

DOI:10.11913/PSJ.2095-0837.2017.40609

吕海舟, 刘琬菁, 何柳, 徐志超, 罗红梅. 植物次生代谢基因簇研究进展[J]. 植物科学学报, 2017, 35(4): 609-621

Lü HZ, Liu WJ, He L, Xu ZC, Luo HM. Advances on the study of gene clusters involved in plant secondary metabolism[J]. *Plant Science Journal*, 2017, 35(4): 609-621

植物次生代谢基因簇研究进展

吕海舟, 刘琬菁, 何柳, 徐志超, 罗红梅*

(中国医学科学院药用植物研究所, 中草药物质基础与资源利用教育部重点实验室, 北京 100193)

摘要: 类似于原核生物的操纵子, 在真核生物(如酵母、真菌、昆虫等)基因组中也出现了彼此功能相关的非同源基因成簇存在的现象。这些基因形成基因簇, 可参与多种次生代谢途径。近年来, 植物中也发现了越来越多的参与次生代谢产物合成的基因簇, 它们已成为植物生物学研究的热点。本文总结并分析了植物中已鉴定的次生代谢基因簇。这些基因簇存在于玉米(*Zea mays* L.)、水稻(*Oryza sativa* L.)、拟南芥(*Arabidopsis thaliana* (L.) Heynh.)、番茄(*Solanum lycopersicum* L.)等植物的基因组中, 分别参与合成苯并噁唑酮类、萜类和生物碱类等次生代谢产物。本文通过解析这些基因簇的组成及结构特点, 对其特征进行总结, 探讨了基因簇形成的分子机理及其调控机制, 对植物次生代谢基因簇在合成生物学及代谢工程学中的研究方向和应用前景进行了展望。

关键词: 次生代谢; 基因簇; 形成与调控; 合成生物学

中图分类号: Q943.2

文献标识码: A

文章编号: 2095-0837(2017)04-0609-13

Advances on the study of gene clusters involved in plant secondary metabolism

Lü Hai-Zhou, Liu Wan-Jing, He Liu, Xu Zhi-Chao, Luo Hong-Mei*

(Key Laboratory of Bioactive Substances and Resources Utilization of Chinese Herbal Medicine, Ministry of Education, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Beijing 100193, China)

Abstract: Genes encoding enzymes involved in the sequential biosynthesis steps of secondary metabolites are clustered together in the prokaryotic genome, also known as ‘operons’. In the genome of eukaryotes (e. g. yeasts, fungi, and insects), a set of functional genes responsible for special metabolite biosynthesis has also been discovered clustered in the chromosome. Recently, several secondary metabolic gene clusters have been identified in plants, such as *Zea mays* L., *Oryza sativa* L., *Arabidopsis thaliana* (L.) Heynh., and *Solanum lycopersicum* L.. In this review, we summarize the identified gene clusters involved in the biosynthesis of benzoxazinoids, terpenoids, and alkaloids in plants, as well as analyze the mechanisms of gene cluster formation and regulation. In addition, the application prospects of these gene clusters are discussed. Increasing knowledge of plant metabolic gene clusters will facilitate the development of synthetic biology and metabolic engineering.

Key words: Secondary metabolism; Gene cluster; Formation and regulation; Synthetic biology

收稿日期: 2016-12-08, 退修日期: 2016-12-28。

基金项目: 国家自然科学基金项目(31570302); 中国医学科学院医学与健康科技创新工程(2016-I2M-3-016)。

This work was supported by grants from the National Natural Science Foundation of China (31570302) and CAMS Innovation Fund for Medical Sciences (CIFMS) (2016-I2M-3-016)。

作者简介: 吕海舟(1992-), 女, 硕士研究生, 研究方向为药用植物功能基因组(E-mail: 18010108227@163.com)。

* 通讯作者(Author for correspondence. E-mail: hmluo@implad.ac.cn)。

基因簇是指在染色体上成簇出现并协同转录的非同源基因, 它们通常编码一些特定化合物生物合成途径中催化连续步骤的酶。全基因组测序已充分证实了基因簇在细菌基因组中广泛存在^[1]。这些基因簇可以合成多种次生代谢产物。次生代谢产物是指那些并非生物正常生长所必需、但具有重要作用的化合物, 例如, 生物体内作为生理反应中的信号分子或具有防御作用的化合物等。由细菌产生的一些次生代谢产物具有可为人类所利用的生物学特性, 有些可以作为抗生素、抗癌药物、杀虫剂、免疫抑制剂或除草剂等^[2]。

除细菌外, 其它真核生物基因组中也出现了基因簇, 在酵母、真菌、昆虫、脊椎动物中均有报道^[3]。例如, 酿酒酵母 (*Saccharomyces cerevisiae*) 中存在参与尿囊素分解代谢途径的基因簇^[4], 动物中存在主要组织相容性复合体基因簇, 它们在动物固有免疫系统及适应性免疫过程中具有对抗病原菌的作用^[5]。近年来, 次生代谢基因簇的研究成为新的植物生物学研究领域, 在特异性化合物的合成生物学研究中备受关注。

真核生物基因簇不同于原核生物操纵子是由于操纵子中的基因多源自于基因水平转移^[6], 而真核生物基因簇的基因来源则更为复杂。此外, 虽然操纵子与基因簇都受到调控因子的调节, 但操纵子在一个启动子的作用下转录为单顺反子 mRNA, 而基因簇中的各个基因会转录为独立的 mRNA^[7]。

在植物基因组中, 具有高度序列相似性的同源基因成簇存在的现象极为普遍^[8], 但近几年发现有些功能相关的非同源基因也可以形成基因簇。至今为止, 基因簇中已鉴定的基因均可以编码参与次生代谢途径的酶, 但参与初生代谢的基因簇尚未见报道。目前, 仍不确定参与植物次生代谢的基因簇是否同真菌和细菌基因簇一样在代谢过程中占主导地位。

本文综述了植物次生代谢基因簇的研究进展, 分析了不同基因簇的结构特点, 探讨了基因簇形成的分子机理及调控机制, 并展望了植物基因簇在合成生物学及代谢工程研究中的应用前景。

1 植物次生代谢基因簇参与多种化合物的生物合成

1997 年, Frey 等^[9]报道了玉米 (*Zea mays*

L.) 中一种防御性化合物的合成基因在基因组中成簇存在, 这是第一个被发现的植物次生代谢基因簇。燕麦 (*Avena strigosa* L.) 中的燕麦根皂苷^[10]和水稻 (*Oryza sativa* L.) 中的抗毒素基因簇^[11]也被相继鉴定。这些基因簇出现在不同的植物种属中, 包括单子叶及双子叶植物, 参与了苯并噁唑啉酮类 (Benzoxazinoids)、生氰糖苷类 (Cyanogenic glucosides, CGs)、萜类 (Terpenoids) 和生物碱类 (Alkaloids) 等不同化合物的合成 (表 1)。

1.1 合成苯并噁唑啉酮类化合物

玉米中发现的 Bx 基因簇, 可以合成苯并噁唑啉酮类物质 DIMBOA (2, 4-dihydroxy-7-methoxy-2H-1, 4-benzoxazin-3 (4H)-one)。最初在玉米的 4 号染色体上发现 5 个基因 (即 Bx1 ~ Bx5) 成簇排列^[9], 随后在该染色体上又发现了 Bx6、Bx7 和 Bx8, 但 Bx8 距离 Bx1 ~ Bx5 较近, 而 Bx6、Bx7 距离 Bx1 ~ Bx5 较远^[12, 13]; Bx9 是 Bx8 的同源基因^[13], 但 Bx9 位于 1 号染色体上; Glu1 与 Glu2 也属于该基因簇, 但位于 10 号染色体上^[14]。

Bx1 是玉米中色氨酸合酶 α -亚基的同源基因^[41], 可以将吲哚-3-甘油磷酸 (indole-3-glycerol phosphate) 转化为吲哚。Bx2 ~ Bx5 均编码 CYP71C, 该类蛋白属于细胞色素 P450 (单氧化酶家族, 可以催化吲哚转化为 DIBOA 的中间步骤^[9])。Bx8 和 Bx9 属于葡糖基转移酶, 可以糖基化 DIBOA 生成 DIBOA-Glc, DIBOA-Glc 在双加氧酶 Bx6 及甲基转移酶 Bx7 的作用下转化为 DIMBOA-Glc。Glu1 与 Glu2 编码 β -葡糖苷酶, 催化 DIBOA-Glc 或 DIMBOA-Glc 生成 DIBOA 或 DIMBOA^[14]。

