西红花苷生源途径解析及其合成生物学研究进展

浦香东¹, 徐志超^{1,2}, 高冉冉¹, 宋经元^{1,2,3*}

1. 中国医学科学院北京协和医学院药用植物研究所, 国家中医药管理局中药资源保护重点研究室, 北京 100193;

2. 中药资源教育部工程研究中心, 北京 100193;

3. 中国医学科学院北京协和医学院药用植物研究所云南分所, 景洪 666100

* 联系人, E-mail: jysong@imlad.ac.cn

2020-07-13 收稿, 2020-08-14 修回, 2020-08-17 接受, 2020-08-17 网络版发表

国家重点研发计划(2019YFC1711100)和中国医学科学院本草基因组协同创新团队(2016-I2M-3-016)资助

摘要 天然产物是新药研发的重要资源, 具有显著的抗疟疾、抗肿瘤、抗病毒感染等作用. 珍稀名贵中药西红花的活性成分为西红花苷, 属于脱辅基类胡萝卜素类化合物, 具有保护中枢神经系统及预防心脑血管疾病的药理活性. 西红花苷资源匮乏、价格高昂, 极大限制其药物开发及临床应用. 本文从西红花苷的化学表征、生物合成途径解析及合成生物学研究进行综述, 展示了西红花苷的光谱学和质谱学特征, 阐述其完整生物合成途径的解析, 介绍了利用微生物细胞工厂生产 5 种西红花苷的最新突破. 据此, 进一步提出开展不同物种西红花苷生物合成途径关键酶的进化机制、合成生物学关键技术探索和条件优化等研究, 以推动西红花苷工业化生产及新药研发, 为西红花苷来源的中药资源可持续发展提供理论基础和方法指导.

关键词 西红花苷, 化学表征, 生物合成途径, 合成生物学, 西红花, 栀子

西红花苷为脱辅基类胡萝卜素类化合物, 是名贵中药西红花的主要活性成分. 西红花具有活血化瘀、凉血解毒、解郁安神的功效^[1], 也被用做食品着色剂, 是世界上最昂贵的香料, 零售价高达 2000~7000 €/kg^[2-4], 有“红色黄金”的美誉^[2,3]. 现代药理研究表明, 西红花苷在抗肿瘤^[5-7]、抗炎^[8-10]、抗氧化^[6]及治疗神经退行性疾病, 如阿尔兹海默症^[11-13]、抑郁症^[14]、失眠^[15]、学习记忆和认知障碍^[16]等方面药理活性显著, 同时还具有保肝利胆、抗糖尿病等作用^[17]. 因此, 其药用价值备受国际关注. 然而, 西红花来源于鸢尾科番红花(*Crocus sativus*)的干燥柱头, 面临资源短缺、采收困难等难题, 大约 110000~170000 朵盛开的番红花才能获得 1 kg 西红花药材. 目前, 西红花苷化合物主要依赖于植物化学提取分离, 面临提取工艺复杂、提纯成本

高、污染严重等问题, 导致西红花苷化合物极其昂贵, 严重限制西红花苷药物的开发及临床使用.

本草基因组学的提出及发展^[18-20]大大促进了药用植物次生代谢途径的解析, 综合利用生物信息学技术和分子生物学手段阐释中药特定次生代谢途径的关键酶编码基因, 为药用天然产物合成生物学研究提供理论依据. 合成生物学是继 DNA 双螺旋结构和基因组测序技术发现及发明后的“第三次生物科学革命”^[21-26], 已在青蒿酸^[27-29]、紫杉醇^[30]等珍稀天然产物生物合成中取得突破性成就, 为中药资源短缺问题的解决提供新途径. 西红花苷生源途径复杂, 涉及类胡萝卜素的合成及其裂解氧化、糖基化修饰等, 其生物合成途径的解析及合成生物学研究受到国内外研究学者的广泛关注^[2,3]. 本文就西红花苷的化学表征、生物合成途径解

引用格式: 浦香东, 徐志超, 高冉冉, 等. 西红花苷生源途径解析及其合成生物学研究进展. 科学通报, 2021, 66: 219-232

Pu X D, Xu Z C, Gao R R, et al. Research progress on biosynthesis and synthetic biology of crocin (in Chinese). Chin Sci Bull, 2021, 66: 219-232, doi: 10.1360/TB-2020-0875

析及合成生物学研究等进行综述,为解决西红花苷资源短缺问题提供参考,推动西红花苷药物开发及应用。

1 西红花苷的化学表征

西红花苷主要在番红花柱头积累,根据其所连接的糖基数目和位置差异,主要分为西红花苷I、西红花苷II、西红花苷III、西红花苷IV和西红花苷V^[4,31,32](图1),极性大小顺序为:西红花苷I>西红花苷II>西红花苷III>西红花苷IV>西红花苷V^[33]。西红花苷分离提取所采用的技术均为常用天然产物分离纯化方法,如柱层析法(SGCC)、制备液相色谱法(PLC)、制备薄层色谱法(PTLC)、高速逆流色谱法(HSCCC)、高速离心色谱法(FCPC)等^[34],但尚无研究报道从植物中提取分离得到西红花苷III,这可能与西红花苷III在植物中的含量低有关。西红花苷分子结构中存在多个共轭双键,常呈现出橙红色,在440~450 nm处有最大吸光值^[35,36]。5种类型西红花苷紫外-可见光谱学特征研究报道显示,在不同检测溶液中,西红花苷的紫外-可见吸收光谱存在差异,但总体影响不大(表1)。近红外光谱及拉曼光谱也被用于西红花苷的光谱学研究^[45,46]。通过收集西红花苷化

合物的光谱学特征信息可以初步确证其结构类型。

西红花苷质谱学研究主要采用电喷雾离子源质谱(electron spray ionization, ESI),通过对Pittenauer等人^[47]及Pu等人^[48]研究中报道的5种西红花苷质谱信息研究发现,在正离子模式下5种类型的西红花苷一般不出现 $[M+H]^+$ 准分子离子峰,而出现 $[M+Na]^+$ 离子峰,质谱裂解规律见表2。两项研究所呈现出的西红花苷裂解规律结果相似,具有很高的参考价值,将大大提高该类成分的结构解析效率。此外,核磁共振波谱法^[33,36,37,41,42,49]以及毛细管电泳法^[50]也被用作西红花苷化学结构的表征研究。

在光照、高温、极端pH条件下,类胡萝卜素类化合物十分不稳定,容易发生降解^[51,52]。西红花苷属于脱辅基类胡萝卜素类化合物,在实际应用中,其稳定性也备受关注。Chen等人^[53,54]通过高温、极端pH、光照等条件下,对西红花苷I、西红花苷II、西红花苷IV和西红花苷V的稳定性进行研究,结果显示,西红花苷分子中的葡萄糖基不利于其结构稳定,随着糖基数目的增加,降解半衰期逐渐缩短。Ketmaro等人^[35]通过在不同酸碱性、不同温度、光照、是否含有氧化剂等条件

图1 (网络版彩色)西红花苷结构类型及其在两种药用植物中的资源分布。(a) 西红花柱头;(b) 栀子果实;(c) 5种西红花苷的结构类型
Figure 1 (Color online) The structure of crocins and its resource distribution in two medicinal plants. (a) *Crocus sativus* stigmas; (b) *Gardenia jasminoides* fruits; (c) the structures of five types of crocins

表 1 西红花苷紫外-可见光谱学特征

Table 1 The UV-VIS spectroscopy characterization of crocins

检测溶剂	光谱学特征($\lambda_{\max}$, nm)	文献
50%乙醇	I(440.0, 465.5或439.5, 464.5) II(440.0, 465.5或439.5, 464.0) III(438.0, 461.0) IV(433.5, 456.0, 326.5或431.0, 453.0) V(431.0, 453.0)	[36]
50%乙醇(EDTA)	I(440, 464); II(440, 464); IV(434, 458); V(432, 456)	[37]
50%乙醇	III(439, 462); V(433, 456)	[38]
33%甲醇	I(440.3, 463.8); II(439.6, 463.2); III(438.9, 462.3); IV(434.3, 458.0); V(433.0, 457.1)	[39]
50%甲醇	I(437, 460); II(434, 455); III(433, 457); IV(439, 462)	[40]
90%甲醇	I(437, 461); IV(430, 452)	[41]
甲醇	I(433, 458)	[42]
超纯水	I(442, 465); II(442, 465); IV(434, 457); V(433, 458)	[43]
50 mmol/L Hepes/NaOH(pH 7.0)	I(442, 468); II(442, 466); III(434, 460); IV(438, 462); V(434, 458)	[44]

