

杨博涵,徐杰,汤丽云,等. 阳春砂花丝、花柱转录组测序及生物信息学分析[J]. 江苏农业科学,2022,50(22):38-45.

doi:10.15889/j.issn.1002-1302.2022.22.005

阳春砂花丝、花柱转录组测序及生物信息学分析

杨博涵¹, 徐杰^{1,2}, 汤丽云³, 苏景⁴, 李子翔¹, 何国振¹

[1. 广州中医药大学中药学院, 广东广州 510006; 2. 广东一方制药有限公司/广东省中药配方颗粒企业重点实验室, 广东佛山 528244;

3. 华南农业大学生命科学学院, 广东广州 510642; 4. 阳春市农业试验场(阳春市春砂仁试验场), 广东阳春 529600]

摘要:为进一步探究阳春砂假合蕊柱形成机制,基于 RNA-Seq 技术对不同生长时期阳春砂花丝、花柱进行转录组测序及生物信息学分析,获得阳春砂转录组数据并筛选出激素合成及信号转导通路基因,为后续研究阳春砂基因功能、代谢途径、花器官发育及运动调控机理等方面提供参考。结果表明,测序数据拼接组装后获得 94 584 条 Unigene,总长度为 92 501 015 bp。将获得的 Unigene 分别在 NR、GO、KOG 和 KEGG 等七大数据库中进行比对,共有 62 174 条 Unigene 得到注释,占全部 Unigene 的 65.73%;NR 数据库中注释到 58 669 条 Unigene,被注释的同源序列主要来自小果野芭蕉;有 45 892 条 Unigene 注释到 KOG 数据库中,涉及 25 个功能分类;GO 数据库注释 Unigene 12 050 条,按照功能分为 3 个大类及 55 个亚类;KEGG 数据库中,有 44 000 条 Unigene 注释到 137 个代谢通路中;SSR 特征分析中共检测到 18 895 个位点,三碱基重复的数目最多,达到 6 313 个,占比为 33.41%。2012 条 Unigene 被注释到 9 种不同激素的合成与信号转导途径中,赤霉素的 Unigene 被注释到最多,共有 505 条。本研究通过构建阳春砂花丝、花柱的转录组数据库并进行生物信息学分析,为阳春砂花器官发育相关基因的发掘及利用、基因组的测序与组装、表达谱的比对提供科学依据,也为后续在分子生物学层面对阳春砂假合蕊柱的形成机制开展深入研究提供理论基础。

关键词: 阳春砂(*Amomum villosum*); 转录组; 花丝; 花柱; 激素

中图分类号: S567.23*9.01 **文献标志码:** A **文章编号:** 1002-1302(2022)22-0038-08

阳春砂(*Amomum villosum*)是姜科(Zingiberaceae)豆蔻属(*Amomum*)多年生常绿草本植物^[1],其干燥成熟果实为四大南药之首——砂仁的主流品种,具有理气安胎、温脾止泻、化湿开胃的功效^[2]。作为传统的大宗药材,阳春砂每年的需求数量大于 300 万 kg,具有重要的药用价值及经济价值^[3]。但阳春砂在生产上存在着严重的低产问题,其自然结实率仅有 1.1%^[4],研究者对此进行了许多研究,认为较低的自然结实率与阳春砂小花的特殊花器结构有着密切的联系。何国振等将阳春砂特殊的花器结构称为假合蕊柱,即雌雄蕊贴合一

起,但并未完全合生,在处于开花的状态下,唇瓣包裹着假合蕊柱,并且呈现出半抱合的状态,与唇瓣间的距离极近,仅有 1.3~2.0 mm,严重阻碍了相应虫媒的授粉^[3]。为提高产量,农户在实际种植中主要采用人工授粉的方式进行劳作,而人工授粉成本高、劳作强度大、根状茎及花序易被严重踩踏损伤等原因是传统人工授粉方式固有的弊端,严重抑制了农户的生产积极性。

细胞伸长或者细胞分裂所致的植物器官运动与内源激素及相关抑制剂作用相关,而细胞结构及生长速率的不对称是花器官运动的基础。有文献报道,腊梅(*Chimonanthus praecox*)花丝的运动来源于其两侧表面细胞生长速率的差异^[5]。姜科植物马来良姜(*Alpinia mutica*)花柱运动部位两侧细胞层数的差异是其花柱卷曲运动的基础^[6],赤霉素(GA)、茉莉酸(JA)、吲哚-3-乙酸(IAA)等激素则有着调控植物雌雄蕊发育的作用^[7-9]。何卓航等研究发现,随着阳春砂小花的生长,其花丝、花柱远、近轴侧的细胞层数出现了差异,这种不对称的结构是雌雄蕊运动的基础,认为阳春砂雌雄蕊的相向运动致使其假合蕊柱的形成^[10];同时对花柱的近

收稿日期:2021-12-12

基金项目:广东省重点领域研发计划“现代种业”重大专项(编号:2020B020221002);广东省林业科技创新项目(编号:2018KJXC033);广东省农业科技创新体系管理项目(编号:0835-190Z22404211);广东省农业科技创新及推广体系建设项目(编号:2021KJ142_2020KJ142_2019KJ142)。