在生物合成苯并噁唑啉酮类化合物的其它禾本科 (Gramineae) 植物, 如小麦 (*Triticum aestivum* L.) 和黑麦 (*Secale cereale* L.) 中^[42], Bx 基因簇中的基因分散在不同的染色体上, 没有完全形成基因簇。在栽培大麦的 4 个品种 (*Hordeum vulgare* ‘Alexis’、*H. vulgare* ‘Baccara’、*H. vulgare* ‘Nuernberg’ 和 *H. vulgare* ‘Tellus’) 基因组中, Bx 基因均已丢失^[14]; 野生大麦 *H. spontaneum* C. Koch. 也丢失了全部相关基因, 但是其它野生大麦 (如 *H. lechleri* L.) 仍然存在 Bx 基因簇, 因而保留了合成苯丙恶嗪的功能 (DIBOA; Bx1 ~ Bx5)^[43]。DIBOA 和 DIMBOA 都可以抵御病原体

表 1 植物中已鉴定的次生代谢基因簇

Table 1 Identified gene clusters involved in plant secondary metabolism

植物 Plant	基因 数目 Gene No.	已鉴定基 因数目 Identified gene No.	细胞色素 P450 数目 CYP450 No.	基因簇长度 Length of gene cluster (kb)	基因密度 Gene density (gene/kb)	染色体 Chromosome	化合物 Compound	化合物类型 Type of compound	参考文献 Reference
玉米 <i>Zea mays</i> L.	11	11	4	17260	1/1917.78	1, 4, 10	DIMBOA	苯并噁唑啉酮类 Benzoxazinoids	[9, 12–14]
百脉根 <i>Lotus japonicus</i> (Regel) K. Larsen	5	4	3	355	1/39.44	3	亚麻苦苷 Linamarin、百脉 根苷 Lotaustrolin	生氰糖苷 Cyanogenic glucosides	[15]
木薯 <i>Manihot esculenta</i> Crantz	5	4	3	83	1/6.38	–	百脉根苷 Lotaustrolin	生氰糖苷 Cyanogenic glucosides	[15]
高粱 <i>Sorghum bicolor</i> (L.) Moench	3	3	2	104	1/14.86	1	蜀黍苷 Dhuririn	生氰糖苷 Cyanogenic glucosides	[16–18]
番茄 <i>Solanum lycopersicum</i> L.	15	4	2	107	1/7.13	8	β-水芹烯 β-phellandrene、Lycosan- talonal	单萜 Monoterpene 二萜 Diterpene	[19–22]
水稻 <i>Oryza sativa</i> L.	5	5	2	170	1/34.00	4	Momilactone A	二萜 Diterpene	[11, 23, 24]
水稻 <i>Oryza sativa</i> L.	10	10	6	245	1/24.50	2	Phytocassanes A-E、 Oryzallides、oryzadiolones	二萜 Diterpene	[25–27]
燕麦 <i>Avena strigosa</i> L.	10	7	1	190	1/38.00	–	燕麦根皂苷 Avenacin	三萜 Triterpene	[28–33]
拟南芥 <i>Arabidopsis thaliana</i> (L.) Heynh.	4	3	2	23	1/5.75	5	Thalianol	三萜 Triterpene	[34]
拟南芥 <i>Arabidopsis thaliana</i> (L.) Heynh.	3	1	2	26	1/8.67	5	Marneral	三萜 Triterpene	[35]
百脉根 <i>Lotus japonicus</i> (Regel) K. Larsen	4	3	2	62	1/15.50	–	20-羟基-白桦脂酸 20-hydroxy-betulinic acid	三萜 Triterpene	[36, 37]
罂粟 <i>Papaver somniferum</i> L.	10	6	4	221	1/22.10	–	诺斯卡品 Noscapine	生物碱 Alkaloids	[38]
番茄 <i>Solanum lycopersicum</i> L.	9	7	2	255	1/28.33	7, 12	α-番茄碱 a-tomatine	生物碱 Alkaloids	[39, 40]
马铃薯 <i>Solanum tuberosum</i> L.	7	5	2	330	1/47.14	7, 12	α-茄碱 a-solanine、 α-卡茄碱 a-chaconine	生物碱 Alkaloids	[39, 40]

和食草动物的侵袭^[44]。值得一提的是, DIMBOA 代谢途径几乎仅存在于禾本科植物中, 但在玄参科 (Scrophulariaceae) 与爵床科 (Acanthaceae) 物种中也有报道^[44]。

1.2 合成生氰糖苷类化合物

生氰糖苷类化合物 (CGs) 是一类氨基酸来源的防御性化合物, 其经特殊的 β -葡萄糖苷酶水解后会生成氰化氢。CGs 包含约 60 种不同的代谢产物。在百脉根 (*Lotus japonicus* (Regel) K. Larsen) 中, 参与亚麻苦苷与百脉根苷生物合成的基因已被鉴定, 它们在基因组中形成基因簇^[15], 位于 3 号染色体上, 这 2 种化合物对应的前体分别是 L-缬氨酸和 L-异亮氨酸。P450 酶 CYP79D3/4 和 CYP736A2 通过 2 步连续催化反应, 将前体转化为羟基乙腈中间体, 再由 UDP-葡萄糖基转移酶 UGT85K3 介导完成葡萄糖基化^[15]。木薯 (*Manihot esculenta* Crantz) 亚麻苦苷和百脉根苷的生物合成基因同样以基因簇形式存在, 且前体化合物同样为 L-缬氨酸和 L-异亮氨酸; 合成的酶包括 CYP79D2、CYP71E7 和 UGT85K4/5, 生物合成反应步骤也与百脉根中的反应步骤类似^[15]。CYP79D1 与 CYP79D2 相邻, 但 CYP79D1 并不是基因簇的成员, 表明 CYP79D1 由 CYP79D2 复制而来。在高粱 (*Sorghum bicolor* (L.) Moench) 中, 生氰糖苷类化合物蜀黍苷由 CYP79A1、CYP71E1 和 UGT85B1 酶按相似的反应步骤, 催化 L-酪氨酸转化合成^[16-18], 该基因簇位于 1 号染色体上。

在上述 3 个合成 CGs 的基因簇中, 基因在基因簇中的排列位置不同, 基因簇编码的关键酶分属于不同的家族 (CYP71 和 CYP736)。此外, 基因簇的大小也不同, 其中百脉根的 CGs 基因簇在染色体上所跨区域最大。同时, 基因密度也不同, 如木薯 CGs 合成基因簇的基因种类最多, 而高粱 CGs 基因簇基因密度最低。另外, 这些基因簇近旁的基因及转座子序列也不同, 说明这些基因簇在 3 个物种中是独立形成的^[15]。

