al. Corrigendum: characterization of anti-leukemia components from *Indigo naturalis* using comprehensive two-dimensional K562/cell membrane chromatography and *in silico* target identification. *Sci Rep*, 2016, 6: 30103
 - 13 Warzecha H, Frank A, Peer M, et al. Formation of the indigo precursor indican in genetically engineered tobacco plants and cell cultures. *Plant Biotech J*, 2007, 5: 185–191
 - 14 王翠玲, 刘建利, 张楠楠, 等. 色胺酮在菘蓝中的分布、生物合成及仿生合成研究. 2006中国药学会学术年会, 2006
 - 15 Oberthür C, Hamburger M. Tryptanthrin content in *Isatis tinctoria* leaves—a comparative study of selected strains and post-harvest treatments. *Planta Medica*, 2004, 70: 642
 - 16 Xia Z Q, Zenk M H. Biosynthesis of indigo precursors in higher plants. *Phytochemistry*, 1992, 31: 2695–2697
 - 17 Frerigmann H, Gigolashvili T. MYB34, MYB51, and MYB122 distinctly regulate indolic glucosinolate biosynthesis in *Arabidopsis thaliana*. *Mol Plant*, 2014, 7: 814–828
 - 18 董娟娥, 梁宗锁, 尉芹, 等. 不同区域松蓝根、叶(板蓝根、大青叶)有效成分含量差异. *应用生态学报*, 2006, 17: 1613–1618
 - 19 孙立新, 唐虹. RP-HPLC法测定板蓝根、大青叶中靛蓝、靛玉红的含量. *沈阳药科大学学报*, 2000, 17: 191–193
 - 20 Radwanski E R, Zhao J, Last R L. *Arabidopsis thaliana* tryptophan synthase alpha: gene cloning, expression, and subunit interaction. *Mol Gen Genet*, 1995, 248: 657–667
 - 21 Zhang R, Wang B, Ouyang J, et al. *Arabidopsis* indole synthase, a homolog of tryptophan synthase alpha, is an enzyme involved in the trp-independent indole-containing metabolite biosynthesis. *J Integr Plant Biol*, 2008, 50: 1070–1077
 - 22 Schneider T R, Gerhardt E, Lee M, et al. Loop closure and intersubunit communication in tryptophan synthase. *Biochemistry*, 1998, 37: 5394–5406
 - 23 Carredano E, Karlsson A, Kauppi B, et al. Substrate binding site of naphthalene 1,2-dioxygenase: functional implications of indole binding. 1 Edited by D. C. Rees. *J Mol Biol*, 2000, 296: 701–712
 - 24 Shi S, Ma F, Sun T, et al. Biotransformation of indole to indigo by the whole cells of phenol hydroxylase engineered strain in biphasic systems. *Appl Biochem Biotechnol*, 2013, 169: 1088–1097
 - 25 Yan X, Chen S. Regulation of plant glucosinolate metabolism. *Planta*, 2007, 226: 1343–1352
 - 26 Meister A. Tryptophanase: Structure, Catalytic Activities, and Mechanism of Action Advances in Enzymology and Related Areas of Molecular Biology, Volume 42. Hoboken: John Wiley & Sons, Inc., 2006. 287
 - 27 Takei K, Yamaya T, Sakakibara H. *Arabidopsis CYP735A1* and *CYP735A2* encode cytokinin hydroxylases that catalyze the biosynthesis of *trans*-Zeatin. *J Biol Chem*, 2004, 279: 41866–41872
 - 28 Park S H, Kim D H, Kim D, et al. Engineering bacterial cytochrome P450 (P450) BM3 into a prototype with human P450 enzyme activity using indigo formation. *Drug Metab Dispos*, 2010, 88: 732–739
 - 29 Wu Z L, Podust L M, Guengerich F P. Expansion of substrate specificity of cytochrome P450 2A6 by random and site-directed mutagenesis. *J Biol Chem*, 2005, 280: 41090–41100
 - 30 Kim J I, Dolan W L, Anderson N A, et al. Indole glucosinolate biosynthesis limits phenylpropanoid accumulation in *Arabidopsis thaliana*. *Plant Cell*, 2015, 27: 1529–1546
 - 31 Pfalz M, Mikkelsen M D, Bednarek P, et al. Metabolic engineering in *Nicotiana benthamiana* reveals key enzyme functions in *Arabidopsis* indole glucosinolate modification. *Plant Cell*, 2011, 23: 716–729

Comprehensive transcriptomic profiling of *Isatis indigotica* leaves reveals biosynthesis of indole active ingredients

CHEN JunFeng¹, LI Qing¹, WANG Yun¹, ZHANG QingLei², ZHANG Lei² & CHEN WanSheng¹

¹ Department of Pharmacy, Changzheng Hospital, Second Military Medical University, Shanghai 200003, China;

² Department of Pharmacognosy, School of Pharmacy, Second Military Medical University, Shanghai 200433, China

Dried *Isatis indigotica* (Daqingye) leaves are widely used in traditional Chinese medicine. Indole derivatives, including indigo, indirubin, and tryptanthrin, are major active ingredients of Daqingye. In plants, the biosynthesis of these indole derivatives is a branch of the tryptophan pathway, which has not been completely elucidated to date. In this study, the transcriptional profiling of methyl jasmonate-treated *I. indigotica* leaves was performed to characterize this pathway and predict its unknown catalytic steps. According to transcriptome annotation, 42 unigenes encoding 16 catalytic enzymes involved in biosyntheses of tryptophan, indole acetic acid, indole glycoside, and terpenoid indole alkaloids were identified. A co-expression network was constructed, which indicated bHLH125 as a core regulator of the biosynthetic pathway. Furthermore, CYP2A6-1 and CYP735A2 were predicted to catalyze the precursors of indigo and indirubin, which showed good binding affinity for indole by molecular docking. This study provides an insight into indole biosynthesis and metabolism, providing potential elements for metabolic engineering of indoles.

***Isatis indigotica*, indole derivatives, transcriptomic profiling, co-expression network**

doi: [10.1360/N052017-00195](https://doi.org/10.1360/N052017-00195)