作者简介:杨博涵(1997—),男,河南洛阳人,硕士研究生,从事药用植物资源利用与生物技术研究。E-mail:13727858519@163.com。

通信作者:何国振,博士,教授,从事药用植物生理与分子生物学研究。E-mail:heguozhen@gzuem.edu.cn。

端以及远端轴侧的 IAA 水平进行测定,结果表明,生长期在保持同样的水平不同横切部位中,IAA 水平基本上呈现远端轴侧小于近端轴侧的趋势,故花柱运动的原因可能是来自于花柱近、远轴侧 IAA 水平的差异,该研究从生理层面初步阐述了阳春砂假合蕊柱的形成机制。然而,若要进一步探明阳春砂假合蕊柱形成机制,则须对阳春砂花器官运动及发育的分子机制作深入研究。

因此,本研究对阳春砂不同生长时期的花丝、花柱进行转录组测序,通过数据拼接、组装的方式建立阳春砂花丝、花柱的转录组数据库,将所得的 Unigene 进行功能注释、分类以及简单重复序列 (SSR) 分子标记。同时,重点关注并筛选可能影响阳春砂雌蕊、雄蕊运动的激素合成与信号转导途径关键基因。以期进一步解析阳春砂假合蕊柱产生的调控机制及分子机制,并为人为干预阳春砂的花器官结构,提高产量奠定理论基础。

1 材料与方法

1.1 试验材料

材料为大田栽培的不同生长时期阳春砂小花的花柱及花丝,阳春砂植株栽培于广东省阳春市合水镇那软村阳春砂种植基地 (22°17'N, 112°01'E), 于 2017 年 5—7 月在该基地开展大田试验。参照陈红的方法^[11],按照长度的不同,将阳春砂小花的生长时期进行了划分(表 1),并于阳春砂花期内,每天上午采摘小花,解剖并分离出不同时期小花的花丝与花柱,迅速置于液氮灌中,随后转移至超低温冰箱 (-80 °C) 进行保存。

表 1 阳春砂小花生长时期划分

时期	小花长度 (cm)
1	[1.0,1.5)
2	[1.5,2.0)
3	[2.0,2.5)
4	[2.5,3.0)
5	[3.0,开花)
6	开花当天
7	花后 1 d

1.2 试验方法

1.2.1 RNA 提取 分别提取阳春砂小花在各生长时期的花柱和与花丝的总 RNA,重复 3 次。利用 NanoDrop 2000 和 Agilent 2100 Bioanalyzer 测定 RNA

的浓度、纯度与完整性,数据检测合格后的 RNA 用于构建转录组数据库。此部分试验与转录组的测序委托深圳华大基因科技服务有限公司完成。

1.2.2 cDNA 文库构建及转录组序列组装 使用 Oligo(dT) 磁珠富集质检合格且所有样品的混合 RNA。加入打断试剂,以片段化的 mRNA 为模板合成 1 链、2 链 cDNA,配制反应体系,使接头与 cDNA 连接。PCR 反应及产物回收、扩增。PCR 产物变性,充分混匀,得到单链环产物,随后 PCR 产物变性,即得到文库。利用 Agilent 2100 Bioanalyzer 和 ABI StepOnePlus real-time PCR System 对所得文库进行检测,检测合格后进行转录组 *de novo* 测序。所得 reads 通过 Trinity 软件进行序列组装,组装序列的质量通过 BUSCO 软件进行评估。

1.2.3 转录组基因功能注释及数据挖掘 利用生物信息学方法分析获得的阳春砂 Unigene,为获得全方位的基因功能信息,对组装所得的 Unigene 在七大功能数据库中进行注释,包括 NR、NT、KOG/COG、GO、KEGG、SwissPro 及 Interpro。使用 MlMicroSatellite (MISA) 工具对 Unigene 进行 SSR 位点的挖掘。并在 KEGG 代谢通路中重点关注植物激素生物合成与信号转导 Unigene 的注释情况。

2 结果与分析

2.1 转录组数据组装

使用 Illumina HiSeq-2000 平台对不同时期的阳春砂花丝及花柱的 cDNA 进行测序。测序的结果显示,从转录组中获得了 81.40 Mb 的原始读数,用过滤软件 SOAPnuke 去除低质量的 reads 后共得到 68.48 Mb Clean reads,最终获得 10.33 Gb 的碱基总数。Q20 和 Q30 的百分比分别为 97.20% 和 93.16% (表 2)。说明转录组的测序质量较高,可以满足后续的生物信息学分析。

表 2 测序数据质量分析

原始序列 (M)	有效序列 (M)	有效碱基数 (Gb)	Q20 (%)	Q30 (%)	有效序列占比 (%)
81.40	68.84	10.33	97.20	93.16	84.57

注:Q20、Q30 为质量值≥20 或 30 的碱基所占百分比。

对转录本中 Clean reads 进行组装,一共获得 138 590 个转录本,包含了 111 312 938 个核苷酸的序列信息,这些片段长度的平均值为 803 bp, N50 为 1 414 bp, GC 含量为 44.28%。对转录本进一步聚

类并去除冗余的序列之后得到 94 584 条 Unigene, 总长度为 92 501 015 bp。Unigene 的 N50、N70 和

GC 含量分别为 1 582bp、1 002 bp 和 44.35%, 均大于其对应平均长度, 说明组装的结果良好(表 3)。

表 3 转录本和单基因簇统计分析

类别	序列总数(条)	序列总长度(bp)	序列平均长度(bp)	N50(bp)	N70(bp)	GC含量(%)
转录本	138 590	111 312 938	803	1 414	870	44.28
单基因簇	94 584	92 501 015	977	1 582	1 002	44.35