表 2 西红花苷的质谱学特征

Table 2 The mass spectrometry information of crocins

离子源	离子模式 ^{a)}	检测溶剂	化合物										文献
			I		II		III		IV		V		
			准分子离子峰 (<i>m/z</i>)	特征性碎片离子峰 (<i>m/z</i>)	准分子离子峰 (<i>m/z</i>)	特征性碎片离子峰 (<i>m/z</i>)	准分子离子峰 (<i>m/z</i>)	特征性碎片离子峰 (<i>m/z</i>)	准分子离子峰 (<i>m/z</i>)	特征性碎片离子峰 (<i>m/z</i>)	准分子离子峰 (<i>m/z</i>)	特征性碎片离子峰 (<i>m/z</i>)	
ESI ^{b)}	+	甲醇	[M+Na] ⁺ : 999.3	907.2, 675.2, 583.1, 347.0, 259.0	[M+Na] ⁺ : 837.2	745.2, 675.2, 583.2, 513.2, 347.0, 259.0	[M+Na] ⁺ : 675.3	583.2, 513.2, 421.1, 259.0	[M+Na] ⁺ : 675.3	583.2, 365.0, 347.0, 259.0	[M+Na] ⁺ : 513.2	421.1, 351.2, 307.1, 259.0, 201.1	[47]
			[M+Na] ⁺ : 999.3686	927.6565, 814.5748, 752.3218, 675.2631, 594.3197, 530.2975, 421.1703, 347.0955, 208.1030, 185.0560	[M+Na] ⁺ : 837.3142	745.2533, 675.2621, 583.1991, 513.2089, 421.1475, 351.1564, 347.0949, 259.0848, 185.0414, 119.0856	[M+Na] ⁺ : 675.2611	584.1980, 513.2087, 421.1470, 351.1557, 259.0933, 185.0411, 119.0643	[M+Na] ⁺ : 675.2614	631.2709, 583.1994, 411.1935, 347.0945, 307.1679, 259.0803	[M+Na] ⁺ : 513.2085	421.1463, 351.1559, 259.0939, 185.0412, 119.0850	[48]

a) +, 表示正离子模式; b) ESI(electron spray ionization), 表示电喷雾离子源

下对栀子(*Gardenia jasminoides*)色泽变化进行研究发
现, 西红花苷在各种条件下的稳定性顺序为: 碱性条件
>4°C水溶液>30°C水溶液>72°C水溶液>酸性条件>过
氧化氢溶液>光照。西红花苷的稳定性研究为其提取分
离纯化提供借鉴, 同时西红花苷稳定性研究提示单体
西红花苷的药理活性可能存在着很大差异, 为西红花

苷药物开发及利用提供理论指导。

2 西红花苷生物合成途径

西红花苷生物合成途径包括三部分: 上游部分经
由2-甲基-D-赤藻糖醇-4-磷酸途径(MEP途径)产生牻牛
儿基牻牛儿基焦磷酸(GGPP)^[3]; 中游部分是类胡萝卜

素代谢途径,即GGPP经过一系列的聚合、脱氢及环化作用生成类胡萝卜素类化合物,如玉米黄质等^[55,56],下游部分是类胡萝卜素的降解、氧化及糖基化反应途径,其关键酶涉及类胡萝卜素裂解双加氧酶(carotenoid cleavage dioxygenases, CCDs)、乙醛脱氢酶(aldehyde dehydrogenases, ALDHs)和糖基转移酶(UDP-glucosyl-transferases, UGTs)(图2)。类胡萝卜素类化合物在CCDs对称裂解催化作用下脱辅基生成西红花酸二醛;西红花酸二醛经ALDHs催化生成西红花酸;西红花酸在一系列UGTs的作用下生成不同类型西红花苷产物(I~V)。尽管这些关键酶的种类明确,但由于其皆属于多基因家族,在西红花苷生物合成途径中行使功能的基因研究仍充满挑战^[2]。随着测序技术的进步,结合多组学及分子生物学研究手段^[19,20],西红花苷生物合成途径解析研究已取得显著进展,将极大促进西红花苷合成生物学研究。

2.1 类胡萝卜素裂解双加氧酶(CCDs)裂解C₄₀类胡萝卜素前体

CCDs是催化类胡萝卜素类化合物裂解的关键酶,

根据其反应底物、反应位点及反应产物的差异,主要分为CCD1、CCD4、CCD7和CCD8四个亚家族,各亚家族成员之间基因结构差异较大、序列相似度很低。CCD1亚家族裂解产生挥发性萜类化合物,在植物香气形成方面起重要作用;CCD4亚家族在植物颜色形成中具有重要价值;CCD7、CCD8参与植物激素独角金内酯的合成,对于植物侧根、侧芽的萌发具有重要作用^[57-60]。

Frusciante等人^[2]通过对不同发育时期番红花柱头转录组测序分析及其中西红花苷、西红花酸含量测定分析,筛选得到关键酶CsCCD2(GenBank: KJ541749)。进化研究发现,在所有CCD亚家族中,CsCCD2与CCD1亚家族具有更近的亲缘关系,但又与CCD1亚家族成员的结构及功能差异较大。进一步通过原核表达体系及玉米胚乳瞬时转化表达体系对其功能进行研究,结果显示,CsCCD2具有裂解玉米黄质生成西红花酸二醛的功能,被认为是西红花苷生物途径中的第一个关键酶基因。Ahrazem等人^[61]从番红花同属植物黄金番红花(*Crocus ancyrensis*)中克隆得到4个CCDs基因,其中包括一个与CsCCD2氨基酸序列具有87%相似度的CaCCD2(GenBank: KP792756)。原核表达体系及水稻

图2 (网络版彩色)西红花苷生物合成途径
Figure 2 (Color online) The biosynthetic pathway of crocins

愈伤组织瞬时转化体系研究发现, CaCCD2与CsCCD2功能一致, 具有裂解玉米黄质7,8(7',8')碳碳双键的活性。此外, 其研究团队通过对大叶醉鱼草(*Buddleja davidii*)的BdCCD4.1/4.3(GenBank: KX374547/KX374549)体外功能验证研究发现其具有裂解玉米黄质7,8(7',8')碳碳双键的功能^[62]。Xu等人^[39]基于梔子基因组、转录组等多组学分析, 挖掘得到关键酶GjCCD4a(GenBank: ARU08109.1), 原核表达及体外功能验证研究结果显示, GjCCD4a具有裂解番茄红素、 β -胡萝卜素和玉米黄质7,8(7',8')碳碳双键的功能, 参与西红花酸二醛的产生; 此外, 基于比较基因组分析, 研究人员清晰阐释茜草科咖啡及梔子分别产生咖啡因和西红花苷的内在分子机制, 揭示梔子西红花苷代谢关键酶基因*GjCCD4a*由基因串联复制产生, 为基因串联复制在植物特定代谢进化中的作用提供参考, 揭示活性成分趋异进化机制^[39]。

植物是自然界最杰出的“化学家”之一, 与动物不同, 它们无法通过移动逃避不利环境, 而是靠自身合成特殊化合物抵御生物或非生物胁迫, 但它们如何获得这类特殊化合物呢? 此处提到的3个不同科属植物中的CCDs都具有催化玉米黄质7,8(7',8')碳碳双键裂解的功能, 其催化机制及进化关系有待深入探索, 并将为不同植物间的趋同进化机制研究提供参考。

2.2 乙醛脱氢酶(ALDHs)催化形成西红花酸

ALDHs是一个氧化还原酶超家族, 可将醛类化合物氧化为对应的羧酸类化合物, 主要参与植物生长发育与应激反应^[63-66]。例如, ZmALDH2B2在玉米(*Zea mays*)花药发育及雄性育性方面发挥着重要作用^[67,68]; OsALDH7对于水稻(*Oryza sativa*)种子成熟及生存能力至关重要^[63]; 过表达*AtALDH3I1*和*AtALDH7B4*能够提高拟南芥(*Arabidopsis thaliana*)对于干旱胁迫、盐胁迫及氧化胁迫的应激能力等^[64]。在西红花苷生物合成途径中, ALDHs主要负责将西红花酸二醛氧化为西红花酸, 从而为西红花苷的合成提供直接前体物质。