1.3 合成萜类化合物的基因簇

1.3.1 合成单萜化合物

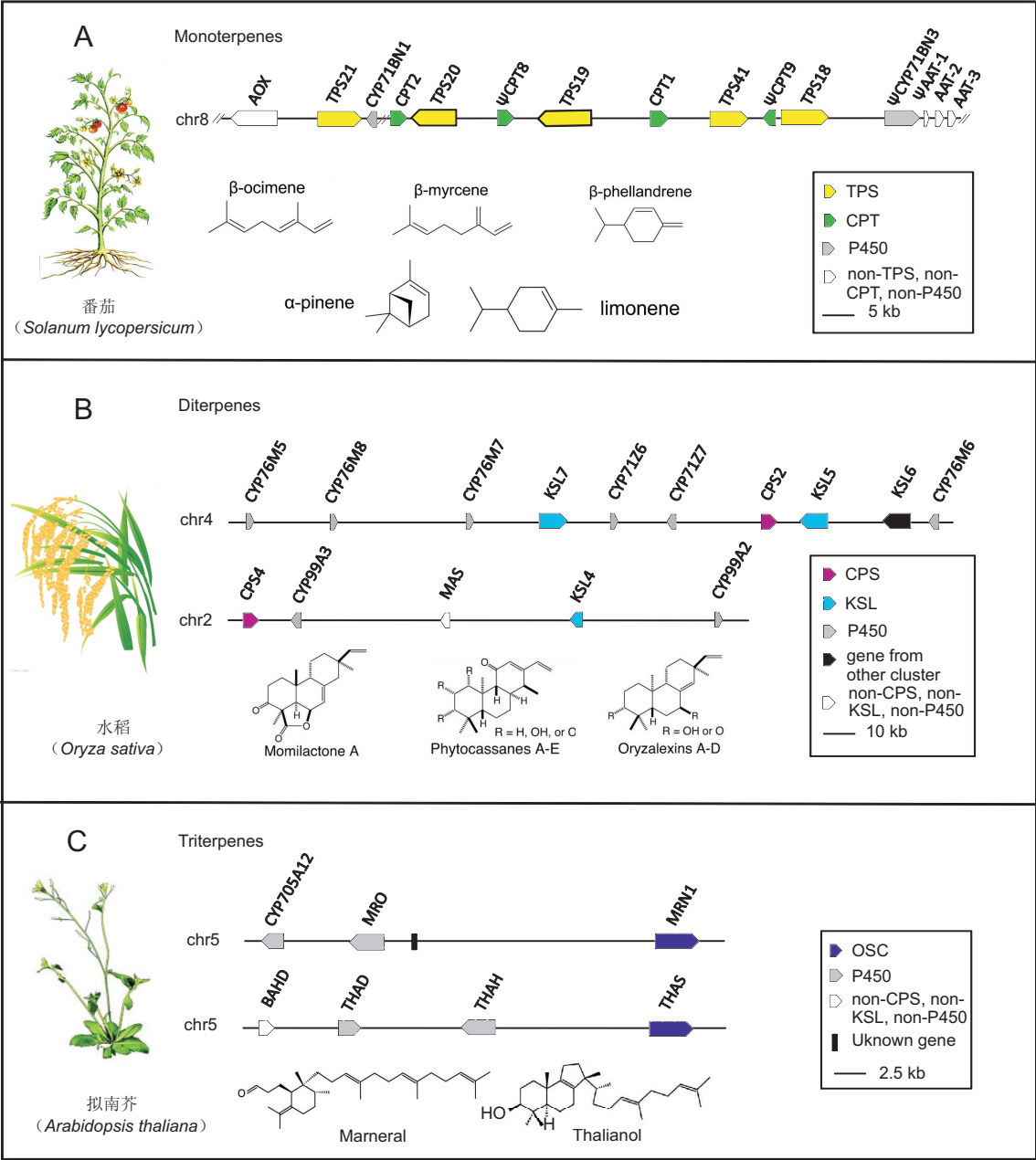
迄今, 高等植物中鉴定到的数量最多的是参与萜类化合物生物合成的基因簇。番茄中发现 1 个参与单萜合成的基因簇^[19], 位于 8 号染色体上 (图 1: A)。该基因簇包括 5 个萜类合酶基因 (*TPS18* ~

TPS21 和 *TPS-41*), 它们利用底物橙花基焦磷酸 (neryl diphosphate, NPP), 即牻牛儿基焦磷酸 (geranyl diphosphate, GPP) 的顺式形式合成单萜。其中, *TPS20*, 也被称为水芹素合酶 1 基因 (phellandrene synthase1, *PHS1*), 编码 1 个质体定位的酶, 催化 NPP 形成 β -水芹烯和几个其它类型的单萜^[20]。此外, 这个基因簇还包含 2 个顺式异戊二烯基转移酶基因 (cis-prenyl transferases, *CPTs*), 其中 *CPT1* 也叫橙花基焦磷酸合酶 1 基因 (neryl diphosphate synthase 1, *NDPS1*), 合成二磷酸前体 NPP^[21]。另外, 这个基因簇中还包含 1 个编码 P450 酶的基因^[21]。

番茄中, SI-CPT1 催化 NPP 合成, SI-TPS19 和 SI-TPS20 利用 NPP 为底物, 这 3 个酶的编码基因在基因簇中距离很近, 并且都在腺毛中高丰度表达, 这与腺毛中确实存在由 NPP 合成的单萜化合物的结果相一致。SI-TPS41 也在腺毛中高表达, 虽然至今为止没有发现其编码蛋白具有活性。对于 SI-TPS18 而言, 虽然也未检测到蛋白活性, 但其编码基因在根中表达量最高^[21]。另外, 顺式异戊二烯基转移酶—SI-CPT2, 可以催化牻牛儿基焦磷酸 (geranyl geranyl diphosphate, GGPP) 的形成, 而 SI-TPS21 可以利用 GGPP 为底物, 在基因簇中它们的编码基因彼此距离较近, 并具有相似表达模式^[21]。有报道表明, SI-CPT2 和 SI-TPS21 以及一个 P450 基因在腺毛外形成了另一个亚基因簇, 催化二萜类化合物的生物合成^[22]。一些番茄品种, 如 *S. pimpinellifolium* L.、*S. pennellii* L.、*S. tuberosum* L. 和 *S. habrochaites* L. 的基因组中均存在该基因簇。其中, *S. pimpinellifolium* 与 *S. lycopersicum* 基因簇的结构最为相似, 而其余 3 个品种基因簇的组成、排列顺序及排列方向等均与 *S. lycopersicum* 和 *S. pimpinellifolium* 基因簇相差较大。

1.3.2 合成二萜化合物

最近发现番茄中 8 号染色体的基因簇除以上介绍的基因外, 还包含了一个 P450 氧化酶基因 *CYP71BN1* (图 1: A)。将 *CPT2* 和 *TPS21* 在大肠杆菌中共表达, 可以合成二萜 lycosantalene; 将 *CPT2*、*TPS21* 和 *CYP71BN1* 在大肠杆菌中共表达, 可以合成二萜 lycosantalanol。Lycosantalanol 是 lycosantalene 的氧化产物。3 个基因均在叶柄



A: 番茄中单萜化合物合成基因簇，引自参考文献 [19–22]；B: 水稻中二萜化合物合成基因簇，引自参考文献 [23–27]；C: 拟南芥中三萜化合物合成基因簇，引自参考文献 [34, 35]。图中植物图片均来源于网络。
A: Gene cluster for biosynthesis of monoterpene in *Solanum lycopersicum*, cited from references [19–22]；B: Gene cluster for biosynthesis of diterpene in *Oryza sativa*, cited from references [23–27]；C: Gene cluster for biosynthesis of triterpene in *Arabidopsis thaliana*, cited from references [34, 35]. Pictures of plants were downloaded from the internet.

图 1 萜类化合物生物合成基因簇示例图

Fig. 1 Examples of gene clusters involved in terpenoid biosynthesis in plants

处表达量最高，但在叶柄的腺毛处却表达量较低。而且二者在野生植株的叶柄或其它组织均未检测到。但在 35S CaMV 启动子启动下，过表达 *CPT2* 的转基因番茄的叶柄中可以检测到它们。这些结果说明在野生型植株的叶柄或者其它组织处可以合成

lycosantalanol 和 lycosantalene，但可能由于含量极低而检测不到，也有可能是这些化合物在植物中被进一步进行了结构修饰。*CPT2*、*TPS21* 和 *CYP71BN1* 在腺毛以外的部位参与二萜化合物的生物合成，基因簇中的其它基因在腺毛中参与单萜

化合物的生物合成,表明该基因簇正在进一步分化成具有独特生理生化功能的“亚基因簇”^[22]。

在水稻中发现了 2 条合成二萜化合物的基因簇,产物分别为 momilactone A^[23] 和 phytocassanes(A-E)^[25](图 1: B)。参与 momilactone A 生物合成的基因簇位于 4 号染色体上。首先 CPS4 和 KSL4 通过 2 步连续的环化步骤,催化底物 GGPP 得到前体化合物,再由 CYP99A2/A3 介导进行氧化反应^[24],最后由 momilactone A synthase(MAS)催化合成 momilactone A^[23]。参与 phytocassanes(A-E)生物合成的基因簇位于 2 号染色体上,包括 CPS2、KSL5/6/7、CYP71Z6/7 和 CYP76M5/6/7/8 共 10 个基因。该基因簇具有多功能性,即可以合成多种化合物,如 CPS2、KSL7、CYP71Z7 和 CYP76M7/8 参与生成抗真菌的化合物 phytocassanes^[25],其中 CYP71Z7 的功能对映 C2-羟化酶^[26];CPS2、KSL6、CYP71Z6 和 CYP76M8 参与生成抗细菌的化合物 oryzalides/oryzadiones,其中 CYP71Z6 的功能为对映-异构贝壳杉烯 C2-羟化酶^[26];所以真菌和细菌感染均可激活该基因簇中相关基因的表达^[25]。CYP76M5/6/8 功能研究表明,CYP76M8 可以替代 CYP76M7 的功能^[27],并且与 CYP76M5/6 共同参与 oryzalexin 的生物合成,而且其上游二萜合酶基因不位于 2 号染色体的基因簇中^[27]。KSL5 也是该基因簇的 1 个成员,催化对映-柯巴基焦磷酸(ent-CPP)生成对映-8,14,15 海松二烯(entpimar-8,14,15-diene),但是其催化的最终产物尚不明确^[27]。