2.2 阳春砂转录组基因总体注释情况

将所获得的结果在七大功能数据库进行注释, 结果见表 4。结果显示, 阳春砂共注释到 94 584 条 Unigene。其中, 共有 58 669 条 Unigene 被 NR 数据库注释, 占比最多, 达到总 Unigene 的 62.03%; NT 数据库有 40 205 条, 占 42.51%; SwissProt 有 40 318 条, 占总 Unigene 的 42.63%; KEGG 数据库与 KOG 数据库分别有 44 000 条和 45 892 条 Unigene, 各占 46.52% 与 48.52%; Interpro 有 45 793 条, 占 48.42%。GO 数据库注释到的基因最少, 仅有 12 050 条, 占总数的 12.74%; 所得比对结果显示, 在七大数据库中均能成功注释的 Unigene 共有 5 691 条, 占总 Unigene 条数的 6.02%。

表 4 转录组基因注释情况统计

数据库名称	注释数量(条)	所占百分比(%)
NR	58 669	62.03
NT	40 205	42.51
SwissProt	40 318	42.63
KEGG	44 000	46.52
KOG	45 892	48.52
Interpro	45 793	48.42
GO	12 050	12.74
Intersection	5 691	6.02
总数	94 584	100

注: Intersection 表示上述 7 个数据库中均能成功注释的 Unigene 数量占 Unigene 总数的比例。

2.2.1 NR 数据库功能注释 NR 数据库的注释结果(图 1)显示, 匹配最多的物种为小果野芭蕉(*Musa acuminata* subsp. *Malaccensis*), 该物种注释到的基因数量最多, 共有 45 257 条, 占比高达 77.14%, 证明阳春砂与该物种的同源性较高; 其他物种依次为油棕(*Elaeis guineensis*)与海藻(*Phoenix dactylifera*), 分别有 3 338、2 435 条 Unigene 被注释, 占比分别为 5.69%、4.15%, 阳春砂与这 2 种植物

的同源性相对较低。而剩余的 13.02% 则分布于其他物种中。

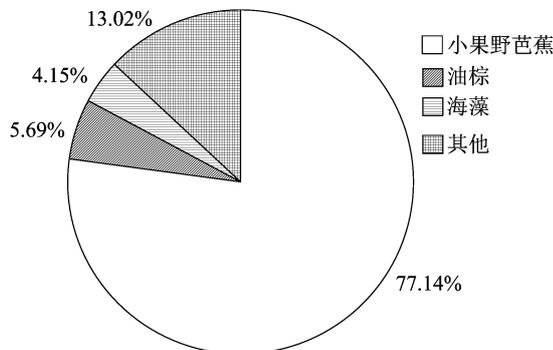
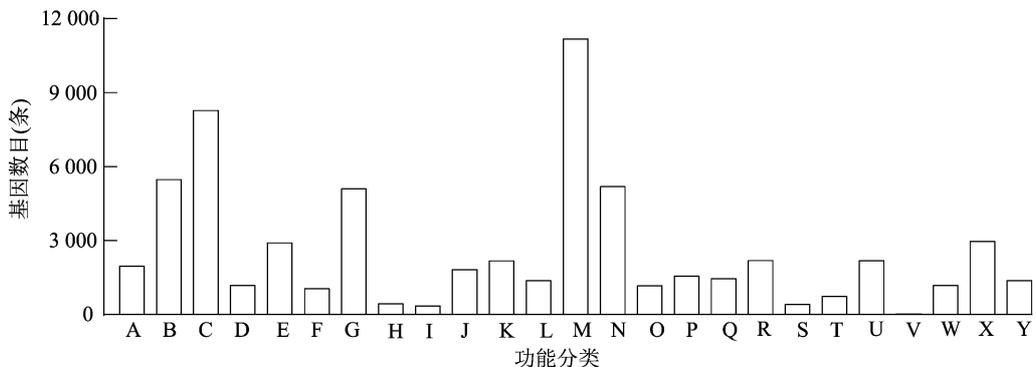


图 1 NR 数据库物种分布统计

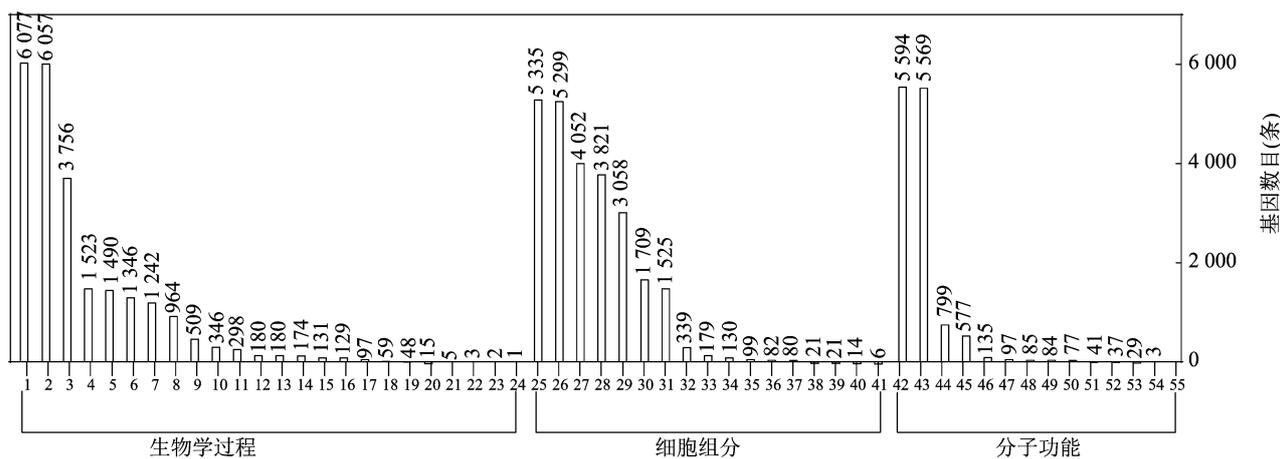
2.2.2 KOG 数据库功能注释 在 KOG 数据库中, 共有 45 892 条 Unigene 被注释到(图 2), 共分为 25 个大类功能区, 包括一般功能预测、信号转导机制、转录等功能。在为数众多且不同的功能分类中, 注释到的基因的数量差异较为显著, 一般功能预测类基因数量最多, 共有 11 177 条 Unigene, 占比为 24.36%, 其次是信号转导机制, 有 8 271 条 Unigene 被注释, 占比为 18.02%; 除此之外, 负责转录功能的 Unigene 有 5 470 个, 占比为 11.92%; 有 1 172 条 Unigene 注释到负责次生代谢产物生物合成、运输和代谢功能区中, 占比为 2.55%。