基于已公开发表的番红花柱头蛋白质组数据, Gómez-Gómez等人^[69]克隆得到5个CsALDHs关键酶基因, 由于其中有2个基因相似度极高, 故只对其中4个基因进行功能验证, 它们分别是CsALDH54788(GenBank: KU577904)、CsALDH3898(GenBank: KU577905)、CsALDH20158(GenBank: KU577906)、CsALDH11367(GenBank: KU577907)。功能验证研究结果显示, 它们

皆具有催化西红花酸二醛生成西红花酸的功能。蛋白同源性及底物亲和力分析结果提示, ALDH2C4亚家族成员对于催化西红花酸二醛生成西红花酸具有更高的特异性^[70]。基于不同发育时期番红花柱头转录组数据^[2]及基因同源性分析, Demurtas等人^[38]筛选得到6个CsALDHs基因, 通过原核表达体系及体外功能验证研究发现, 只有CsALDH3I1(GenBank: MF596165)能够催化西红花酸二醛产生西红花酸。Tan等人^[71]通过对不同发育时期番红花柱头转录组测序分析, 筛选得到3个CsALDHs基因, 利用产西红花酸二醛的酵母细胞为底盘, 将CsALDHs基因整合入其中进行功能验证研究。结果显示, CsALDH3具有氧化西红花酸二醛至西红花酸的功能。Xu等人^[39]基于梔子多组学分析及梔子不同组织部位西红花苷含量分析, 共筛选到19个*GjALDHs*基因, 结合基因表达及其与关键酶基因*GjCCD4a*的共表达分析, 锚定关键酶GjALDH2C3, 体外功能验证研究发现, GjALDH2C3(GenBank: KY631926.1)具有催化西红花酸二醛转化至西红花酸的功能。

现已鉴定得到多个具有催化西红花酸二醛生成西红花酸活性的ALDHs蛋白, 经氨基酸序列比对分析后发现, CsALDH3I1和CsALDH54788氨基酸序列相似度高达98.8%, CsALDH20158和CsALDH11367氨基酸序列相似度为84.9%, 其余各序列之间相似度较低(图3(a), (b))。为了筛选出高效催化基因元器件应用于后续的西红花苷合成生物学研究, 有必要对它们的催化特性展开深入研究。

2.3 糖基转移酶(UGTs)催化西红花酸生成西红花苷

糖基化反应是生物系统中最常见及最重要的反应类型之一, 它可以提高亲脂性化合物的亲水性, 从而改善化合物的药代动力学特征^[72,73]。糖基转移酶是负责糖基化反应的重要生物催化剂, 通过催化糖基供体的糖基转移至糖基受体中, 实现糖基复合物的合成^[74,75]。在西红花苷生物合成途径中, 糖基转移酶主要负责西红花酸的糖基化, 通过将糖基供体尿苷二磷酸葡萄糖(uridine diphosphate glucose, UDPG)中的葡萄糖基转移至西红花酸, 进而生成西红花苷。

Dufresne等人^[76]利用西红花球茎诱导产生的愈伤组织提取粗酶, 并于体外进行西红花酸催化反应, 研究结果显示, 粗酶可催化产生西红花苷。结合不同类型西红花苷的合成动力学规律, 研究人员推测, 西红花酸至

图3 (网络版彩色)西红花苷生物合成关键酶氨基酸序列比对分析。(a) ALDH氨基酸序列比对分析; (b) ALDH氨基酸序列相似度分析; (c) UGT氨基酸序列比对分析; (d) UGT氨基酸序列相似度分析

Figure 3 (Color online) Amino acid sequence analysis of key enzymes in crocin biosynthetic pathway. (a) Amino acid sequence analysis of ALDH; (b) amino acid sequence identity analysis of ALDH; (c) amino acid sequence analysis of UGT; (d) amino acid sequence identity analysis of UGT

西红花苷的生物合成途径需要两个不同糖基转移酶的催化。此后,其研究团队进一步结合蛋白电泳及蛋白分离纯化技术对此推测进行验证,确证该生物合成途径至少需要两个不同的糖基转移酶参与^[77]。Moraga等人^[44]根据糖基转移酶家族所具有的保守序列,从番红花柱头中克隆得到两个糖基转移酶基因,原核表达及体外功能验证发现,CsUGT2具有催化西红花酸、西红花苷V和西红花苷IV的功能,但其催化产物皆不是常见西红花苷类型,而是西红花酸和至少9个葡萄糖基结合的化合物。Nagatoshi等人^[33]基于对栀子细胞悬浮培养及生物转化作用研究,克隆得到15个西红花苷生物合成关键糖基转移酶基因,原核表达及体外功能验证研究发现,GjUGT75L6(GenBank: AB733668)和GjUGT94E5(GenBank: AB733669)分别具有西红花苷初级糖基化(在底物羧基上加葡萄糖基)及次级糖基化(在底物葡萄糖基上继续加入糖基)的功能,但GjUGT75L6和GjUGT94E5的表达模式与西红花苷在栀子中的积累模式差异显著,这意味着它们可能并非栀子西红花苷生物合成途径中的关键酶。Wang等人^[78]重复这两个基因的体外功能验证研究发现,GjUGT75L6对西红花酸具有微弱的催化活力,而GjUGT94E5不具

有催化功能。基于已报道的关键酶CsUGT2^[44]以及番红花柱头的转录组数据^[2,79]。Demurtas等人^[38]筛选得到一个关键酶CsUGT74AD1(GenBank: MF596166),原核表达体系及体外功能验证研究发现,其具有催化西红花酸产生西红花苷III和西红花苷V的功能。基于多组学研究及栀子不同组织部位西红花苷含量分析,Xu等人^[39]在栀子中共筛选到173个GjUGTs基因,结合基因表达及其与西红花苷代谢途径上游关键酶基因GjCCD4a的共表达分析,最终筛选得到关键酶GjUGT74F8(GenBank: MN944054)和GjUGT94E13(GenBank: MN944055)。功能验证研究显示,这两个糖基转移酶能够实现西红花酸至5种西红花苷的转化,且它们的表达模式与栀子中西红花苷的积累模式一致。

西红花苷生物合成关键糖基转移酶从最初的基原植物细胞培养催化反应研究到结合组学的基原植物糖基转移酶筛选研究,在国内外已持续20余年,受到广泛关注。近期,基于本草基因组学的指导及分子生物学的技术支撑,药用植物西红花苷生物合成途径被清晰阐释,筛选得到的关键糖基转移酶(CsUGT74AD1、GjUGT74F8和GjUGT94E13)能够实现5种西红花苷体外生产。这一突破为西红花苷合成生物学研究提供重

要基因元器件(图3(c), (d)), 为解决西红花苷资源匮乏、价格昂贵等问题奠定基础。

3 西红花酸及西红花苷合成生物学研究

西红花苷化学提取需要消耗大量有机化学试剂, 分离纯化耗时耗力, 随着日益增长的用药需求, 逐渐出现资源匮乏、供不应求等问题, 不利于西红花苷资源的可持续发展。此外, 目前仅能实现西红花苷III的化学合成, 其具有合成路线长(主要包括2,7-二甲基-2,4,6-辛三烯-1,8-二醛的合成、C5-磷酸酯的合成、西红花酸二甲酯的合成、西红花酸二醛的合成、西红花酸的合成、西红花苷III的合成等)、操作繁琐、反应收率低等问题^[80]。因此, 西红花苷合成生物学研究受到国内外研究学者的广泛关注。通过西红花苷生物合成途径及其合成生物学研究或将大大改善其化学提取及合成所带来的危害及污染, 实现西红花苷的绿色生产。另一方面, 还可

实现基原植物中含量较低西红花苷的合成及生产。

合成生物学诞生于21世纪, 发展迅猛, 是生物学研究的一个重要分支^[26,81,82]。合成生物学主要通过DNA合成、基因组编辑、基因元器件标准化、元器件组装设计以及生物工程等手段实现, 常用于改造的模式微生物主要为酵母和大肠杆菌^[81,83]。天然产物主要源于自然界的馈赠, 其中不乏具有宝贵药用价值的化合物, 是药物研发的重要宝库^[84]。合成生物学的产生及应用能够大大缓解天然产品价格昂贵、资源匮乏等问题^[48,85], 如青蒿素^[27-29]、依托泊苷^[86]、阿片类药物^[87,88]、大麻素^[89]、灯盏花素^[90]、人参皂苷^[91,92]等(图4)。