1.3.3 合成三萜化合物

三萜化合物可以参与植物防御反应,某些化合物还可用作抗癌药物^[45]。已经发现,超过 1000 种双子叶植物体内含有三萜化合物,而单子叶植物却几乎只有燕麦中含有。野生燕麦中发现的一个基因簇^[10],包含皂苷缺乏(saponin-deficient, SAD)基因,可参与三萜化合物燕麦根皂苷的生物合成^[28]。虽然已知 10 个 SAD 基因中有 9 个(SAD1~SAD3、SAD5~SAD10)^[28]在细胞减数分裂过程中共分离,但是 SAD3~SAD6 以及 SAD8 在染色体上的确切位置仍不明确。燕麦根皂苷生物合成的第一步由 SAD1 催化,其底物为 2,3-oxidosqualene。SAD1 是 1 个 β -香树精合酶,

它可能进化自一个环阿屯醇环化酶相似基因(cycloartenol cyclase-like)的复制和变异^[10]。环阿屯醇环化酶(cycloartenol cyclase, CAS1)是单子叶和双子叶植物中参与固醇和类固醇激素(油菜素内酯)生物合成的氧化鲨烯环化酶(oxidosqualene cyclases, OSCs)家族中的成员^[29]。燕麦根皂苷合成途径中的第二个酶—SAD2,属于单子叶植物特有的 P450 亚家族 CYP51H^[30],参与 β -香树精的氧化。SAD3 和 SAD4 负责糖基化反应^[28],SAD7 负责酰化反应^[31],SAD9 是甲基转移酶^[32],SAD10 是糖基转移酶^[33]。SAD5、SAD 6 和 SAD 8 在燕麦根皂苷生物合成过程中的作用尚不明确^[10]。

拟南芥中有 2 个基因簇,参与三萜化合物—thalianol^[34]和 marneral^[35]的生物合成,它们均位于 5 号染色体上(图 1: C)。Thalianol 和 marneral 的底物均为 2,3-oxidosqualene,催化该底物的酶分别为 thalianol 合酶(THAS)和 marneral 合酶(MRN1),它们都属于 OSCs 中保守的十字花科家族分支 II。Thalianol 可以被羟化酶 CYP708A2(THAH)和去饱和酶 CYP705A5(THAD)转化成去饱和的 thalian-diol。基因簇中的酰基转移酶 BAHD 的活性尚未被确认,虽然它与另外 3 个基因共表达^[34]。Marneral 基因簇中包括 3 个基因,第二个基因编码 CYP71A16(MRO),可能起氧化作用^[35],第三个基因编码 CYP705A12,与 THAD 都属于去饱和酶,可能催化相似的反应^[35]。过表达 THAS 和 MRN1 可导致拟南芥的矮化表型^[34,35],而 thas 和 mrn1 敲除突变体后分别表现出根伸长和开花延迟的性状^[34]。但到目前为止,这 2 个基因簇合成的化合物在植物中的具体作用仍不明确。

百脉根中存在另一个合成三萜化合物的基因簇^[36],包含 AMY2(编码 OSC 酶)、2 个 P450s(CYP88D5、CYP71D353)以及 SDRt。SDRt 与拟南芥的短链脱氢酶 HCF173 具有较高同源性^[37]。AMY2 合成三萜二氢羽扇豆醇(dihydrolupeol),再被 CYP71D353 转化成 20-羟基羽扇豆醇(20-hydroxylupeol)和 20-羟基-白桦脂酸(20-hydroxy-betulinic acid)。基因簇中 CYP88D5 的具体功能尚未被鉴定,SDRt 被认为具有调控作用^[36]。最近研究表明二氢羽扇豆醇可调控百脉根中的早期结瘤过程^[46],因此该基因簇在根和根瘤发育期会过量

表达。

1.4 合成生物碱类化合物

目前,已有研究鉴定到2个基因簇合成特殊的生物碱:罂粟的诺斯卡品(Noscapine)^[38]、番茄和马铃薯中均有的甾族糖苷生物碱(SGA)^[39]。生物碱是一类数目庞大的天然化合物。诺斯卡品属于异喹啉类生物碱中的四氯苯酚亚类化合物,结构上与众所熟知的吗啡和可待因(吗啡喃亚类)不同。与吗啡或可待因相比,诺斯卡品不会使人成瘾,并可以作为抗肿瘤药物,抑制癌细胞的微管蛋白聚合,诱导其凋亡^[47]。编码参与诺斯卡品生物合成酶的基因簇包括10个基因,分别是O-甲基转移酶和P450s。该基因簇只存在于富含诺斯卡品的罂粟物种中^[38],该途径的最初底物是斯氏紫堇碱,来自L-酪氨酸途径。10个基因中只有6个确定了功能,另外4个基因(CYP82X1、CYP82Y1、PSAT1和PSMT3)及几步氧化反应和酰化过程尚未被解析。TNMT编码1个四氢原小檗碱-顺式-N-甲基转移酶,最近研究表明,该基因也参与其它生物碱的合成^[38]。而在罂粟中,参与吗啡和可待因生物合成途径的酶,并没有被该基因簇上的基因所编码^[38]。

相比诺斯卡品而言,SGAs是蔬菜谷物里可以使人 and 动物产生不良生理反应的化合物,可以导致神经紊乱和胃肠功能紊乱,如马铃薯中的 α -茄碱和 α -卡茄碱、番茄中的 α -番茄碱^[39]。参与以上化合物合成的基因在马铃薯和番茄中都形成基因簇,并且都分布在7号和12号染色体上。在番茄基因组中,7号染色体上有7个基因,包括糖基转移酶GAME1(SGT1),3个UDP糖基转移酶GAME2(SGT3)、GAME17和GAME18,CYP72A亚家族的1个P450酶基因GAME6,2-氧戊二酸盐依赖的双加氧酶GAME11,还有1个细胞纤维素合酶类似蛋白的编码基因;而马铃薯的7号染色体上不包含GAME17和GAME18这2个基因,其余5个基因与番茄中是相同的^[40]。而且在这2个物种中,位于12号染色体上的2个基因也是相同的,它们是CYP88D亚家族的GAME4和转氨酶基因GAME12^[40]。基因GAME7和GAME8属于CYP72亚家族,它们在这2个物种中均与该基因簇的基因共表达,但是却位于基因组的其他染色体上。胆固醇是甾族生物碱合成的可能前体,

GAME7在SGA生物合成中的功能尚未被鉴定,GAME8可能介导22-氢化胆固醇的羟基化过程^[39]。

2 基因簇特点

2.1 普遍特征

一般而言,1个次生代谢基因簇中有至少3个彼此相邻的用于合成特异化合物的非同源基因。在这些非同源基因中,有的基因编码的酶用于合成次生代谢物的骨架,这样的基因/酶被称作“信号”基因/酶(signature genes/enzymes)^[48]。其它基因编码的酶称作“剪裁”酶(tailoring enzymes),用于对上述骨架化合物进行修饰,催化得到最终产物。“信号”酶基因通过在周围集合“剪裁”酶基因来形成一个代谢基因簇。以番茄中的单萜合成基因簇为例,其中TPS、CPT及P450等非同源基因在染色体上彼此相邻,TPS基因编码“信号”酶,合成单萜类化合物的骨架,CPT及P450基因编码“剪裁”酶,修饰该骨架,最终得到单萜类化合物。