2.2.3 GO 数据库功能注释 GO 数据库中的功能分类注释结果见图 3, 结果表明, 阳春砂花丝及花柱中共有 12 050 条 Unigene 注释到不同的功能节点上, 共涉及到生物学过程、细胞组分和分子功能 3 个大类, 55 个亚类。归入到细胞组分的 17 个亚类中, 以细胞、细胞部分、膜和细胞器功能的 Unigene 数量最多, 分别有 5 335、5 229、4 052、3 821 条。在分子功能的 14 个亚类中, 催化活性和结合注释数量最多, 分别有 5 594、5 569 条。而涉及生物学过程的 24 个亚类中, 以代谢过程(6 077 条)和细胞进程(6 057 条)为主。结果表明, 同一个 Unigene 可以注释到多个功能节点上, 因此同一个功能分支的总注



A—翻译、生物起源及核糖体结构；B—转录；C—信号转导机制；D—次生代谢产物的生物合成、分解代谢与运输；E—RNA加工和修饰；F—复制、重组及修复；G—翻译后的修饰、蛋白质的转换与代谢；H—核苷酸的转运及代谢；I—核结构；J—脂质运输和代谢；K—细胞内的运输、分泌及囊泡运输；L—无机离子的转运和代谢；M—一般功能预测；N—未知函数；O—细胞外结构；P—能量产生及转化；Q—防御机制；R—细胞骨架；S—辅酶的转运及代谢；T—染色质的结构及动力学；U—细胞壁/细胞膜/包膜的生物发生；V—细胞的运动；W—细胞周期控制、细胞分裂及染色体分裂；X—碳水化合物的代谢及运输；Y—氨基酸转运与代谢

图2 转录组 Unigene 的 KOG 功能分布统计



1—代谢过程；2—细胞过程；3—单有机体过程；4—生物调节；5—本地化；6—生物的调节过程；7—应激反应；8—细胞组分或生物发生；9—信号；10—发展过程；11—多细胞生物的生长过程；12—再生作用；13—生殖进程；14—生物负调控进程；15—多细胞生物进程；16—生物正调控进程；17—解毒；18—免疫过程；19—生长；20—有节奏的过程；21—细胞死亡；22—生物黏附；23—运动；24—生物相；25—细胞；26—细胞部分；27—膜；28—细胞器；29—膜部分；30—细胞器部分；31—大分子复合物；32—膜封闭腔；33—细胞骨架；34—胞外区；35—超分子复合体；36—细胞连接；37—共聚体；38—病毒体；39—病毒部分；40—胞外区部分；41—核苷酸；42—催化活性；43—结合；44—翻译调节活性；45—结构分子的活性；46—核酸转录因子的活性；47—抗氧化活性；48—分子功能调节剂；49—电子载体活性；50—信号传感器的活性；51—转录因子活性，蛋白结合；52—分子转导活性；53—营养库活性；54—蛋白标识；55—金属伴侣活性

图3 阳春砂 Unigene 的 GO 数据库功能分类

释数大于注释到该功能分支的总 Unigene 数。

2.2.4 KEGG 数据库功能注释 对转录组测序获得的 Unigene 进行 KEGG 代谢通路富集分析后发现,共有 44 000 条 Unigene 得到注释,涉及到 137 个通路。对富集显著的前 20 条通路进行分析(表 5),其中,共有 8 660 条 Unigene 被代谢途径通路所注释,占比最多,达到注释总数的 19.68%。而次生代谢产物的生物合成通路共有 4 191 条 Unigene 被注释到,占注释总数的 9.53%,另外,本研究主要关注的植物激素转导通路一共有 1 656 条 Unigene 被注释到,占比为 3.76%。

在所有的代谢通路中,共有 23 个与次生代谢相关的代谢通路,对其所有的 Pathway 进行分析后(表 6)发现,一共有 4 191 个 Unigene 被次生代谢产物的生物合成途径所注释;而在本研究主要关注的植物激素合成途径中,62 个 Unigene 被油菜素内酯生物合成途径所注释,另有 59 个 Unigene 被玉米素生物合成途径所注释。

2.2.5 SSR 特征分析 转录组测序所得 Unigene 的 SSR 的检测最终结果显示,共有 18 895 个 SSR 被检测到(图 4)。总共有 14 种多碱基重复 SSR,其中,三核苷酸重复的 SSR 数目最多,共有 6 313 个,