3.1 西红花酸合成生物学研究

西红花酸为西红花苷合成的直接前体物质, 其合成生物学研究将为西红花苷的生产奠定基础。Chai等人^[93]通过在产 β -胡萝卜素的酿酒酵母(*Saccharomyces*

图4 (网络版彩色)重要天然产物合成生物学研究

Figure 4 (Color online) The synthetic biology of important natural products

cerevisiae)菌株中引入*PsCrtZ*(*Pantoea stewartii*)和*CsCCD2*基因,得到西红花酸产量为351 µg/L的酵母工程菌.之后,又通过引入*ScALD*基因,提升西红花酸产量至1219 µg/L.最后,通过发酵培养进一步提升西红花酸产量至6278 µg/L. Wang等人^[78]将关键酶*CsCCD2*和*NcALD8*(*Neurospora crassa*)整合入大肠杆菌(*Escherichia coli*),通过优化得到西红花酸产量为4.42 mg/L的工程化大肠杆菌菌株. Tan等人^[71]通过整合*CsBCH*、*CsCCD2*和*CsALDH3*基因至产β-胡萝卜素的酿酒酵母中,实现了西红花酸的生产. Martí等人^[40]通过病毒转导系统将*CsCCD2*和*BdCCD4.1*基因导入烟草,实现西红花酸在植物中的生产(表3).

3.2 西红花苷合成生物学研究

Raghavan等人^[95]将植物来源的糖基转移酶(甜叶菊(*Stevia rebaudiana*)的*SrUGT76G1*, 拟南芥(*Arabidopsis thaliana*)的*At71C125571C2*和*At71C125571E1*, 西红花的*CsUGTUN1671*、*CsUGTUN4522*、*CsUGT2*和*CsVrUGT2*)整合进大肠杆菌中,实现西红花苷的生产. Martí等人^[40]通过病毒转导系统将*CsCCD2*和*BdCCD4.1*基因导入烟草植株,成功检测到多种西红花苷的产生.

微生物糖基转移酶具有很高的底物杂泛性,能够催化多种底物生成对应的糖基化产物,如人参皂苷^[96]、根皮素^[97]、红景天苷^[98]等.参与西红花苷生物

合成的微生物糖基转移酶研究也备受关注.基于已报道的西红花苷生物合成糖基转移酶, Ding等人^[94]在NCBI蛋白质数据进行微生物糖基转移酶的搜索,根据序列相似度分析筛选得到23个候选基因.原核表达及功能验证发现,来源于芽孢杆菌属(*Bacillus subtilis* 168和*Bacillus cereus* WQ9-2)的*Bs-GT*(GenBank: WP_003220110.1)和*Bc-GTA*(GenBank: EEL00975.1)皆具有催化西红花酸生成西红苷III和西红花苷V的功能,将其导入大肠杆菌,外源饲喂西红花酸(50 mg/L),可以实现西红苷III和西红苷V的生产. *Bs-GT*和*Bc-GTA*催化西红花酸生成西红苷的转化率分别为15.3%和6.7%. Wang等人^[78]基于其实验室数据库,挖掘参与西红花苷生物合成的微生物糖基转移酶,发现来源于枯草芽孢杆菌(*Bacillus subtilis*)的3个糖基转移酶(*BsYjiC*、*BsYdhE*和*BsYojK*)可能具有催化西红花酸生产5种西红花苷的功能(产物结构有待确证),将它们整合至西红花酸产量为4.42 mg/L的工程化大肠杆菌菌株,培养后仅检测到微量西红苷V的产生.

基于栀子西红花苷生物合成途径研究的突破, Pu等人^[39]将从栀子中挖掘得到的两个糖基转移酶(*GjUGT74F8*和*GjUGT94E13*)整合至大肠杆菌中,通过外源饲喂西红花酸(50 mg/L)及葡萄糖的方式,实现5种西红花苷的生产.结合过表达葡萄糖-6-磷酸变位酶*pgm*基因和UDP-葡萄糖焦磷酸化酶*galU*基因,加强大肠杆菌内源UDP-葡萄糖生物合成途径,提高其内源

表3 西红花酸及西红花苷的合成生物学比较分析

Table 3 Comparative analysis on the synthetic biology of crocetin and crocins

产物	底盘细胞	合成方式	基因元器件	产量(mg/L)或转化率	文献
西红花酸	酿酒酵母	从头合成	<i>PsCrtZ</i> , <i>CsCCD2ScALD</i>	1.219(摇瓶水平) 6.278(发酵水平)	[93]
	酿酒酵母	从头合成	<i>CsBCH</i> , <i>CsCCD2</i> , <i>CsALDH3</i>	可检测	[71]
	大肠杆菌	从头合成	<i>CsCCD2</i> , <i>NcALD8</i>	4.42(发酵水平)	[78]
	烟草细胞	从头合成	<i>CsCCD2</i> 或 <i>BdCCD4.1</i>	可检测	[40]
V	大肠杆菌	从头合成	<i>CsCCD2</i> , <i>NcALD8</i> , <i>BsYjiC</i> , <i>BsYdhE</i> , <i>BsYojK</i>	可检测	[78]
III, V	大肠杆菌	底物饲喂(50 mg/L西红花酸)	<i>Bs-GT</i> 或 <i>Bc-GTA</i>	15.3%或6.7%(摇瓶水平)	[94]
I, II, III, IV, V	大肠杆菌	从头合成	<i>SrUGT76G1</i> , <i>At71C125571C2</i> , <i>At71C125571E1</i> , <i>CsUGT2</i> , <i>CsVrUGT2</i>	可检测	[95]
	大肠杆菌	底物饲喂(50 mg/L西红花酸)	<i>GjUGT74F8</i> , <i>GjUGT94E13</i>	I~V: 12.48, 17.21, 44.99, 1.08, 15.82(摇瓶水平)	[48]
	烟草细胞	从头合成	<i>CsCCD2</i> 或 <i>BdCCD4.1</i>	可检测	[40]

UDP-葡萄糖含量, 进而提高西红花苷产量. 工程菌条件优化实验表明, 50 mmol/L $\text{NaH}_2\text{PO}_4\text{-Na}_2\text{HPO}_4$ 缓冲液 (pH 8.0) 加5%葡萄糖是西红花酸糖基化效率最高的反应条件. 通过GjUGT94E13和GjUGT74F8的催化, 西红花酸(50 mg/L)到5种西红花苷的最大转化率可达到66.1%^[48,94].

综上所述, 植物及微生物来源的UGT皆可实现西红花苷的生产, 但栀子中挖掘得到的UGT具有更高的催化活性. 通过底物饲喂的方法进行合成生物学研究为5种西红花苷的合成生物学研究提供重要参考.

4 展望

中药活性成分是其发挥疗效的物质基础, 也是新药研发的重要来源, 传统的化学提取或化学合成难以满足科研及新药研发的需求^[20]. 近几年, 大宗常用中药如人参^[99]、三七^[100,101]、丹参^[102]、灵芝^[103]、红景天^[104]、灯盏花^[105]、穿心莲^[106]、甘草^[107]、黄花蒿^[108]、天麻^[109]、栀子^[39]等基因组草图或精细图先后公布, 极大推动中药活性成分生物合成途径的解析及合成生物学研究.

西红花苷具有广泛的药理活性及良好的成药潜力, 但由于其资源分布的特殊性, 导致其资源匮乏、价格昂贵, 亟须寻找替代资源或开展合成生物学研究. 西红花苷主要在鸢尾科番红花的柱头和茜草科栀子的成熟果实中积累, 且其在番红花柱头中的含量至少为其在同等干重栀子果实中含量的3倍^[32,43,49]. 另一方面, 相较于采收困难、种植条件苛刻的西红花来说, 栀子具有产量高、易于栽培等优势, 因此被认为是西红花的理想替代资源^[110,111]. 此外, 研究显示在鸢尾科黄金番红花的柱头和花被片^[61]、马钱科密蒙花(*Buddleja officinalis*)和大叶醉鱼草的花^[62,112]、木犀科夜花(*Nyc-*

tanthes arborescens)的花萼^[113]、菊科牛蒡(*Arctium lappa*)的叶^[114]、百部科蔓生百部(*Stemona japonica*)的根^[115], 以及豆科含羞草(*Mimosa pudica*)的全草^[116]中也检测到了西红花苷的存在. 西红花苷生物合成途径解析的研究主要集中在西红花、栀子及大叶醉鱼草中, 其中栀子中西红花苷生物合成途径解析的相对完整, 西红花中的西红花苷生物合成途径尚缺少关键UGTs, 大叶醉鱼草中的西红花苷生物合成途径缺少关键ALDHs和UGTs, 有待进一步研究突破. 另一方面, 不同科属物种中西红花苷生物合成途径的解析, 将为自然界不同物种间的趋同进化研究提供参考.

