

陈 衡,潘相文,王飞飞,等. 大豆异黄酮生物合成与调控的分子机制研究进展[J]. 土壤与作物,2021,10(2):126 – 142.
CHEN H,PAN X W,WANG F F,et al. Molecular mechanisms of isoflavone biosynthesis and regulation in soybean;a review[J]. Soils and Crops,2021,10(2):126 – 142.

大豆异黄酮生物合成与调控的分子机制研究进展

陈 衡^{1,2,3}, 潘相文^{1,2}, 王飞飞^{1,2}, 刘长锴^{1,2}, 王 雪^{1,2,3}, 李彦生^{1,2}, 张秋英^{1,2}
(1. 中国科学院 东北地理与农业生态研究所 大豆分子设计育种重点实验室, 黑龙江 哈尔滨 150081;
2. 中国科学院 种子创新研究院, 北京 100101; 3. 中国科学院大学, 北京 100049)

摘 要: 大豆异黄酮,是由肉桂酰辅酶 A 产生的 3-苯基衍生物,属植物次生代谢产物。它具有降血压、缓解更年期综合症等诸多功效,被称为“植物雌激素”。本文在阐述影响大豆子粒中异黄酮含量的 QTL 位点的基础上,重点综述了异黄酮生物合成途径中的多种酶和相关基因,以及调控该途径中相关酶系基因表达的转录因子和相关基因,并分析探讨了该领域一些亟待解决的问题和未来发展的方向,以期为深入研究大豆异黄酮合成的分子机制及高异黄酮大豆分子育种提供理论参考。

关键词: 大豆异黄酮; 植物苯丙氨酸代谢途径; 转录因子; QTL

中图分类号: S642. 2 **文献标识码:** A **文章编号:** 2095 – 2961 (2021) 02 – 126 – 17

DOI: 10. 11689/j. issn. 2095 – 2961. 2021. 02. 002

Molecular mechanisms of isoflavone biosynthesis and regulation in soybean :a review

CHEN Heng^{1,2,3}, PAN Xiangwen^{1,2}, WANG Feifei^{1,2}, LIU Changkai^{1,2}, WANG Xue^{1,2,3}, LI Yansheng^{1,2}, ZHANG Qiuying^{1,2}
(1. Key Laboratory of Molecular Breeding and Design, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Harbin 150081, China; 2. Innovation Academy for Seed Design, Chinese Academy of Sciences, Beijing 100101, China;
3. University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: Soy isoflavone, a kind of plant secondary metabolite, is a group of 3 – phenyl derivative synthesized by cinnamyl – CoA. Soy isoflavone is referred to as phytoestrogen, for it reduces blood pressure, alleviates menopausal symptoms and has numerous other features as well. To provide a reference for further research of isoflavone biosynthesis and molecular breeding of soybean with high isoflavone contents, this review focused the research development on the enzymes and genes related to the biosynthesis of isoflavone, and highlighted transcription factors and their genes controlling this biosynthesis pathway based on a brief overview of QTLs influencing the isoflavone contents in soybean. Some questions and future research avenue of this research area were presented as well.

Key words: soy isoflavone; phenylpropanoid biosynthetic pathway; transcription factor; QTLs

0 引 言

大豆 (*Glycine max* (L.) Merrill), 古称菽,是双子叶植物纲、豆科、大豆属的一年生草本。原产于我国,已有 5000 年的栽培历史,是我国和世界最重要的几种栽培作物之一。

大豆不仅富含蛋白质和脂肪酸,还含有异黄酮及维生素 E 等多种营养成分。其中,大豆异黄酮 (Soy isoflavone, SI) 因具有降血压、降低血液胆固醇、改善人脑的认知能力和缓解更年期综合症等功能,而被称为植物雌激素^[1]。可以说,大豆异黄酮与蛋白质、维生素 E 一样,是人类最重要的营养物质。