至今为止,植物中所有已知的基因簇均参与了次生代谢途径,但没有关于初生代谢基因簇的报道。然而,这些基因簇与初生代谢具有一定相关性。基因簇中的“信号基因”似乎直接或间接地来自初生代谢基因的复制或基因的“新功能化”,尤其是与植物激素合成途径上的基因关系更为密切^[49]。例如,参与三萜 β -香树精、thalianol、marneral生物合成的合酶基因来自环阿屯醇合酶基因的复制^[34]。玉米基因簇中的Bx1基因可能来自于色氨酸合酶(合成植物生长激素)的 α -亚单元的复制,因为它们有60%的序列同源性^[41]。水稻中phytocassane合成基因簇中的CPS2和KSL7以及momilactone合成基因簇中的CPS4和KSL4都可能来自赤霉素生物合成途径基因的复制和变异^[25]。

所有基因簇中均存在P450s(表2),有些家族的P450s在单子叶和双子叶植物中均存在,如CYP79家族的成员,其中CYP79D3和CYP79A1分别在百脉根和高粱中参与生氰糖苷的生物合成^[15];而其它家族的P450s,如CYP51H亚家族,即参与燕麦根皂苷生物合成的SAD2,只存在于单子叶植物中^[30]。

表 2 植物次生代谢基因簇中的 P450s
Table 2 P450s identified in gene clusters involved in plant secondary metabolism

植物 Plant	细胞色素 P450 CYP450	P450 的功能 Function of CYP450	家族分类 Classification	参考文献 Reference
玉米 <i>Zea mays</i> L.	CYP71C1/2/3/4	氧化	CYP71 族	[9,12–14]
百脉根 <i>Lotus japonicus</i> (Regel) K. Larsen	CYP79D3/4、CYP736A2	–	CYP71 族	[15]
木薯 <i>Manihot esculenta</i> Crantz	CYP79D2、CYP71E7、CYP71E	–	CYP71 族	[15]
高粱 <i>Sorghum bicolor</i> (L.) Moench	CYP79A1、CYP71E1	–	CYP71 族	[16–18]
番茄 <i>Solanum lycopersicum</i> L.	CYP71BN1/3	氧化	CYP71 族	[19–22]
水稻 <i>Oryza sativa</i> L.	CYP99A2/3	氧化	CYP71 族	[11,23–27]
	CYP71Z6/7、CYP76M5/6/7/8	羟化	CYP71 族	
燕麦 <i>Avena strigosa</i> L.	CYP51H10	氧化	CYP51 族	[10,28–33]
	CYP708A2	羟化	CYP71 族	[34,35]
拟南芥 <i>Arabidopsis thaliana</i> (L.) Heynh.	CYP705A5/12	去饱和	CYP85 族	
	CYP71A16	氧化	CYP71 族	
百脉根 <i>Lotus japonicus</i> (Regel) K. Larsen	CYP71D353	羟化	CYP71 族	[36,37]
	CYP88D5	–	CYP85 族	
罂粟 <i>Papaver somniferum</i> L.	CYP82X1/2、CYP82Y1、CYP719A21	–	CYP71 族	[38]
	CYP72A	–	CYP72 族	[39,40]
番茄 <i>Solanum lycopersicum</i> L.	CYP88D	–	CYP85 族	
	CYP72A	–	CYP72 族	[39,40]
马铃薯 <i>Solanum tuberosum</i> L.	CYP88D	–	CYP85 族	

2.2 其它特征

植物基因簇除了具有共有特征外，还有一些不具有规律性的特征。例如，拟南芥、罂粟和百脉根中发现了基因簇附近存在相邻的转座子与还原转座子序列^[34–36,38]。某些基因簇位于端粒或邻近着丝粒的位置，这 2 个区域均为染色体高度动态变化的区域，基因重组、重排及异位的概率都较高，这有利于复制的基因进行分化从而促进基因簇的形成^[35,50]。还有一些基因簇，如百脉根中的合成生氰糖苷类化合物的基因簇，包含在次生代谢途径中没有功能的其它基因；而有些基因簇，如燕麦中燕麦根皂苷基因簇和拟南芥中 thalianol 基因簇，基因紧密排列，且基因之间没有其它基因的插入^[10,15,21]。马铃薯和番茄中的甾体生物碱基因簇，每个基因簇分成 2 个亚基因簇，位于不同的染色体上，但依然具有共调控的特性^[39]。

一些代谢基因簇可以合成多种产物。例如，百脉根中合成生氰糖苷类化合物的基因簇编码的酶可以使用不同的前体氨基酸作为起始底物，催化合成亚麻苦苷和/或百脉根苷^[15]。番茄中甾体生物碱基因簇的主要产物是 α-茄碱和 α-卡茄碱，这些代谢物只有糖基部分不同，表明对基因簇合成的骨架进行不同修饰可以形成不同的产物^[39]。

值得一提的是，基因簇中的大多数基因是协同转录且共表达的，这样使得代谢途径中的酶可以在特定组织和时间内具有协同催化的特点。然而，即使所有基因簇中的基因在某些高度特异的情况下共表达，但个别基因也会在其它情况下分别转录^[15,21]。如水稻中多功能的 phytocassane 合成基因簇，部分重叠的 2 个基因簇形成 1 个大基因簇，对专属于亚基因簇的基因却表现出不同于大基因簇的转录表达谱^[27,51]。番茄中 1 个基因簇也正逐渐分化成 2 个亚基因簇，它们的生化功能和转录表达谱都具有差异^[22]。由此可见，一个多功能基因簇中的基因并不完全服从于基因簇统一的共调控，这表明共调控机制不是促进基因簇形成的主要因素。但基因簇中所有基因都参与特殊的代谢途径，因此可以推测共同参与特殊代谢途径是促进基因簇形成的动力。与此假说一致的是，进化树分析表明水稻中 2 个二萜生物合成基因簇都经历了独立的进化过程^[25]，表明编码相关酶的基因形成基因簇的进化压力是持久而稳定的。

3 基因簇调控机制

目前，对基因簇调控机制的解析并不充分。水稻转录因子 TGAP1 调控 momilactone 和 phyto-

cassanes 的生物合成^[52]。然而, TGAP1 似乎是二萜类化合物的全局调控因子, 而不是特异性地调控这 2 个基因簇^[53]。CYP76M7 是参与水稻 phytocassanes 生物合成的基因, 其启动子可被稻瘟病诱导。因此, 可以考虑利用该启动子过表达水稻中的防御基因, 用于培养具有持续抗性的水稻品系^[54]。

另外, 玉米 DIMBOA 主要在幼嫩植株中具有较高的浓度, 可以起防御作用。最近研究发现了一个 3.9 kb 的顺式作用元件—DICE (distal ciselement), 位于 *Bx* 基因簇中 *Bx1* 基因上游大约 140 kb 处, 该区域属于减数分裂重组的热点区域, 可以促进 *Bx1* 基因的持续表达, 使 *Bx1* 在玉米各个发育阶段可以保持较高的转录水平, 促进 DIMBOA 的合成^[55]。

真菌中, 最初在烟曲霉菌 (*Aspergillus fumigatus*) 中发现了全局调控因子 LaeA, 已有报道表明 LaeA 可调控多种次生代谢产物的合成, 但目前植物中还未发现这样的全局调控因子。

除调控因子外, 染色质重塑也会调控基因簇的表达。组蛋白 2 变体 H2A.Z 对于拟南芥 thalianol 基因簇正常表达具有重要的调节作用, 它可以通过使染色质结构变得松散来促进基因的表达。通过突变使 H2A.Z 不能结合到核小体上导致基因簇中的基因表达量显著降低^[56]。ARP6 是 SWR1 染色质重塑复合体的组分, 是调节 H2A.Z 沉积进核小体的必要因子^[57,58]。*apr6* 突变体中, 基因簇区域核小体结构的稳固性高于野生型, 基因簇中基因的表达水平与野生型相比也显著降低。这说明基因簇的表达可以通过染色质重塑进行调节。通过研究基因簇的调控机制, 可为实现利用基因簇对植物中次生代谢途径的定向调控提供理论指导。

4 基因簇形成机制

植物中, 部分基因形成基因簇的直接原因与植物的生长发育相关, 大多数基因簇合成的代谢产物参与植物防御反应。基因簇形成的机制即使在近缘物种中也各不相同^[14], 基因簇在近缘物种(如 *Bx* 基因簇在玉米、黑麦、小麦中)或远缘物种中(如生氰糖苷途径在高粱、木薯、百脉根中)均是独立形成的, 说明至少这些基因簇的形成不是完全的偶然事件。事实上, 基因最初聚集在一起可能源于偶