合成生物学研究所得到的西红花苷生产菌株最大产量仍处于毫克级别, 不足以供工业化生产应用. 因此, 西红花苷合成生物学研究仍需进一步优化, 特提出以下建议供相关研究人员参考: (1) 综合利用现有西红花苷合成相关基因元器件及新发现的基因元器件(如: 栀子GjCCD4a、GjALDH2C3、GjUGT74F8和GjUGT94E13, 及西红花CsALDH3II和CsUGT74AD1等), 继续开展西红花苷合成生物学研究. (2) 结合组学研究, 对不同科属物种中西红花苷生物合成途径进行深入研究. (3) 结合酶学分析及突变研究, 筛选更多可用于西红花苷合成生物学研究的高效基因元器件. (4) 通过调整底盘细胞代谢流等方式, 提高西红花苷产量. (5) 通过对西红花苷中上游类胡萝卜素代谢途径的研究, 进一步提高底盘细胞类胡萝卜素产量. (6) 通过转录水平调控活性成分合成途径酶基因及相关基因表达^[117,118], 实现西红花苷高效生产.

多学科交融及发展将大大促进药用天然产物合成生物学研究的发展与进步, 为缓解其资源匮乏及药物开发及应用奠定基础, 对中药资源可持续发展具有重大意义.

参考文献

- Chinese Pharmacopoeia Commission. Pharmacopoeia of the People's Republic of China (in Chinese). Beijing: China Medical Science Press, 2020 [国家药典委员会. 中华人民共和国药典. 北京: 中国医药科技出版社, 2020]
- Frusciante S, Diretto G, Bruno M, et al. Novel carotenoid cleavage dioxygenase catalyzes the first dedicated step in saffron crocin biosynthesis. *Proc Natl Acad Sci USA*, 2014, 111: 12246–12251
- Ji A, Jia J, Xu Z, et al. Transcriptome-guided mining of genes involved in crocin biosynthesis. *Front Plant Sci*, 2017, 8: 518
- Melnik J P, Wang S, Marcone M F. Chemical and biological properties of the world's most expensive spice: Saffron. *Food Res Int*, 2010, 43: 1981–1989
- Khorasanchi Z, Shafiee M, Kermanshahi F, et al. *Crocus sativus* a natural food coloring and flavoring has potent anti-tumor properties. *Phytomedicine*, 2018, 43: 21–27
- Rahaiee S, Hashemi M, Shojaosadati S A, et al. Nanoparticles based on crocin loaded chitosan-alginate biopolymers: Antioxidant activities,

- bioavailability and anticancer properties. *Int J Biol Macromol*, 2017, 99: 401–408
- 7 Sawant A V, Srivastava S, Prassanawar S S, et al. Crocin, a carotenoid, suppresses spindle microtubule dynamics and activates the mitotic checkpoint by binding to tubulin. *Biochem Pharmacol*, 2019, 163: 32–45
- 8 Pashirzad M, Shafiee M, Avan A, et al. Therapeutic potency of crocin in the treatment of inflammatory diseases: Current status and perspective. *J Cell Physiol*, 2019, 234: 14601–14611
- 9 Fernández-Albarral J A, Ramírez A I, de Hoz R, et al. Neuroprotective and anti-inflammatory effects of a hydrophilic saffron extract in a model of glaucoma. *Int J Mol Sci*, 2019, 20: 4110
- 10 Sung Y Y, Kim H K. Crocin ameliorates atopic dermatitis symptoms by down regulation of Th2 response via blocking of NF- κ B/STAT6 signaling pathways in mice. *Nutrients*, 2018, 10: 1625
- 11 Finley J W, Gao S. A perspective on *Crocus sativus* L. (Saffron) constituent crocin: A potent water-soluble antioxidant and potential therapy for Alzheimer's disease. *J Agric Food Chem*, 2017, 65: 1005–1020
- 12 Naghizadeh B, Mansouri M T, Ghorbanzadeh B, et al. Protective effects of oral crocin against intracerebroventricular streptozotocin-induced spatial memory deficit and oxidative stress in rats. *Phytotherapy Res*, 2013, 20: 537–542
- 13 Batarseh Y S, Bharate S S, Kumar V, et al. *Crocus sativus* extract tightens the blood-brain barrier, reduces amyloid β load and related toxicity in 5XFAD mice. *ACS Chem Neurosci*, 2017, 8: 1756–1766
- 14 Talaei A, Hassanpour Moghadam M, Sajadi Tabassi S A, et al. Crocin, the main active saffron constituent, as an adjunctive treatment in major depressive disorder: A randomized, double-blind, placebo-controlled, pilot clinical trial. *J Affect Disorders*, 2015, 174: 51–56
- 15 Masaki M, Aritake K, Tanaka H, et al. Crocin promotes non-rapid eye movement sleep in mice. *Mol Nutr Food Res*, 2012, 56: 304–308
- 16 Hosseinzadeh H, Sadeghnia H R, Ghaeni F A, et al. Effects of saffron (*Crocus sativus* L.) and its active constituent, crocin, on recognition and spatial memory after chronic cerebral hypoperfusion in rats. *Phytother Res*, 2012, 26: 381–386
- 17 Han Y, Xie G Y, Li R, et al. Research progress on pharmacological activities of crocin (in Chinese). *Drugs Clin*, 2017, 32: 1806–1814 [韩宇, 谢国勇, 李冉, 等. 西红花苷药理活性的研究进展. *现代药物与临床*, 2017, 32: 1806–1814]
- 18 Song J Y, Xu Z C, Chen S L. Introduction to the thematic issue of herbgonomics (in Chinese). *Sci Sin Vitae*, 2018, 48: 349–351 [宋经元, 徐志超, 陈士林. 本草基因组学专辑简介. *中国科学: 生命科学*, 2018, 48: 349–351]
- 19 Xin T, Zhang Y, Pu X, et al. Trends in herbgonomics. *Sci China Life Sci*, 2019, 62: 288–308
- 20 Chen S L, Song J Y. Herbgonomics (in Chinese). *China J Chin Mater Med*, 2016, 41: 3881–3889 [陈士林, 宋经元. 本草基因组学. *中国中药杂志*, 2016, 41: 3881–3889]
- 21 Lander E S, Linton L M, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*, 2001, 409: 860–921
- 22 Venter J C, Adams M D, Myers E W, et al. The sequence of the human genome. *Science*, 2001, 291: 1304–1351
- 23 Crick F. Central dogma of molecular biology. *Nature*, 1970, 227: 561–563
- 24 Watson J D, Crick F H C. Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature*, 1953, 171: 737–738
- 25 Chai M Z, Jia B, Li B Z, et al. Advances in synthetic genome and genome rearrangement (in Chinese). *Chin Bull Life Sci*, 2019, 31: 46–53 [柴梦哲, 贾斌, 李炳志, 等. 人工基因组合成与重排研究进展. *生命科学*, 2019, 31: 46–53]
- 26 Zhang L Y, Chang S H, Wang J. Synthetic biology: From the first synthetic cell to see its current situation and future development. *Chin Sci Bull*, 2011, 56: 229–237 [张柳燕, 常素华, 王晶. 从首个合成细胞看合成生物学的现状与发展. *科学通报*, 2010, 55: 3477–3488]
- 27 Ro D K, Paradise E M, Ouellet M, et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*, 2006, 440: 940–943
- 28 Paddon C J, Westfall P J, Pitera D J, et al. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature*, 2013, 496: 528–532
- 29 Zeng Q P, Bao F. Research achievements in the synthetic biology and metabolic engineering of artemisinin (in Chinese). *Chin Sci Bull*, 2011, 56: 2289–2297 [曾庆平, 鲍飞. 青蒿素合成生物学及代谢工程研究进展. *科学通报*, 2011, 56: 2289–2297]
- 30 Li B J, Wang H, Gong T, et al. Improving 10-deacetylbaccatin III-10- β -O-acetyltransferase catalytic fitness for taxol production. *Nat Commun*, 2017, 8: 15544
- 31 Alavizadeh S H, Hosseinzadeh H. Bioactivity assessment and toxicity of crocin: A comprehensive review. *Food Chem Toxicol*, 2014, 64: 65–80
- 32 Kyriakoudi A, Tsimidou M Z, O'Callaghan Y C, et al. Changes in total and individual crocetin esters upon *in vitro* gastrointestinal digestion of saffron aqueous extracts. *J Agric Food Chem*, 2013, 61: 5318–5327
- 33 Nagatosh M, Terasaka K, Owaki M, et al. UGT75L6 and UGT94E5 mediate sequential glucosylation of crocetin to crocin in *Gardenia jasminoides*. *FEBS Lett*, 2012, 586: 1055–1061
- 34 Zhu A Y, Xia Q, Li X B, et al. Research progress of purification, determination and structure-activity relationship of crocins from gardenia and saffron (in Chinese). *Chin J Pharm Anal*, 2018, 38: 735–747 [朱安运, 夏婧, 李小波, 等. 栀子与西红花中西红花苷纯化、分析及构效关系研究进展. *药物分析杂志*, 2018, 38: 735–747]
- 35 Ketmaro P, Muangsiri W, Werawatganone P. UV spectroscopic characterization and stabilities of natural colorants from roselle calyx, lac resin and