表5 次生代谢相关代谢通路的注释情况

代谢通路编号	代谢通路	Unigene 数目 (条)
ko01100	代谢途径	8660 (19.68%)
ko01110	次生代谢产物的生物合成	4191 (9.53%)
ko03013	RNA 转运	1668 (3.79%)
ko04075	植物激素的信号转导	1656 (3.76%)
ko04626	植物与病原物的相互作用	1354 (3.08%)
ko04144	胞吞作用	1198 (2.72%)
ko04016	MAPK 信号通路	1191 (2.71%)
ko03015	mRNA 监视通路	1167 (2.65%)
ko04141	内质网蛋白的加工	1098 (2.50%)
ko03040	剪接体	1032 (2.35%)
ko03010	核糖体	1003 (2.28%)
ko01200	碳代谢	992 (2.25%)
ko00230	嘌呤代谢	965 (2.19%)
ko00240	嘧啶代谢	953 (2.17%)
ko01230	氨基酸的生物合成	930 (2.11%)
ko00500	淀粉和蔗糖的代谢	892 (2.03%)
ko00940	苯丙素的生物合成	793 (1.80%)
ko03018	RNA 降解	747 (1.7%)
ko04120	泛素介导蛋白质水解	726 (1.65%)
ko00520	氨基糖和核苷酸糖代谢	618 (1.40%)

表6 次生代谢相关代谢通路的注释情况

代谢通路编号	代谢通路	Unigene 数目 (条)
ko00110	次生代谢产物的生物合成	4 191
ko00940	苯丙素的生物合成	793
ko00900	萜类骨架的生物合成	266
ko00860	卟啉和叶绿素代谢	205
ko00130	辅酶 Q 和其他类萜合成	165
ko00906	类胡萝卜素的生物合成	160
ko00941	黄酮类的生物合成	123
ko00901	吲哚生物碱的生物合成	94
ko00073	角质、小檗碱和蜡的生物合成	89
ko00950	异喹啉生物碱的生物合成	81
ko00960	托烷、哌啶和吡啶生物碱的生物合成	80
ko00904	二萜的生物合成	76
ko00909	倍半萜类的生物合成	68
ko00905	油菜素内酯的生物合成	62
ko00908	玉米素生物合成	59
ko00903	柠檬烯、蒽烯降解	52
ko00261	单环 β -内酰胺的合成	41
ko00944	黄酮和黄酮醇的生物合成	31
ko00943	异黄酮的生物合成	24
ko00942	花青素的生物合成	13
ko00902	单萜的生物合成	10
ko00402	苯并噁类的生物合成	9
ko00965	甜菜红色素的生物合成	4

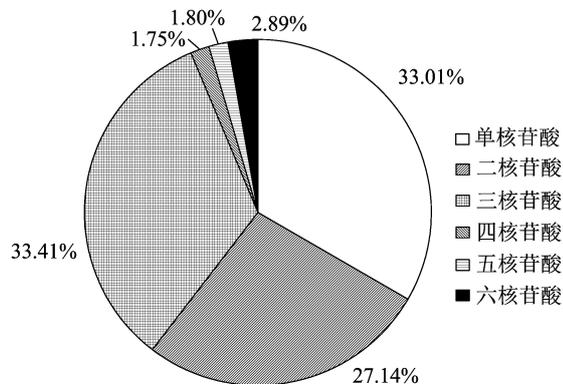


图4 阳春砂 SSR 统计分析

占 SSR 位点总数的 33.41%；三核苷酸重复中 AGG/CCT 重复单元最多,共有 1 434 个;单核苷酸和二核苷酸重复分别有 6 238 个 (33.01%) 和 5 128 个 (27.14%),二核苷酸中重复单元最多的是 AG/CT,共有 2 909 个;六核苷酸重复数目共有 546 个 (2.89%),四核苷酸重复 330 个 (1.75%) 和五核苷酸重复 340 个 (1.80%)。

2.3 重点关注通路的 Unigene 注释情况

2.3.1 植物激素信号转导途径中 Unigene 注释情况

在 KEGG 代谢通路中,共有 1 760 条 Unigene 注释到植物激素的信号转导途径中(表 7)。赤霉素的信号转导途径中注释到的 5 个基因,对应的 Unigene 最多,共有 438 个;其中,植物光敏色素互作因子 (*PIF*) 对应的 Unigene 最多,达到 318 个;而水杨酸信号转导途径注释到 3 个基因,包括非表达病程相关蛋白 (*NPRI*)、*TGA* 转录因子及病程相关蛋白 1 (*PRI*) 基因,但对应的 Unigene 最少,仅有 74 个;茉莉酸的信号转导途径中注释到的茉莉酸 ZIM 结构域蛋白 (*JAZ*)、茉莉酸氨基酸合成酶 (*JARI*)、*MYC2* 转录因子和茉莉酸受体蛋白 (*COI1*) 一共对应了 176 条 Unigene;乙烯和油菜素内酯信号转导途径注释到的基因最多,均为 8 个,分别注释到 139 条及 191 条 Unigene;其他植物激素中,生长素、细胞分裂素以及脱落酸的信号转导途径分别注释到 329、245、168 条 Unigene;统计结果中,各激素信号转导途径关键基因均有注释,如细胞分裂素中的磷酸转移蛋白 (*AHP*)、生长素中的生长素载体 (*AUX1*)、生长素/吲哚乙酸蛋白 (*Aux/IAA*) 转录家族以及脱落酸信号转导途径中的 *ABF* 转录因子等。综上,研究较为深入的植物激素基本均在 KEGG 通路中有所注释,且其信号转导途径的关键基因都有一定数量的 Unigene 被注释到。