然, 但基因簇的维持或扩张需要正向选择压力, 表明基因簇参与合成的代谢产物有利于植物促进植物的生长发育过程。

基因形成基因簇的优势主要包括: 促进基因功能改善、基因有益性状稳定遗传、基因簇中的基因共同被调控等。编码蛋白复合体的基因形成基因簇后可以完善蛋白复合体中各组分的最优比例, 使复合体更好地发挥功能^[6]。大部分植物基因簇合成的代谢产物可以防御病原体和食草动物, 这有助于植物抵御不良环境侵袭, 因此, 基因形成基因簇有利于促进植物有益性状的稳定遗传^[34,48,59]。另外, 对于完整代谢途径中的基因簇而言, 各基因成簇存在可以避免有毒中间化合物的积累。有关真菌基因簇的研究表明, 由成对相邻基因(基因在染色体上相邻, 它们的产物在生物合成途径上也相邻)合成的中间体中 40% 为有毒物质^[60]。基因簇的另一优势是通过不同调控因子和染色质重塑等调节作用, 实现基因簇中各基因的共表达和共调控, 从而使代谢途径的调控更具有协同性^[48]。

5 基因簇应用

细菌生物合成基因簇的发现为抗生素及其它生物活性产物的发现奠定了基础。而基因组测序技术使得挖掘微生物基因组中候选基因簇成为可能^[61], 并有利于发现新的代谢途径和代谢产物^[62]。与微生物相比, 大多数植物次生代谢基因簇是通过遗传学和生物化学方法发现并鉴定获得。目前, 先进的基因组测序技术以及完善的基因簇挖掘方法使得更多植物中的基因簇得以发掘和鉴定^[63,64]。基于基因组信息发现的候选基因簇, 有利于解析次生代谢途径的分子机制。

天然产物合成生物学是在基因组学研究的基础上, 解析天然产物的生物合成途径, 发掘途径上的酶基因等生物元件, 将生物元件导入底盘细胞中, 重构天然产物的生物合成途径, 将这些底盘细胞构建成可以制备天然产物的“细胞工厂”, 利用“细胞工厂”快速繁殖的特点, 制备出结构优化、产量丰富的天然产物^[65]。天然产物合成生物学的研究策略包括以下几个方面: 天然产物代谢途径的解析; 相关生物元件的发掘、设计与标准化; 底盘细胞设计与构建等^[66]。

在药用植物中, 天然产物的类型多种多样, 涉

及到的次生代谢途径也较复杂。目前,除了紫杉醇(Paclitaxel)^[67]、青蒿素(Artemisinin)^[68]、丹参酮(Tanshinone)^[69]、人参皂苷(Ginsenoside)^[70]等少数次生代谢途径上的基因解析较为清楚外,对大多数天然产物的生物合成途径的了解都比较匮乏。越来越多植物次生代谢基因簇的发现,加快了次生代谢途径的解析进程,也为合成生物学研究提供了丰富的可供选择和改造的生物学元件,推动了天然产物合成生物学研究的发展。同时,基于合成生物学的理念和技术,将植物次生代谢途径中所有酶基因导入特定的底盘细胞,有望实现天然产物定向、高效的异源合成。

研究表明,青蒿素的前体已在酵母菌株中通过基因工程的手段实现了工业化水平的合成^[71]。除酵母外,多种异源表达平台已被建立。例如,天蓝色链霉菌(*Streptomyces coelicolor*)及构巢曲霉菌(*Aspergillus nidulans*)可为来自放线菌(*Actinomycetes*)和丝状真菌的基因提供宿主^[72,73]。除了微生物系统,一些植物系统,如烟草属植物 *Nicotiana tabacum* L. 和 *N. benthamiana* Domin 也可以用作基因异源表达的宿主^[74],用于合成植物次生代谢产物。

6 小结与展望

植物次生代谢基因簇已成为植物生物学及天然产物合成生物学的研究热点,为解析复杂代谢途径的分子机制提供了崭新的研究视角和方法。这些基因簇既具有普遍特征,又可能具有独特性。通过对已有基因簇的分析和了解,有助于我们发现植物中新的次生代谢基因簇,并为利用合成生物学及代谢工程手段定向、高效合成次生代谢产物的研究奠定基础。

测序技术的飞速发展使越来越多的植物基因组信息得到解析,这为基因簇的高效挖掘和鉴定带来便利条件,本课题组通过对丹参(*Salvia miltiorrhiza* Bunge)基因组进行测序,预测到了 4 个可能参与二萜类化合物合成的基因簇,其中 *SmCPS1* 和 *SmCPS2* 基因簇可能参与丹参酮的合成^[75]。这些预测的基因簇里均包含 1 个用于合成产物骨架的二萜合酶基因和几个对骨架起修饰作用的 *P450* 基因^[75]。另外,解析基因簇表达调控的分子机制,探索推动代谢基因簇形成及维持的进化力量,为进

一步开发基因簇功能,并使之应用于相关次生代谢物的工业化生产提供了理论基础。基因簇还可应用于合成生物学及代谢工程研究中,使那些被广泛用作药品、芳香剂及调味品等的大量植物次生代谢产物得到工业化生产,解决天然产物的来源问题。对植物而言,有些具有生物活性的代谢产物在植物体内大量积累可能会对植物本身产生毒性,因此,可以利用次生代谢基因簇的表达调控机制优化合成途径,有效控制产物合成的种类和数量,减轻代谢产物对植物潜在的毒性。

参考文献:

- [1] Worthen DB. Streptomyces in nature and medicine: the antibiotic makers[J]. *Oxford Univ Pr*, 2008, 63(2): 273–274.
- [2] O'Connor SE. Engineering of secondary metabolism[J]. *Annu Rev Genet*, 2015, 49: 71–94.
- [3] Hurst LD, Pál C, Lercher MJ. The evolutionary dynamics of eukaryotic gene order[J]. *Nat Rev Genet*, 2004, 5(4): 299–310.
- [4] Wong S, Wolfe KH. Birth of a metabolic gene cluster in yeast by adaptive gene relocation[J]. *Nat Genet*, 2005, 37(7): 777–782.
- [5] Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, et al. Gene map of the extended human MHC[J]. *Nat Rev Genet*, 2004, 5(12): 889–899.
- [6] Rocha EPC. The organization of the bacterial genome[J]. *Annu Rev Genet*, 2008, 42(1): 211–233.
- [7] Blumenthal T. Trans-splicing and Operons in *C. elegans* [M/OL]. Wormbook; the Online Review of *C. elegans*, 2012, 20, 1–11. <http://www.wormbook.org>.
- [8] Richard GF, Kerrest A, Dujon B. Comparative genomics and molecular dynamics of DNA repeats in eukaryotes[J]. *Microbiol Mol Biol Rev*, 2008, 72(4): 686–727.
- [9] Frey M, Chomet P, Glawischnig E, Stettner C, Grün S, et al. Analysis of a chemical plant defense mechanism in grasses[J]. *Science*, 1997, 277(5326): 696–699.
- [10] Qi X, Bakht S, Leggett M, Maxwell C, Melton R, Osbourn A. A gene cluster for secondary metabolism in oat: implications for the evolution of metabolic diversity in plants[J]. *Proc Natl Acad Sci USA*, 2004, 101(21): 8233–8238.
- [11] Wilderman PR, Xu M, Jin Y, Coates RM, Peters RJ. Identification of syn-pimara-7, 15-diene synthase reveals functional clustering of terpene synthases involved in rice phytoalexin/allelochemical biosynthesis[J]. *Plant Physiol*, 2004, 135(4): 2098–105.
- [12] Jonczyk R, Schmidt H, Osterrieder A, Fiesselmann A,