- gardenia fruit. *J Health Res*, 2010, 24: 7–13
- 36 Van Calsteren M R, Bissonnette M C, Cormier F, et al. Spectroscopic characterization of crocetin derivatives from *Crocus sativus* and *Gardenia jasminoides*. *J Agric Food Chem*, 1997, 45: 1055–1061
 - 37 Dufresne C, Cormier F, Dorion S, et al. Glycosylation of encapsulated crocetin by a *Crocus sativus* L. cell culture. *Enzyme Microb Technol*, 1999, 24: 453–462
 - 38 Demurtas O C, Frusciante S, Ferrante P, et al. Candidate enzymes for saffron crocin biosynthesis are localized in multiple cellular compartments. *Plant Physiol*, 2018, 177: 990–1006
 - 39 Xu Z, Pu X, Gao R, et al. Tandem gene duplications drive divergent evolution of caffeine and crocin biosynthetic pathways in plants. *BMC Biol*, 2020, 18: 63
 - 40 Martí M, Diretto G, Aragonés V, et al. Efficient production of saffron crocins and picrocrocin in *Nicotiana benthamiana* using a virus-driven system. *Metab Eng*, 2020, 61: 238–250
 - 41 Pfister S, Meyer P, Steck A, et al. Isolation and structure elucidation of carotenoid-glycosyl esters in gardenia fruits (*Gardenia jasminoides* Ellis) and saffron (*Crocus sativus* Linne). *J Agric Food Chem*, 1996, 44: 2612–2615
 - 42 Choi H. Isolation and characterization of the major colorant in gardenia fruit. *Dyes Pigment*, 2001, 49: 15–20
 - 43 Carmona M, Zalacain A, Sánchez A M, et al. Crocetin esters, picrocrocin and its related compounds present in *Crocus sativus* stigmas and *Gardenia jasminoides* fruits. Tentative identification of seven new compounds by LC-ESI-MS. *J Agric Food Chem*, 2006, 54: 973–979
 - 44 Moraga A R, Nohales P F, Pérez J A F, et al. Glucosylation of the saffron apocarotenoid crocetin by a glucosyltransferase isolated from *Crocus sativus* stigmas. *Planta*, 2004, 219: 955–966
 - 45 Zalacain A, Ordoudi S A, Díaz-Plaza E M, et al. Near-infrared spectroscopy in saffron quality control: Determination of chemical composition and geographical origin. *J Agric Food Chem*, 2005, 53: 9337–9341
 - 46 Anastasaki E G, Kanakis C D, Pappas C, et al. Quantification of crocetin esters in saffron (*Crocus sativus* L.) using Raman spectroscopy and chemometrics. *J Agric Food Chem*, 2010, 58: 6011–6017
 - 47 Pittenauer E, Koulakiotis N S, Tsaibopoulos A, et al. In-chain neutral hydrocarbon loss from crocin apocarotenoid ester glycosides and the crocetin aglycon (*Crocus sativus* L.) by ESI-MSⁿ ($n = 2, 3$). *J Mass Spectrom*, 2013, 48: 1299–1307
 - 48 Pu X, He C, Yang Y, et al. *In vivo* production of five crocins in the engineered *Escherichia coli*. *ACS Synth Biol*, 2020, 9: 1160–1168
 - 49 Chen Y, Zhang H, Tian X, et al. Antioxidant potential of crocins and ethanol extracts of *Gardenia jasminoides* ELLIS and *Crocus sativus* L.: A relationship investigation between antioxidant activity and crocin contents. *Food Chem*, 2008, 109: 484–492
 - 50 Ozaki A, Kitano M, Furusawa N, et al. Genotoxicity of gardenia yellow and its components. *Food Chem Toxicol*, 2002, 40: 1603–1610
 - 51 Aman R, Schieber A, Carle R. Effects of heating and illumination on *trans-cis* isomerization and degradation of β -carotene and lutein in isolated spinach chloroplasts. *J Agric Food Chem*, 2005, 53: 9512–9518
 - 52 Gouveia L, Empis J. Relative stabilities of microalgal carotenoids in microalgal extracts, biomass and fish feed: Effect of storage conditions. *Innov Food Sci Emerg Technol*, 2003, 4: 227–233
 - 53 Chen Y, Cai L, Zhao C, et al. Spectroscopic, stability and radical-scavenging properties of a novel pigment from gardenia. *Food Chem*, 2008, 109: 269–277
 - 54 Cai S N, Chen Y. Stability investigation of different crocins (in Chinese). *Food Sci Technol*, 2012, 37: 212–215 [蔡仕宁, 陈阳. 不同结构西红花苷稳定性的研究. 食品科技, 2012, 37: 212–215]
 - 55 Pu X, Li Z, Tian Y, et al. The honeysuckle genome provides insight into the molecular mechanism of carotenoid metabolism underlying dynamic flower coloration. *New Phytol*, 2020, 227: 930–943
 - 56 Ruiz-Sola M Á, Rodríguez-Concepción M. Carotenoid biosynthesis in *Arabidopsis*: A colorful pathway. *Arabidopsis Book*, 2012, 10: e0158
 - 57 Walter M H, Strack D. Carotenoids and their cleavage products: Biosynthesis and functions. *Nat Prod Rep*, 2011, 28: 663–692
 - 58 Alder A, Jamil M, Marzorati M, et al. The path from β -carotene to carlactone, a strigolactone-like plant hormone. *Science*, 2012, 335: 1348–1351
 - 59 Shu H M, Guo S Q, Gong Y Y, et al. Advances in crocin biosynthesis and CCD gene family (in Chinese). *Chin Agric Sci Bull*, 2016, 32: 91–95 [束红梅, 郭书巧, 巩元勇, 等. 西红花苷的生物合成及CCD家族基因研究进展. 中国农学通报, 2016, 32: 91–95]
 - 60 Liu Y C, Zhang C, Dong B, et al. Advances of CCD subfamily in higher plants (in Chinese). *J Agric Biotechnol*, 2019, 27: 720–734 [刘玉成, 张超, 董彬, 等. 高等植物CCD亚家族基因研究进展. 农业生物技术学报, 2019, 27: 720–734]
 - 61 Ahrazem O, Rubio-Moraga A, Berman J, et al. The carotenoid cleavage dioxygenase CCD2 catalysing the synthesis of crocetin in spring crocuses and saffron is a plastidial enzyme. *New Phytol*, 2016, 209: 650–663
 - 62 Ahrazem O, Diretto G, Argandoña J, et al. Evolutionarily distinct carotenoid cleavage dioxygenases are responsible for crocetin production in *Buddleja davidii*. *J Exp Bot*, 2017, 68: 4663–4677
 - 63 Shin J H, Kim S R, An G. Rice aldehyde dehydrogenase7 is needed for seed maturation and viability. *Plant Physiol*, 2009, 149: 905–915
 - 64 Kotchoni S O, Kuhns C, Ditzer A, et al. Over-expression of different aldehyde dehydrogenase genes in *Arabidopsis thaliana* confers tolerance to