表 7 转录组中植物激素信号转导途径 Unigene 注释情况

植物激素	基因名称	Unigene 数目(条)	代谢通路编号	
赤霉素	<i>GID1</i>	19	K14493	
	<i>GID2</i>	15	K14495	
	<i>DELLA</i>	86	K14494	
	<i>PIF4</i>	136	K16189	
	<i>PIF3</i>	182	K12126	
水杨酸	<i>NPR1</i>	25	K14508	
	<i>TGA</i>	41	K14431	
	<i>PR1</i>	8	K13449	
乙烯	<i>ETR</i>	14	K14509	
	<i>CTR1</i>	41	K14510	
	<i>SIMKK</i>	7	K13413	
	<i>MPK6</i>	7	K14512	
	<i>EIN2</i>	18	K14513	
	<i>EBF1/2</i>	15	K14515	
	<i>EIN3</i>	26	K14514	
	<i>ERF1/2</i>	11	K14516	
	油菜素内酯	<i>BAK1</i>	11	K13416
		<i>BRI1</i>	111	K13415
<i>BKII</i>		5	K14449	
<i>BSK</i>		21	K14500	
<i>BIN2</i>		11	K14502	
<i>BZR1/2</i>		14	K14503	
<i>TCH4</i>		5	K14504	
<i>CYCD3</i>		13	K14505	
生长素		<i>AUX1</i>	13	K13946
		<i>TIR1</i>	20	K14485
	<i>AUX/IAA</i>	77	K14484	
	<i>ARF</i>	131	K14486	
	<i>GH3</i>	24	K14487	
	<i>SAUR</i>	64	K14488	
	细胞分裂素	<i>CRE1</i>	41	K14489
<i>AHP</i>		21	K14490	
<i>B-ARR</i>		159	K14491	
<i>A-ARR</i>		24	K14492	
脱落酸	<i>PYL</i>	21	K14496	
	<i>PP2C</i>	43	K14497	
	<i>SNRK2</i>	45	K14498	
	<i>ABF</i>	59	K14432	
茉莉酸	<i>JAR1</i>	12	K14506	
	<i>COI1</i>	5	K13463	
	<i>JAZ</i>	72	K13464	
	<i>MYC2</i>	87	K13422	

(表 8)显示,赤霉素注释的酶基因最多,包括内根-贝壳杉烯合成酶(*KS*)、古巴焦磷酸合成酶(*CPS*)、内根-贝壳杉烯氧化酶(*KO*)及 GA_3 氧化酶(*GA_3ox*)等 8 个合成通路中的关键酶基因,一共对应 67 条 Unigene;油菜素内酯与水杨酸合成途径中均有 1 个酶基因被注释到,为异分支酸合酶(*ICS*)与油菜素内酯-6-氧化酶 2(*CYP85A2*),分别对应 9 条和 24 条 Unigene;茉莉酸合成途径 2 个关键酶基因丙二烯氧化物合成酶(*AOS*)以及丙二烯氧化物环化酶(*AOC*)分别注释到 6 条和 7 条 Unigene,生长素一共注释到 42 条 Unigene,色氨酸转氨酶(*TAA1*)及吲哚-3-丙酮酸单加氧酶(*YUC*)分别注释到了其中 22 条和 20 条;乙烯合成途径中的乙烯合成前体合成酶(*ACS*)与乙烯合成前体氧化酶(*ACO*)分别对应 9 和 14 条 Unigene;脱落酸及玉米素的生物合成途径中均注释到 3 个关键酶基因,分别有 29 条和 45 条 Unigene。

表 8 转录组中植物激素生物合成途径 Unigene 注释情况

植物激素	基因名称	Unigene 数目(条)	代谢通路编号
赤霉素	<i>CPS</i>	7	K04120
	<i>KS</i>	9	K04121
	<i>KO</i>	2	K04122
	<i>KA0</i>	5	K04123
	<i>GA_{20ox}</i>	16	K05282
	<i>GA_{3ox}</i>	3	K04124
	<i>GA_{13ox}</i>	8	K20666
	<i>GA_{2ox}</i>	17	K04125
水杨酸	<i>ICS</i>	24	K14759
油菜素内酯	<i>CYP85A2</i>	9	K12640
茉莉酸	<i>AOS</i>	6	K01723
	<i>AOC</i>	7	K10525
生长素	<i>TAA1</i>	22	K16903
	<i>YUC</i>	20	K11816
乙烯	<i>ACS</i>	9	K01762
	<i>ACO</i>	14	K05933
脱落酸	<i>ZEP</i>	5	K09838
	<i>NCED</i>	12	K09840
	<i>AO</i>	12	K09842
玉米素	<i>IPT</i>	3	K10760
	<i>CYP735A</i>	2	K10717
	<i>CKX</i>	40	K00279

2.3.2 植物激素合成途径中 Unigene 注释情况

转录组中激素合成途径酶基因 Unigene 注释情况

3 讨论与结论

近年来,转录组测序技术发展迅速,同时伴随

着多个生物信息学分析平台的加入,大量研究即便在没有基因组数据支撑的前提下,通过转录组测序所得的结果也可在挖掘植物未知基因、明确相关生理功能的代谢途径及基因调控机制等方面提供海量相关信息。无需参考基因组的数据便可对目标物种进行分子生物学方面的研究,已成为中草药材基因信息挖掘的重要研究手段^[12]。利用该技术,前人已从多种植物中挖掘出有关花器官发育的基因,并分析了相关基因的调控机制。李梅等利用雌蕊缺失茶树(*Camellia sinensis*)花的花、花蕾、花芽作为材料,基于转录组测序技术,共发掘出了 34 个花器官发育差异表达的相关基因,包括 *KNOX* 家族基因、*WUS* 类基因、花器官发育 ABCDE 模型的相关基因等,证明茶树花即使是在雌蕊缺失的情况下,有关雌蕊发育的基因也是同样存在^[13]。而在长瓣兜兰(*Paphiopedilum dianthum*)花蕾和花朵的转录组测序中发现,相较于花蕾期,花朵时期中调控花朵发育相关的基因,如 *AG*、*C2H2-ZEP* 等基因的表达量明显下调,大量基因表达下调的原因可能是在花朵时期,花的各部分器官已完成分化,调控器官分化基因的相关任务已完成^[14]所致。位明明等通过该技术初步明确了多个参与编码橡胶树(*Hevea brasiliensis*)花器官发育的基因家族,如 *Mads-box*、*MYB* 及 *AP2* 等基因家族^[15]。