- Schullehner K, *et al.* Elucidation of the final reactions of DIMBOA-glucoside biosynthesis in maize; characterization of Bx6 and Bx7 [J]. *Plant Physiol*, 2008, 146(3): 1053–1063.
- [13] von Rad U, Hüttel R, Lottspeich F, Gierl A, Frey M. Two glucosyltransferases are involved in detoxification of benzoxazinoids in maize[J]. *Plant J*, 2001, 28(6): 633–642.
- [14] Sue M, Nakamura C, Nomura T. Dispersed benzoxazinone gene cluster; molecular characterization and chromosomal localization of glucosyltransferase and glucosidase genes in wheat and rye[J]. *Plant Physiol*, 2011, 157(3): 985–997.
- [15] Takos AM, Knudsen C, Lai D, Kannangara R, Mikkelsen L, *et al.* Genomic clustering of cyanogenic glucoside biosynthetic genes aids their identification in *Lotus japonicus* and suggests the repeated evolution of this chemical defence pathway[J]. *Plant J*, 2011, 68(2): 273–286.
- [16] Koch BM, Sibbesen O, Halkier BA, Svendsen I, Møller BL. The primary sequence of cytochrome P450_{tyr}, the multifunctional N-hydroxylase catalyzing the conversion of L-tyrosine to p-hydroxyphenylacetaldehyde oxime in the biosynthesis the cyanogenic glucoside dhurrin in *Sorghum bicolor* (L.) Moench[J]. *Arch Biochem Biophys*, 1995, 323(1): 177–186.
- [17] Bak S, Nielsen HL, Halkier BA. The presence of CYP79 homologues in glucosinolate-producing plants shows evolutionary conservation of the enzymes in the conversion of amino acid to aldoxime in the biosynthesis of cyanogenic glucosides and glucosinolates[J]. *Plant Mol Biol*, 1998, 38(5): 725–734.
- [18] Jones PR, Møller BL, Hoj PB. The UDP-glucose: p-hydroxymandelonitrile-O-glucosyltransferase that catalyzes the last step in synthesis of the cyanogenic glucoside dhurrin in *Sorghum bicolor*; isolation, cloning, heterologous expression, and substrate specificity [J]. *J Biol Chem*, 1999, 274(50): 35483–35491.
- [19] Falara V, Akhtar TA, Nguyen TT, Spyropoulou EA, Bleeker PM, *et al.* The tomato terpene synthase gene family [J]. *Plant Physiol*, 2011, 157(2): 770–789.
- [20] Schilmiller AL, Schauvinhold I, Larson M, Xu R, Charbonneau AL, *et al.* Monoterpenes in the glandular trichomes of tomato are synthesized from a neryl diphosphate precursor rather than geranyl diphosphate[J]. *Proc Natl Acad Sci USA*, 2009, 106(26): 10865–10870.
- [21] Matsuba Y, Nguyen TT, Wiegert K, Falara V, Gonzales-Vigil E, *et al.* Evolution of a complex locus for terpene biosynthesis in *Solanum*[J]. *Plant Cell*, 2013, 25(6): 2022–2036.
- [22] Matsuba Y, Zi J, Jones AD, Peters RJ, Pichersky E. Biosynthesis of the diterpenoid lycosantalanol via neryl neryl diphosphate in *Solanum lycopersicum* [J]. *PLoS One*, 2015, 10(3): e0119302.
- [23] Shimura K, Okada A, Okada K, Jikumaru Y, Ko KW, *et al.* Identification of a biosynthetic gene cluster in rice for momilactones[J]. *J Biol Chem*, 2007, 282(47): 34013–34018.
- [24] Wang Q, Hillwig ML, Peters RJ. CYP99A3: functional identification of a diterpene oxidase from the momilactone biosynthetic gene cluster in rice[J]. *Plant J*, 2011, 65(1): 87–95.
- [25] Swaminathan S, Morrone D, Wang Q, Fulton DB, Peters RJ. CYP76M7 is an ent-cassadiene C11 ahydroxylase defining a second multifunctional diterpenoid biosynthetic gene cluster in rice[J]. *Plant Cell*, 2009, 21(10): 3315–3325.
- [26] Wu Y, Hillwig ML, Wang Q, Peters RJ. Parsing a multifunctional biosynthetic gene cluster from rice: Biochemical characterization of CYP71Z6&7 [J]. *FEBS Lett*, 2011, 585(21): 3446–3451.
- [27] Wang Q, Hillwig ML, Okada K, Yamazaki K, Wu Y, *et al.* Characterization of CYP76M5-8 indicates metabolic plasticity within a plant biosynthetic gene cluster [J]. *J Biol Chem*, 2012, 287(9): 6159–6168.
- [28] Papadopoulou K, Melton RE, Leggett M, Daniels MJ, Osbourn AE. Compromised disease resistance in saponin-deficient plants [J]. *Proc Natl Acad Sci USA*, 1999, 96(22): 12923–12928.
- [29] Corey EJ, Matsuda SP, Bartel B. Isolation of an *Arabidopsis thaliana* gene encoding cycloartenol synthase by functional expression in a yeast mutant lacking lanosterol synthase by the use of a chromatographic screen [J]. *Proc Natl Acad Sci USA*, 1993, 90(24): 11628–11632.
- [30] Qi X, Bakht S, Qin B, Leggett M, Hemmings A, *et al.* A different function for a member of an ancient and highly conserved cytochrome P450 family: from essential sterols to plant defense[J]. *Proc Natl Acad Sci USA*, 2006, 103(49): 18848–18853.
- [31] Mugford ST, Qi X, Bakht S, Hill L, Wegel E, *et al.* A serine carboxypeptidase-like acyltransferase is required for synthesis of antimicrobial compounds and disease resistance in oats [J]. *Plant Cell*, 2009, 21(8): 2473–2484.
- [32] Mugford ST, Louveau T, Melton R, Qi X, Bakht S, *et al.* Modularity of plant metabolic gene clusters: a trio of linked genes that are collectively required for acylation of triterpenes in oat [J]. *Plant Cell*, 2013, 25(3): 1078–1092.
- [33] Owatworakit A, Townsend B, Louveau T, Jenner H, Rejzek M, *et al.* Glycosyltransferases from oat (*Avena*) implicated in the acylation of avenacins [J]. *J Biol Chem*,