- abiotic stress and protects plants against lipid peroxidation and oxidative stress. *Plant Cell Environ*, 2006, 29: 1033–1048
- 65 Sunkar R, Bartels D, Kirch H H. Overexpression of a stress-inducible aldehyde dehydrogenase gene from *Arabidopsis thaliana* in transgenic plants improves stress tolerance. *Plant J*, 2003, 35: 452–464
- 66 Kim N H, Hwang B K. Pepper aldehyde dehydrogenase CaALDH1 interacts with Xanthomonas effector AvrBsT and promotes effector-triggered cell death and defence responses. *J Exp Bot*, 2015, 66: 3367–3380
- 67 Liu F, Cui X, Horner H T, et al. Mitochondrial aldehyde dehydrogenase activity is required for male fertility in maize. *Plant Cell*, 2001, 13: 1063–1078
- 68 Liu F, Schnable P S. Functional specialization of maize mitochondrial aldehyde dehydrogenases. *Plant Physiol*, 2002, 130: 1657–1674
- 69 Gómez-Gómez L, Parra-Vega V, Rivas-Sendra A, et al. Unraveling massive crocins transport and accumulation through proteome and microscopy tools during the development of saffron stigma. *Int J Mol Sci*, 2017, 18: 76
- 70 Gómez-Gómez L, Pacios L F, Diaz-Perales A, et al. Expression and interaction analysis among saffron ALDHs and crocetin dialdehyde. *Int J Mol Sci*, 2018, 19: 1409
- 71 Tan H, Chen X, Liang N, et al. Transcriptome analysis reveals novel enzymes for apo-carotenoid biosynthesis in saffron and allows construction of a pathway for crocetin synthesis in yeast. *J Exp Bot*, 2019, 70: 4819–4834
- 72 Thibodeaux C, Melançon Iii C, Liu H. Natural-product sugar biosynthesis and enzymatic glycodiversification. *Angew Chem Int Ed*, 2008, 47: 9814–9859
- 73 Gachon C M M, Langlois-Meurinne M, Saindrenan P. Plant secondary metabolism glycosyltransferases: The emerging functional analysis. *Trends Plant Sci*, 2005, 10: 542–549
- 74 Lim E K. Plant glycosyltransferases: Their potential as novel biocatalysts. *Chem Eur J*, 2005, 11: 5486–5494
- 75 Lairson L L, Henrissat B, Davies G J, et al. Glycosyltransferases: Structures, functions, and mechanisms. *Annu Rev Biochem*, 2008, 77: 521–555
- 76 Dufresne C, Cormier F, Dorion S. *In vitro* formation of crocetin glucosyl esters by *Crocus sativus* callus extract. *Planta Med*, 1997, 63: 150–153
- 77 Côté F, Cormier F, Dufresne C, et al. A highly specific glucosyltransferase is involved in the synthesis of crocetin glucosylesters in *Crocus sativus* cultured cells. *J Plant Physiol*, 2001, 158: 553–560
- 78 Wang W, He P, Zhao D, et al. Construction of *Escherichia coli* cell factories for crocin biosynthesis. *Microb Cell Fact*, 2019, 18: 120
- 79 Jain M, Srivastava P L, Verma M, et al. *De novo* transcriptome assembly and comprehensive expression profiling in *Crocus sativus* to gain insights into apocarotenoid biosynthesis. *Sci Rep*, 2016, 6: 22456
- 80 Zhang H, Zhang J G, Zhang L, et al. Study on synthesis of active compounds from saffron (in Chinese). *Chem Res Appl*, 2009, 21: 794–800 [张宏, 张俊国, 张伶, 等. 西红花有效成分合成研究进展. 化学研究与应用, 2009, 21: 794–800]
- 81 Cameron D E, Bashor C J, Collins J J. A brief history of synthetic biology. *Nat Rev Microbiol*, 2014, 12: 381–390
- 82 Alberts B. A grand challenge in biology. *Science*, 2011, 333: 1200
- 83 Hanson A D, Jez J M. Synthetic biology meets plant metabolism. *Plant Sci*, 2018, 273: 1–2
- 84 Xu Y, Cheng J F. Small molecule drug development based on derivatization of natural products (in Chinese). *Chin Sci Bull*, 2017, 62: 908–919 [徐悦, 程杰飞. 基于天然产物衍生优化的小分子药物研发. 科学通报, 2017, 62: 908–919]
- 85 Zhao X, Park S Y, Yang D, et al. Synthetic biology for natural compounds. *Biochemistry*, 2019, 58: 1454–1456
- 86 Lau W, Sattely E S. Six enzymes from mayapple that complete the biosynthetic pathway to the etoposide aglycone. *Science*, 2015, 349: 1224–1228
- 87 Galanie S, Thodey K, Trenchard I J, et al. Complete biosynthesis of opioids in yeast. *Science*, 2015, 349: 1095–1100
- 88 Fan C Y, Liu L Y, Shen Y, et al. Progress of biosynthesis of Morphine and its industrial manufacture (in Chinese). *Chin Sci Bull*, 2016, 61: 1436–1444 [范楚珏, 刘龙英, 沈玥, 等. 吗啡的合成生物学研究和工业化生产. 科学通报, 2016, 61: 1436–1444]
- 89 Luo X, Reiter M A, d’Espaux L, et al. Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature*, 2019, 567: 123–126
- 90 Liu X, Cheng J, Zhang G, et al. Engineering yeast for the production of breviscapine by genomic analysis and synthetic biology approaches. *Nat Commun*, 2018, 9: 448
- 91 Yan X, Fan Y, Wei W, et al. Production of bioactive ginsenoside compound K in metabolically engineered yeast. *Cell Res*, 2014, 24: 770–773
- 92 Wang P, Wei Y, Fan Y, et al. Production of bioactive ginsenosides Rh2 and Rg3 by metabolically engineered yeasts. *Metab Eng*, 2015, 29: 97–105
- 93 Chai F, Wang Y, Mei X, et al. Heterologous biosynthesis and manipulation of crocetin in *Saccharomyces cerevisiae*. *Microb Cell Fact*, 2017, 16: 54
- 94 Ding F, Liu F, Shao W, et al. Efficient synthesis of crocins from crocetin by a microbial glycosyltransferase from *Bacillus subtilis* 168. *J Agric Food Chem*, 2018, 66: 11701–11708
- 95 Raghavan S, Hansen J, Sonkar S, et al. Methods and materials for recombinant production of saffron compounds. WO Patent, 2013021261A3, 2013