本研究通过 RNA-Seq 技术构建了阳春砂不同时期花丝以及花柱总样品的转录组数据库,共获得 94 584 条 Unigene, Unigene 的平均长度为 842 bp, N50 为 1 582 bp, GC 含量为 44.35%。其中,共有 62 174 条 Unigene 被注释到 GO、KEGG、NR 和 KOG 等七大数据库,占总 Unigene 的 65.73%。注释率较高,测序所得结果与拼接质量较好。这些注释的 Unigene 为阳春砂的代谢途径、基因功能分类、花器官发育及雌雄蕊运动分析等方面提供参考依据。此外,一共有 32 410 条 Unigene 未被注释到,占总 Unigene 的 34.27%,推测这些 Unigene 可能为阳春砂中的非编码 RNA 序列或现有基因数据库尚未完善所致。在 NR 数据库中共注释到 58 669 条 Unigene,其中 77.14% 的 Unigene 被注释到姜科植物小果野芭蕉中,两者比对所得同源基因数目较多的原因可能是亲缘关系较近。转录组所得 Unigene 在 GO 功能分类中,注释到生物学过程、细胞组分以及分子功能三大类,分别有 24、17、14 个亚类,主要集中在代谢过程、催化活性、细胞部分和细胞过程

等功能。而在 KOG 数据库的分析结果中,共有 45 892 条 Unigene 得到了注释,在 25 个功能大类中,一般功能预测的 Unigene 最多,共有 11 117 条,信号转导机制、转录以及翻译后的修饰、蛋白质转换与代谢功能次之,核结构与细胞运动最少,分别仅有 346 条和 75 条 Unigene。通过 KEGG 数据库分析发现,被注释到 44 000 条 Unigene 参与了 23 类共计 137 个 KEGG 代谢通路,这些通路主要集中在代谢途径、次生代谢产物的生物合成、RNA 转运以及植物激素信号转导等,可以通过所注释的通路全面了解阳春砂花丝及花柱的代谢途径信息。同时,在阳春砂花丝、花柱转录组中一共挖掘到 18 895 个 SSR 位点,在 6 种不同的核苷酸重复类型中均有分布,证明本研究转录组中位点的类型丰富,分布密度较大,具有良好的多态性潜能;在二核苷酸和三核苷酸中,重复最多的基序是分别是 AG/CT 和 AGG/CTT,该结果与王焕的研究结果^[16]一致。

植物激素是植物通过自身代谢所产生的一些有机信号分子,可在低浓度下产生明显的生理效应,并在合成部位发挥功能^[17-18]。而花器官的运动受到激素调节,Luo 等研究发现,外源施加 IAA 后,会明显影响姜科植物花柱的弯曲程度^[19]。有些植物激素也会调控姜科植物蓝猪耳二长雄蕊的翻转与伸长^[20]。此外,其他植物激素同样也在雌雄蕊发育中起着重要的调控作用。JA 可调控花药的开裂,JA 生物合成通路的相关基因 *DADI* 与 *OPR3* 在生长素感知缺陷的拟南芥突变体中的表达量提高,可提前使得花药产生开裂,说明此过程是一种负向调节,即生长素可以通过介导 JA 的生物合成来调节花药的发育^[21-22]。在模式植物拟南芥的其他研究中,茉莉酸与赤霉素信号转导相关基因也在拟南芥雄蕊的发育过程中起到了重要的调控作用^[8,23-24]。在本研究重点关注的植物激素合成与信号转导途径中一共有 2012 条 Unigene 被注释,囊括了赤霉素、生长素、茉莉酸及脱落酸等 9 种重要的植物激素,为进一步从分子层面阐述阳春砂花器官的发育及运动提供了大量候选基因。

笔者所在课题组于 2012 年提出通过生理手段改变阳春砂的花器结构,以适应更多种类的昆虫传粉而提高产量的研究思路^[25]。近 10 年来,已经在阳春砂的花芽分化规律^[26]、花器结构特征^[3]、假合蕊柱的形成过程^[10]、某些生殖生物学特性^[27]、落果规律^[28]等多方面进行了研究,明晰了阳春砂生产过

程中由于特殊花器官结构导致传粉昆虫种类受限和低产的原因。因此,下一步的研究方向需对本研究中所获有关阳春砂花器官的发育及运动的关键候选基因进行更深层次的功能分析与验证,并对阳春砂假合蕊柱形成与相关基因表达模式的关系进行分析与探讨,不仅可为深入开展阳春砂花器官发育与运动相关基因的克隆及相互作用的模式分析提供数据基础,也为进一步深入揭示阳春砂假合蕊柱的分子调控机制提供了理论依据,以期形成人为干预的效果进而改变阳春砂花器官的形态,拓宽阳春砂唇瓣与假合蕊柱之间的距离,为阳春砂的授粉虫媒打造一个相对有利的授粉环境,从而提高自然结实率,达到增加农民收入的目的。