20 世纪 40 年代,许多澳洲西部的母羊在食用了富含异黄酮类物质的三叶草后出现了一系列生殖问

题，这一现象引起了研究者 Bennetts 等^[2]的兴趣，他研究报道称大豆异黄酮具有植物雌激素的作用，从而引起科学界的关注。之后，研究者们也发现大豆异黄酮对猫科动物、啮齿类动物的生殖也有相似的影响，由此，“大豆异黄酮具有植物雌激素的作用”的观点开始被广泛接受。到了 21 世纪初，Lamartiniere 等^[3]研究报道了“早期食用大豆异黄酮可有效降低成年动物乳腺癌发病率”，这堪称是该领域一个里程碑式的发现。

随着研究的进一步深入，大豆异黄酮的理化性质被解析出来。研究表明，大豆异黄酮的密度为 $1.239\text{ g}\cdot\text{cm}^{-3}$ ，沸点均在 $367\text{ }^{\circ}\text{C}$ (760 mmHg) 左右。大豆包含 4 大类异黄酮，分别是苷元 (Aglycones)、糖苷 (Glycosides)、乙酰糖苷 (Acetylglycosides) 和丙二酰糖苷 (Malonylglycosides)^[2]；共有 12 种，分别为：大豆黄素 (Daidzein, DAI)，染料木素 (Genistein, GEN)，黄豆黄素 (Glycitein, GLY)，大豆黄素苷元 (Daidzin, DAI - G)，染料木苷 (Genistin, GEN - G)，黄豆黄素苷元 (Glycitin, GLY - G)，乙酰大豆黄素苷元 (Acetyldaidzin, DAI - GA)，乙酰染料木苷 (Acetylgenistin, GEN - GA)，乙酰黄豆黄素苷元 (Acetylglycitin, GLY - GA)，丙二酰大豆黄素苷元 (Malonyldaidzin, DAI - GM)，丙二酰染料木苷 (Malonylgenistin, GEN - GM) 以及丙二酰黄豆黄素苷元 (Malonylglycitin, GLY - GM)^[3-5]。

研究表明，丙二酰糖苷及其衍生物 (DAI - GM、GEN - GM、GLY - GM) 是大豆中异黄酮最主要的存在形式，占总异黄酮 (Total isoflavone, TIF) 的 82.5%；糖苷类 (DAI - G、GLY - G、GEN - G)、苷元类 (DAI、GEN、GLY) 和乙酰糖苷类 (DAI - GA、GEN - GA、GLY - GA) 的含量分别占 TIF 的 16.48%，0.88% 和 0.81%^[5-8]。

已有研究证实，植物苯丙氨酸代谢途径是大豆异黄酮合成的最主要途径。植物中，一切含苯丙烷骨架的物质都是由苯丙氨酸代谢途径直接或间接合成的^[9-10]。该代谢过程可产生黄酮类 (Flavone) 和木质素 (Lignin) 等重要的次生代谢产物。在大豆中，大豆黄素、黄豆黄素和染料木素等 3 种苷元型大豆异黄酮由植物苯丙氨酸途径直接合成的，其他类型的大豆异黄酮，则是由该代谢产物运输至高尔基体后，被葡萄糖基化或丙二酰葡萄糖基化而合成的^[11]。

苯丙氨酸途径的最初底物为苯丙氨酸 (Phenylalanine, Phe)，它在苯丙氨酸裂解酶 (Phenylalanine ammonia - lyase, PAL) 作用下生成肉桂酸 (Cinnamate)，再经肉桂酸 - 4 - 羟化酶 (Cinnamic acid 4 - hydroxylase, C4H)、4 - 香豆酰 CoA 连接酶 (4 - Coumarate: CoA ligase, 4CL) 催化生成香豆酰 CoA。此时，部分香豆酰 CoA 经查尔酮合酶 (Chalcone synthase, CHS) 和查尔酮还原酶 (Chalcone reductase, CHR) 催化后生成异甘草素 (Isoliquiritigenin)。异甘草素在查尔酮异构酶 (Chalcone isomerase, CHI) 催化下异构化，生成甘草素 (Liquiritigenin) 后可生成大豆黄素；异甘草素亦可经一系列反应后可生成黄豆黄素。另一部分香豆酰 CoA 被查尔酮合酶催化后可生成柚皮素 (Naringenin)，再经异黄酮合酶 (Isoflavone synthase, IFS) 可生成染料木素^[9,12-14]。