- 96 Zhuang Y, Yang G Y, Chen X, et al. Biosynthesis of plant-derived ginsenoside Rh2 in yeast via repurposing a key promiscuous microbial enzyme. *Metab Eng*, 2017, 42: 25–32
- 97 Pandey R P, Li T F, Kim E H, et al. Enzymatic synthesis of novel phloretin glucosides. *Appl Environ Microbiol*, 2013, 79: 3516–3521
- 98 Fan B, Chen T, Zhang S, et al. Mining of efficient microbial UDP-glycosyltransferases by motif evolution cross plant kingdom for application in biosynthesis of salidroside. *Sci Rep*, 2017, 7: 463
- 99 Xu J, Chu Y, Liao B, et al. *Panax ginseng* genome examination for ginsenoside biosynthesis. *Gigascience*, 2017, 6: 1–15
- 100 Chen W, Kui L, Zhang G, et al. Whole-genome sequencing and analysis of the Chinese herbal plant *Panax notoginseng*. *Mol Plant*, 2017, 10: 899–902
- 101 Zhang D, Li W, Xia E H, et al. The medicinal herb *Panax notoginseng* genome provides insights into ginsenoside biosynthesis and genome evolution. *Mol Plant*, 2017, 10: 903–907
- 102 Xu H, Song J, Luo H, et al. Analysis of the genome sequence of the medicinal plant *Salvia miltiorrhiza*. *Mol Plant*, 2016, 9: 949–952
- 103 Chen S, Xu J, Liu C, et al. Genome sequence of the model medicinal mushroom *Ganoderma lucidum*. *Nat Commun*, 2012, 3: 913
- 104 Fu Y, Li L, Hao S, et al. Draft genome sequence of the Tibetan medicinal herb *Rhodiola crenulata*. *Gigascience*, 2017, 6: 1–5
- 105 Yang J, Zhang G, Zhang J, et al. Hybrid *de novo* genome assembly of the Chinese herbal fleabane *Erigeron breviscapus*. *Gigascience*, 2017, 6: 1–7
- 106 Sun W, Leng L, Yin Q, et al. The genome of the medicinal plant *Andrographis paniculata* provides insight into the biosynthesis of the bioactive diterpenoid neoandrographolide. *Plant J*, 2019, 97: 841–857
- 107 Mochida K, Sakurai T, Seki H, et al. Draft genome assembly and annotation of *Glycyrrhiza uralensis*, a medicinal legume. *Plant J*, 2017, 89: 181–194
- 108 Shen Q, Zhang L, Liao Z, et al. The genome of *Artemisia annua* provides insight into the evolution of asteraceae family and artemisinin biosynthesis. *Mol Plant*, 2018, 11: 776–788
- 109 Yuan Y, Jin X, Liu J, et al. The *Gastrodia elata* genome provides insights into plant adaptation to heterotrophy. *Nat Commun*, 2018, 9: 1615
- 110 Xie X W, Zhang J H, Ma A H. Evolved investigation of Chinese traditional medicine-Gardenoside (in Chinese). *Lishizhen Med Mat Med Res*, 2000, 11: 943–945 [谢学建, 张俊慧, 马爱华. 中药栀子研究进展. 时珍国医国药, 2000, 11: 943–945]
- 111 Li W P, Zhang Y, Ding Z S. The research progress of *Crocus sativus* (in Chinese). *J Beijing Union Univ (Nat Sci)*, 2011, 25: 55–58 [李伟平, 张云, 丁志山. 西红花的研究进展. 北京联合大学学报(自然科学版), 2011, 25: 55–58]
- 112 Liao Y H, Houghton P J, Hoult J R S. Novel and known constituents from *Buddleja* species and their activity against leukocyte eicosanoid generation. *J Nat Prod*, 1999, 62: 1241–1245
- 113 Pawar N D, Panchal S S, Kumaravelu J, et al. Crocin rich extract of *nyctanthes arborescens* flower calyx induces anti-angiogenic activity. *J Nat Prod*, 2016, 6: 40–48
- 114 Tang Y, Lou Z, Yang L, et al. Screening of antimicrobial compounds against *Salmonella typhimurium* from burdock (*Arctium lappa*) leaf based on metabolomics. *Eur Food Res Technol*, 2015, 240: 1203–1209
- 115 Yang X Z, Tang C P, Ke C Q, et al. Chemical constituents of *Stemona japonica* (in Chinese). *Nat Prod Res Dev*, 2008, 20: 399–402 [杨新洲, 唐春萍, 柯昌强, 等. 蔓生百部的化学成分研究. 天然产物研究与开发, 2008, 20: 399–402]
- 116 Patel N K, Bhutani K K. Suppressive effects of *Mimosa pudica* (L.) constituents on the production of LPS-induced pro-inflammatory mediators. *EXCLI J*, 2014, 13: 1011–1021
- 117 Ji A J, Luo H M, Xu Z C, et al. Research and perspectives on AP2/ERF transcription factors in medicinal plants (in Chinese). *Chin Sci Bull*, 2015, 60: 1272–1284 [季爱加, 罗红梅, 徐志超, 等. 药用植物转录因子AP2/ERF研究与展望. 科学通报, 2015, 60: 1272–1284]
- 118 Shi M, Wang Y, Zhou W, et al. Research progress in terms of the biosynthesis and regulation of terpenoids from medicinal plants (in Chinese). *Sci Sin-Vitae*, 2018, 48: 352–364 [时敏, 王瑶, 周伟, 等. 药用植物萜类化合物的生物合成与代谢调控研究进展. 中国科学: 生命科学, 2018, 48: 352–364]

Summary for “西红花苷生源途径解析及其合成生物学研究进展”

Research progress on biosynthesis and synthetic biology of crocin

Xiangdong Pu¹, Zhichao Xu^{1,2}, Ranran Gao¹ & Jingyuan Song^{1,2,3*}

¹ Key Lab of Chinese Medicine Resources Conservation, State Administration of Traditional Chinese Medicine of China, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100193, China;

² Engineering Research Center of Chinese Medicine Resource, Ministry of Education, Beijing 100193, China;

³ Yunnan Branch, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Jinghong 666100, China

* Corresponding author, E-mail: jysong@implad.ac.cn

Natural products are important resources for drug discovery and have marvelous roles in the treatment of various diseases, such as malaria, cancers, and viral infections. Crocins are the colored apocarotenoids derived from the stigmas of *Crocus sativus* and often used as a spice/colorant in the food industry. Pharmacologically, crocins confer protection to the central nervous system (especially in the treatment of Alzheimer's disease) and prevent cardiovascular/cerebrovascular diseases. Unfortunately, the demand for crocins based on their pharmacological benefits combined with a complex harvesting process of *C. sativus* has driven the retail price of *C. sativus* stigmas up to 2000–7000 €/kg. Thus, their expense and limited supply greatly restrict the development and clinical application of crocin-based pharmaceuticals. It is therefore urgent to develop an alternative, sustainable way to bolster the supply of crocins.

In recent years, synthetic biology has been widely used to overcome shortages of plant-based natural products, including artemisinic acid, etoposide aglycone, opioids, cannabinoids, breviscapine, and ginsenosides. As such, the genetic tools required for crocin biosynthesis have attracted a great deal of attention. The crocin biosynthetic pathway includes three stages: Stage (I), the precursor geranylgeranyl pyrophosphate (GGPP) is generated by the methylerythritol-4-phosphate (MEP) biosynthetic pathway; stage (II), C₄₀-carotenoids are produced via the carotenoid biosynthetic pathway; stage (III), crocins are synthesized by three key catalytic steps, namely, carotenoid cleavage, aldehyde oxidation, and glucosyl group transfer. A series of studies have successfully decoded the crocin biosynthetic pathway and led to its reconstitution in *Saccharomyces cerevisiae* and *Escherichia coli*.

Hereby, we summarized the chemical characterization of crocins using spectroscopy and mass spectrometry, and elucidated stage (III) of the crocin biosynthetic pathway. Carotenoid cleavage dioxygenases (CCDs) degrade carotenoids. CsCCD2, CaCCD2 (*Crocus ancyrensis*), BdCCD4.1 (*Buddleja davidii*), BdCCD4.3 and GjCCD4a (*Gardenia jasminoides*) can cleave the 7,8 (7',8') double bond of zeaxanthin to yield crocetin dialdehyde. Subsequently, crocetin dialdehyde was oxidized to crocetin by the catalysis of aldehyde dehydrogenases (ALDHs), including CsALDH54788, CsALDH3898, CsALDH20158, CsALDH11367, CsALDH311, CsALDH3, and GjALDH2C3. Finally, the biosynthesis of crocins from crocetin was catalyzed by UDP-glycosyltransferases (UGTs), including CsUGT74AD1, GjUGT74F8, GjUGT75L6, GjUGT94E5, GjUGT94E13, Bs-GT (*Bacillus subtilis* 168), and Bc-GTA (*Bacillus cereus* WQ9-2). And then, we displayed the research progress on the synthetic biology of crocins. With the rapid updating of genetic tools, we have the potential to generate more types of crocins. The *in vivo* production of crocins is highlighted in this review. The engineered *E. coli* with *GjUGT74F8* and *GjUGT94E13* can catalyze crocetin into five types of crocins. Furthermore, we put forward some suggestions that might help to improve the bioproduction of crocins, such as the exploration of evolutionary mechanisms of the key enzymes and the optimization of reaction conditions, etc.

Taken together, this review will facilitate the industrial production of crocins and the development of crocin-based pharmaceuticals, providing a theoretical basis and method guidance for the sustainable development of crocin-based traditional Chinese medicine.

crocins, chemical characterization, biosynthetic pathway, synthetic biology, *Crocus sativus*, *Gardenia jasminoides*

doi: [10.1360/TB-2020-0875](https://doi.org/10.1360/TB-2020-0875)