参考文献:

- [1]中国科学院中国植物志编辑委员会. 中国植物志:第四十六卷[M]. 北京:科学出版社,1981:125.
- [2]国家药典委员会. 中华人民共和国药典(一部)[M]. 2020年版. 北京:中国医药科技出版社,2020:264-265.
- [3]何国振,高伟,苏景,等. 药用植物阳春砂花器形态特征[J]. 植物学报,2014,49(3):313-321.
- [4]彭建明,李荣英,李戈,等. 阳春砂仁授粉特性与传粉昆虫的研究[J]. 云南中医学院学报,2012,35(4):51-55.
- [5]Du W, Qin K Z, Wang X F. The mechanism of stamen movement in *Chimonanthus praecox* (Calycanthaceae): differential cell growth rates on the adaxial and abaxial surfaces of filaments after flower opening[J]. Plant Systematics and Evolution, 2012, 298(3):561-567.
- [6]罗银玲,任盘宇,李庆军. 马来良姜花柱卷曲运动的结构基础[J]. 植物学报,2009,44(2):191-196.
- [7]王燕. 高温诱导番茄柱头外露的生理及分子基础的研究[D]. 杭州:浙江大学,2015:50-56.
- [8]Song S S, Qi T C, Huang H, et al. Regulation of stamen development by coordinated actions of jasmonate, auxin, and gibberellin in *Arabidopsis*[J]. Molecular Plant, 2013, 6(4):1065-1073.
- [9]Dill A, Jung H S, Sun T P. The DELLA motif is essential for gibberellin-induced degradation of RGA[J]. Proceedings of the National Academy of Sciences of the United States of America, 2001, 98(24):14162-14167.
- [10]何卓航,陈红,汤丽云,等. 阳春砂小花假合蕊柱的形成过程[J]. 中国实验方剂学杂志,2020,26(9):176-183.
- [11]陈红. 阳春砂小花形态结构与假合蕊柱形成机理的初步研究[D]. 广州:广州中医药大学,2015:14.
- [12]王尧龙,黄璐琦,袁媛,等. 药用植物转录组研究进展[J]. 中国中药杂志,2015,40(11):2055-2061.
- [13]李梅,陈林波,田易萍,等. 雌蕊缺失茶树花3个发育期的数字基因表达谱分析[J]. 茶叶科学,2017,37(1):97-107.
- [14]颜凤霞,王莲辉,田凡,等. 长瓣兜兰花2个不同时期转录组分析[J]. 种子,2021,40(5):91-97.
- [15]位明明,王亚杰,李维国. 橡胶树生殖发育相关基因的RNA-Seq转录组分析[J]. 分子植物育种,2017,15(1):99-106.
- [16]王焕. MeJA影响阳春砂挥发性萜类和转录组变化的研究[D]. 广州:广州中医药大学,2015:39.
- [17]黎家,李传友. 新中国成立70年来植物激素研究进展[J]. 中国科学:生命科学,2019,49(10):1227-1281.
- [18]戴毅,田龙果,潘贞志,等. 激素和非生物逆境胁迫调控植物硝酸盐转运蛋白功能的研究进展[J]. 江苏农业学报,2020,36(6):1595-1604.
- [19]Luo Y L, Bi T J, Li D, et al. Effects of indole-3-acetic acid and auxin transport inhibitors on the style curvature of three *Alpinia* species (Zingiberaceae)[J]. Acta Physiologiae Plantarum, 2012, 34(5):2019-2025.
- [20]蔡丹,孙姝兰. 植物生长物质对蓝猪耳二长雄蕊运动的影响[J]. 华南师范大学学报(自然科学版),2014,46(6):93-97.
- [21]Ishiguro S, Kawai-Oda A, Ueda J, et al. The *DEFECTIVE IN ANther DEHISCENCE* gene encodes a novel phospholipase A1 catalyzing the initial step of jasmonic acid biosynthesis, which synchronizes pollen maturation, anther dehiscence, and flower opening in *Arabidopsis*[J]. The Plant Cell, 2001, 13(10):2191-2209.
- [22]Cecchetti V, Altamura M M, Brunetti P, et al. Auxin controls *Arabidopsis* anther dehiscence by regulating endothecium lignification and jasmonic acid biosynthesis[J]. The Plant Journal, 2013, 74(3):411-422.
- [23]Song S S, Qi T C, Huang H, et al. The jasmonate-ZIM domain proteins interact with the R2R3-MYB transcription factors MYB21 and MYB24 to affect jasmonate-regulated stamen development in *Arabidopsis*[J]. The Plant Cell, 2011, 23(3):1000-1013.
- [24]Qi T C, Huang H, Song S S, et al. Regulation of jasmonate-mediated stamen development and seed production by a bHLH-MYB complex in *Arabidopsis*[J]. The Plant Cell, 2015, 27(6):1620-1633.
- [25]汤丽云,何国振,苏景,等. 道地春砂仁产业发展的策略研究[J]. 中国农学通报,2012,28(8):94-99.
- [26]何国振,高伟,苏景,等. 药用植物阳春砂花芽分化规律研究[J]. 国际医药卫生导报,2012,18(15):2142-2144.
- [27]Yang R P, Wang J, Gao W, et al. Research on the reproductive biological characteristics of *Amomum villosum* Lour. and *Amomum longiligulare* T. L. Wu[J]. PLoS One, 2021, 16(8):e0250335.
- [28]吕秉鼎,胡佳佳,汤丽云,等. 阳春砂落果规律及其生理机制[J]. 植物生理学报,2021,57(2):429-438.