本文评述了大豆异黄酮合成代谢过程中的相关酶，重点综述了影响异黄酮含量的 QTL 位点，异黄酮的生物合成途径中的多种酶系基因，以及调控该途径中相关酶系基因表达的转录因子和相关基因的研究进展，以期为大豆异黄酮的合成分子机制研究，及高异黄酮大豆品系选育等提供理论参考。

1 影响大豆异黄酮含量的 QTL 位点

一般认为控制大豆异黄酮含量的基因位于数量性状基因座 (Quantitative trait locus, QTL)。母体遗传效应 (Maternal genetic effect) 是限制大豆异黄酮含量的最主要遗传因素。胚效应 (Embryogenetic effect) 对大豆异黄酮含量的影响也十分显著，而胞质效应 (Cytoplasmic effect) 对其影响不甚明显。由于现存的大豆品种的生长环境、遗传信息均有较大差异，所以已知的大豆异黄酮及其组分含量 QTL 无法形成稳定表现型。有关大豆异黄酮的 QTL 部分工作，具体整理在表 1 中^[15-21]。

表 1 大豆异黄酮的相关数量性状基因座
Table 1 The quantitative trait locus of SI

研究者 Researcher	重组自交系 Recombinant inbred lines	数量性状基因座 Quantitative trait locus	连锁群 Linkage group
Meksem 等 ^[15-16]	Essex × Forrest	12	第 3、5、9、11 和 12 号染色体
Primomo 等 ^[17]	AC756 × RCAT Angora	17	第 1、3、5、6、7、9、12、13、16 和 18 号染色体
Zhang 等 ^[18]	Xiaoheidou × GR8836	22	第 3、4、7、9、10、12、13 和 17 号染色体
Pei 等 ^[19]	LHD2 × NHZ	24	第 2、3、5、7、8、9、13、14、15、16、18、19 和 20 染色体
Cai 等 ^[20]	Huachun2 × Wayao	15	第 2、5、7、9、10、11、13、16、17 和 19 染色体
Wu 等 ^[21]	Zhongdou27 × Jiunong20	37	第 1、2、3、4、5、6、7、9、10、12、13、14、15、19 和 20 染色体

QTL 定位时，大多采用构建重组自交系（Recombinant inbred lines，RIL）的方式，其构建群体所需的时间比构建近等基因系（Near – isogenic line，NIL）和染色体片段代换系（Chromosome segment substitution lines，CSSLs）而言更短，且定位的精确度较 F2 家系、回交群体和双单倍体群体（Double haploid，DH）更准确。而遗传图谱构建时，多采用测序后建立分子标记的方式进行构建。早期测序成本相对较高，之后出现了 SLAF – Seq 技术使得图谱构建的门槛显著降低。随着高通量测序技术的发展，目前，重测序的成本已经大幅降低，因此测序结果最完整、最精确的重测序成为了遗传图谱构建时的第一选择。早期的分子标记技术，主要是限制性内切酶片段长度多态标记（Restriction fragment length polymorphism，RFLP）和简单重复序列标记（Simple sequence repeat，SSR）等标记；如今，单核苷酸多态性标记（Single nucleotide polymorphism，SNP）凭借其密度高、分布广泛及遗传稳定等诸多优点而最受研究者的青睐。而随着测序和分子标记技术的快速发展，全基因组关联分析（Genome – wide association studies，GWAS）在大豆研究中的应用也逐渐增多，这种用于连锁作图、揭示表型性状与基因型间内在关系的高效方法，也开始在 QTL 的定位中逐渐兴起，并取代了连锁分析（Linkage analysis）^[22-23]。

大豆的异黄酮含量属于典型的数量性状，且存在至少 10 个主效 QTL 位点。而且，由于品种之间差异较大，定位的位点也存在较大差异。因此，研究不同品种影响大豆异黄酮含量的 QTL 定位工作仍具有较大意义。而且，不同品种之间共有的 QTL 以及特有的 QTL 也值得研究。此外，QTL 定位区间的精细化，以及位点对应基因的探究与验证等研究也是很重要的工作。相信，随着生物信息学的发展，这些工作将会越来越完善。