- 2013, 288(6): 3696–3704.
- [34] Field B, Osbourn AE. Metabolic diversification-independent assembly of operon-like gene clusters in different plants[J]. *Science*, 2008, 320(5875): 543–547.
- [35] Field B, Fiston-Lavier AS, Kemen A, Geisler K, Quesneville H, Osbourn AE. Formation of plant metabolic gene clusters within dynamic chromosomal regions[J]. *Proc Natl Acad Sci USA*, 2011, 108(38): 16116–16121.
- [36] Krokida A, Delis C, Geisler K, Garagounis C, Tsikou D, et al. A metabolic gene cluster in *Lotus japonicus* discloses novel enzyme functions and products in triterpene biosynthesis[J]. *New Phytol*, 2013, 200(3): 675–690.
- [37] Link S, Engelmann K, Meierhoff K, Westhoff P. The atypical short-chain dehydrogenases HCF173 and HCF244 are jointly involved in translational initiation of the psbA mRNA of *Arabidopsis*[J]. *Plant Physiol*, 2012, 160(4): 2202–2218.
- [38] Winzer T, Gazda V, He Z, Kaminski F, Kern M, et al. A *Papaver somniferum* 10-gene cluster for synthesis of the anticancer alkaloid noscapine[J]. *Science*, 2012, 336(6089): 1704–1708.
- [39] Itkin M, Heinig U, Tzfadia O, Bhide AJ, Shinde B, et al. Biosynthesis of antinutritional alkaloids in *solanaceous* crops is mediated by clustered genes[J]. *Science*, 2013, 341(6142): 175–179.
- [40] Itkin M, Rogachev I, Alkan N, Rosenberg T, Malitsky S, Masini L, Meir S, Iijima Y, Aoki K, de Vos R, Prusky D, Burdman S, Beekwilder J, Aharoni A. Glycoalkaloid metabolism1 is required for steroidal alkaloid glycosylation and prevention of phytotoxicity in tomato[J]. *Plant Cell*, 2011, 23(12): 4507–4525.
- [41] Kramer V, Koziel M. Structure of a maize tryptophan synthase a subunit gene with pith enhanced expression[J]. *Plant Mol Biol*, 1995, 27(6): 1183–1188.
- [42] Dutarte L, Hilliou F, Feyereisen R. Phylogenomics of the benzoxazinoid biosynthetic pathway of Poaceae: gene duplications and origin of the Bx cluster[J]. *BMC Evol Biol*, 2012, 12: 64–82.
- [43] Grün S, Frey M, Gierl A. Evolution of the indole alkaloid biosynthesis in the genus *Hordeum*; distribution of gramine and DIBOA and isolation of the benzoxazinoid biosynthesis genes from *Hordeum lechleri*[J]. *Phytochemistry*, 2005, 66(11): 1264–1272.
- [44] Niemeyer HM. Hydroxamic acids 4-hydroxy-1,4-benzoxazin-3-ones, defence chemicals in the gramineae[J]. *Phytochemistry*, 1988, 27(11): 3349–3358.
- [45] Liby KT, Yore MM, Sporn MB. Triterpenoids and rexinoids as multifunctional agents for the prevention and treatment of cancer[J]. *Nat Rev Cancer*, 2007, 7(5): 357–369.
- [46] Delis C, Krokida A, Georgiou S, Peña-Rodríguez LM, Kavroulakis N, Ioannou E, Roussis V, Osbourn AE, Papadopoulou KK. Role of lupeol synthase in *Lotus japonicus* nodule formation[J]. *New Phytol*, 2011, 189(1): 335–346.
- [47] Ye K, Ke Y, Keshava N, Shanks J, Kapp JA, Tekmal RR, Petros J, Joshi HC. Opium alkaloid noscapine is an antitumor agent that arrests metaphase and induces apoptosis in dividing cells[J]. *Proc Natl Acad Sci USA*, 1998, 95(4): 1601–1606.
- [48] Osbourn A. Secondary metabolic gene clusters: evolutionary toolkits for chemical innovation[J]. *Trends Genet*, 2010, 26(10): 449–457.
- [49] Chu HY, Wegel E, Osbourn A. From hormones to secondary metabolism: the emergence of metabolic gene clusters in plants[J]. *Plant J*, 2011, 66(1): 66–79.
- [50] Field B, Osbourn A. Order in the playground: formation of plant gene clusters in dynamic chromosomal regions[J]. *Mob Genet Element*, 2012, 2(1): 46–50.
- [51] Okada A, Shimizu T, Okada K, Kuzuyama T, Koga J, Shibuya N, Nojiri H, Yamane H. Elicitor induced activation of the methylerythritol phosphate pathway toward phytoalexins biosynthesis in rice[J]. *Plant Mol Biol*, 2007, 65(1–2): 177–187.
- [52] Okada A, Okada K, Miyamoto K, Koga J, Shibuya N, Nojiri H, Yamane H. OsTGAP1, a bZIP transcription factor, coordinately regulates the inductive production of diterpenoid phytoalexins in rice[J]. *J Biol Chem*, 2009, 284(39): 26510–26518.
- [53] Yamane H. Biosynthesis of phytoalexins and regulatory mechanisms of it in rice[J]. *Biosci Biotechnol Biochem*, 2013, 77(6): 1141–1148.
- [54] Vijayan J, Devanna BN, Singh NK, Sharma TR. Cloning and functional validation of early inducible *Magnaporthe oryzae* responsive CYP76M7 promoter from rice[J]. *Front Plant Sci*, 2015, 6: 371.
- [55] Zheng L, McMullen MD, Bauer E, Schön CC, Gierl A, Frey M. Prolonged expression of the *BX1* signature enzyme is associated with a recombination hotspot in the benzoxazinoid gene cluster in *Zea mays*[J]. *J Exp Bot*, 2015, 66(13): 3917–3930.
- [56] Nützmann HW, Osbourn A. Regulation of metabolic gene clusters in *Arabidopsis thaliana*[J]. *New Phytol*, 2015, 205(2): 503–510.
- [57] Deal RB, Topp CN, McKinney EC, Meagher RB. Repression of flowering in *Arabidopsis* requires activation of *Flowering Locus C* expression by the histone variant H2A.Z[J]. *Plant Cell*, 2007, 19(1): 74–83.
- [58] Kumar SV, Wigge PA. H2A.Z-containing nucleosomes mediate the thermosensory response in *Arabidopsis*[J]. *Cell*, 2010, 140(1): 136–147.

- [59] Takos AM, Rook F. Why biosynthetic genes for chemical defense compounds cluster[J]. *Trends Plant Sci*, 2012, 17(7): 383–388.
- [60] McGary KL, Slot JC, Rokas A. Physical linkage of metabolic genes in fungi is an adaptation against the accumulation of toxic intermediate compounds[J]. *Proc Natl Acad Sci USA*, 2013, 110(28): 11481–11486.
- [61] Winter JM, Behnken S, Hertweck C. Genomics-inspired discovery of natural products[J]. *Curr Opin Chem Biol*, 2011, 15(1): 22–31.
- [62] Feng Z, Kallifidas D, Brady SF. Functional analysis of environmental DNA-derived type II polyketide synthases reveals structurally diverse secondary metabolites[J]. *Proc Natl Acad Sci USA*, 2011, 108(31): 12629–12634.
- [63] Castillo DA, Kolesnikova MD, Matsuda SP. An effective strategy for exploring unknown metabolic pathways by genome mining[J]. *J Am Chem Soc*, 2013, 135(15): 5885–5894.
- [64] Chen SL, Song JY, Sun C, Xu J, Zhu YJ, Verpoorte R, Fan TP. Herbal genomics: Examining the biology of traditional medicines[J]. *Science*, 2015, 347(S6219): 27–29.
- [65] Benner SA, Sismour AM. Synthetic biology[J]. *Nat Rev Genet*, 2005, 6(7): 533–43.
- [66] 陈士林, 朱孝轩, 李春芳, 王勇, 姚辉, 孙超, 宋经元. 中药基因组学与合成生物学[J]. *药学报*, 2012, 47(8): 1070–1078.
- Chen SL, Zhu XX, Li CF, Wang Y, Yao H, Sun C, Song JY. Genomics and synthetic biology of traditional Chinese medicine[J]. *Acta Pharmaceutica Sinica*, 2012, 47(8): 1070–1078.
- [67] Guerra-Bubb J, Croteau R, Williams RM. The early stages of taxol biosynthesis: an interim report on the synthesis and identification of early pathway metabolites[J]. *Nat Prod Rep*, 2012, 29: 683–696.
- [68] Dietrich JA, Yoshikuni Y, Fisher KJ, Woolard FX, Ockey D, McPhee DJ, Renninger NS, Chang MC, Baker D, Keasling JD. A novel semi-biosynthetic route for artemisinin production using engineered substrate-promiscuous P450 (BM3)[J]. *ACS Chem Biol*, 2009, 4: 261–267.
- [69] Zhou YJ, Gao W, Rong Q, Jin G, Chu H, Liu W, Yang W, Zhu Z, Li G, Zhu G, Huang L, Zhao ZK. Modular pathway engineering of diterpenoid synthases and the mevalonic acid pathway for miltiradiene production[J]. *J Am Chem Soc*, 2012, 134(6): 3234–3241.
- [70] Dai Z, Liu Y, Zhang X, Shi M, Wang B, Wang D, Huang L, Zhang X. Metabolic engineering of *Saccharomyces cerevisiae* for production of ginsenosides[J]. *Metab Eng*, 2013, 20: 146–156.
- [71] Siddiqui MS, Thodey K, Trenchard I, Smolke CD. Advancing secondary metabolite biosynthesis in yeast with synthetic biology tools[J]. *Fems Yeast Res*, 2012, 12(2): 144–170.
- [72] Gomez-Escribano JP, Bibb MJ. Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters[J]. *Microb Biotechnol*, 2011, 4(2): 207–215.
- [73] Chiang YM, Oakley CE, Ahuja M, Entwistle R, Schultz A, et al. An efficient system for heterologous expression of secondary metabolite genes in *Aspergillus nidulans*[J]. *J Am Chem Soc*, 2013, 135(20): 7720–7731.
- [74] Kumar S, Hahn FM, Baidoo E, Kahlon TS, Wood DF, et al. Remodeling the isoprenoid pathway in tobacco by expressing the cytoplasmic mevalonate pathway in chloroplasts[J]. *Metab Eng*, 2012, 14(1): 19–28.
- [75] Xu HB, Song JY, Luo HM, Zhang YJ, Li QS, et al. Analysis of the genome sequence of the medicinal plant *Salvia miltiorrhiza*[J]. *Mol Plant*, 2016, 9(6): 949–952.

(责任编辑:周媛)