2 异黄酮合成途径中的相关酶及基因

异黄酮合成过程中，主要有苯丙氨酸途径和类黄酮合成途径两个阶段，并受到多个酶的催化，因此，该过程受到多个基因的调控。

2.1 异黄酮合成途径中的相关酶

2.1.1 苯丙氨酸代谢途径的关键酶。PALs 作为苯丙氨酸代谢中的重要酶系，普遍存在于各种植物和真菌中^[24]。Koukol 和 Conn 等^[24]解析了苯丙氨酸生成肉桂酸和 NH₃的脱氨反应，并分离出一种脱氨酶，能够催化 L – 苯丙氨酸发生脱氨反应。Young 等^[25]认为该酶受酪氨酸（Tyrosine，Tyr）的抑制。现今，PAL 酶已从许多植物组织中纯化和分离，其分子量约为 240 ~ 330 kD，是由 4 个相同亚基构成的四聚体^[10]。

C4H 酶是苯丙氨酸代谢途径的关键酶之一，当体系中存在 NADPH 时，该酶可催化反式肉桂酸的氧化过程，并生成对 – 香豆酸。C4H 隶属于细胞色素 P450 – 依赖型的单加氧酶（Cytochrome P450 – dependent monooxygenase）CYP73 家族^[26]，是植物中较早发现并被研究的一种典型 P450 单加氧酶^[26-27]。P450s 是

一种膜结合蛋白,广泛存在于真核生物和原核生物中。而植物中,P450s 家族的成员参与了许多次生代谢中的氧化过程,间接影响了植物的抗逆及抗病等功能。除 C4H 外,IFS 和异黄酮 2'-羟化酶(Isoflavone 2'-hydroxylase, I2'H)也属于 P450s 家族,但 C4H 的存在较二者更加广泛^[28]。可以说,P450s 是植物中非常重要的一类酶,而 C4H 则是重中之重。

4CL 是木质素生物合成的关键酶之一,它位于苯丙氨酸途径与木质素特异合成途径的转折点上,催化肉桂酸及其羟基或甲氧基衍生物生成相应辅酶 A 酯的过程^[29-30]。4CLs 主要分成 2 类,I 类包括:拟南芥 At4CL1、At4CL2,杂种杨 Ptd4CL1、Ptd4CL2 和 Pt4CL1,其主要伴随木质素以及其他苯丙烷类衍生物的生物合成;II 类主要包括:拟南芥 At4CL3 及杨树 Pt4CL2,其主要伴随着类黄酮的生物合成。大豆中,II 类所占比例较高^[30]。

2.1.2 类黄酮代谢途径的关键酶。CHS 是首个被发现的 III 型聚酮合酶(Polyketide synthase, PKS),是类黄酮合成途径的限速酶,催化 3 分子丙二酰-CoA 和 1 分子对香豆酰-CoA 结合形成查尔酮的过程^[31]。查尔酮作为类黄酮、异黄酮及花青素等重要次生代谢物生物合成途径的中间产物^[32],影响着诸多下游代谢途径的进行,也对许多植物的抗逆及抗病等性状有着重要影响。CHS 蛋白是由两个 40~50 kDa 亚基构成的同源二聚体,单个亚基仍具有催化活性,二者之间起协同作用且高度保守。CHS 蛋白具有典型的 $\alpha\beta\alpha\beta\alpha$ 的结构,这一结构也是 β -酮乙基合酶(β -Ketoacyl synthase)的典型特征^[31-33]。

CHR 是植物中参与黄酮和异黄酮代谢的重要酶,它与 CHS 共同作用。Welle 等^[34]从大豆栽培品种 Harosoy63 中首次分离出 CHR,并证明其与查尔酮合成酶协同作用催化香豆-CoA 生成异甘草素。CHR 是一个由 318 个氨基酸组成的蛋白质,分子量在 35 kDa 左右;其二维结构包含无规卷曲(45.28%), α -螺旋(45.28%)和延长链(10.38%);三维结构显示,CHR 包含一个 AKR 结构域(Aldo-ket-red domain),因而具有醛酮还原酶活性^[35]。

CHI 是一种 24~29 kDa 的蛋白质,存在组织特异性。在大豆、葡萄及玉米等许多植物中都发现了该酶,但 CHI 保守性较低,氨基酸序列同源性仅为 50% 左右。CHI 在植物中分为 I 型和 II 型,在豆科植物中,I 型和 II 型 CHI 同时存在;而在其他植物中,几乎只有 I 型 CHI;在真菌中,也存在 CHI 酶系,一般将其定义为第 III 类。CHI 酶活性的高低直接影响黄酮类化合物的含量,在番茄和牵牛花中过表达 CHI 基因,黄酮类化合物产量比对照分别提高了 48 倍和 79 倍^[36-37]。

IFS 是合成大豆异黄酮最重要的酶。IFS 在 NADPH 和 O₂ 的参与下,可催化底物的芳香基团重排过程。虽然该过程目前还未被完全解析,但可以肯定的是 IFS 可以直接催化染料木素和大豆黄酮的合成反应,且该酶的底物柚皮素和甘草素几乎是所有大豆异黄酮的前体物质。与 C4H 一样,IFS 也是 P450s 的一种,但属于 CYP93C 家族^[38-39]。虽然 IFS 的分布较 C4H 而言并不广泛,但 IFS 因其直接催化异黄酮的生成,而在异黄酮的合成过程中具有无可比拟的重要性。

综上所述,大豆异黄酮的合成包含苯丙氨酸途径和类黄酮合成途径两个阶段,受到多个酶的影响。在苯丙氨酸途径中,PAL 酶为限速酶,C4H 酶和 4CL 酶是影响异黄酮前体物质合成的关键酶;而在类黄酮途径中,CHS 酶、CHR 酶、CHI 酶与大豆异黄酮的合成密切相关,IFS 酶则直接参与异黄酮的合成。

有关上述酶系及代谢途径的研究仍存在一些挑战。例如,针对“PALs 作为苯丙氨酸途径限速酶”的观点,如今也有研究,通过光照和遮光处理大豆后,发现光照组的 PAL 基因的 mRNA 和酶活性均有增加,但在异黄酮等下游次级代谢产物的积累量中却没有体现,这表明可能有其他的限速酶参与该过程^[10]。同样,一些发现也激发了人们的思考:既然 IFS 可以直接参与合成异黄酮过程,那么可否通过调控 IFS 基因的表达从而操纵整个异黄酮的合成代谢过程,进而改善豆科作物的农艺性状与营养价值?进一步而言,能否通过调节 IFS 以及其他关联基因的表达量,以改善豆科作物的抗逆境能力?同时,IFS 也可在其他植物

中（如玉米、拟南芥等）表达并生成异黄酮，因此在非豆科植物中的异黄酮分子（模块）设计育种也是一个非常有前景的领域。

2.2 调控大豆异黄酮合成的相关基因

酶系的合成受相关酶系基因的调控，植物中大豆异黄酮的合成受到多个基因影响的观点已被人认可。控制大豆异黄酮合成的基因中，最重要的是异黄酮合酶基因（*IFS*），它直接影响大豆黄素和染料木素的合成；其次，调控类黄酮途径的 *CHS*、*CHR* 和 *CHI* 等 3 个基因，对异黄酮的合成也有很大影响；此外，*PAL*、*C4H*、*4CL* 等 3 个调控植物苯丙氨酸代谢的基因，也会影响大豆异黄酮的合成。

2.2.1 直接影响大豆异黄酮合成的基因。*IFS* 基因的产物，直接催化大豆黄素和染料木素的合成反应。该基因主要存在于豆类中，基因拷贝数在 1~3 之间，编码区含有 1 个内含子（约 135~218 bp 长），cDNA 序列高度保守，在 400~600 bp 和 80~120 bp 区段完全相同，由 cDNA 推导的蛋白氨基酸序列同源性在 92%~97% 之间。Subramanian 等^[40] 分析了大豆 *IFS* 的启动子结构，发现 *IFS* 启动子的 TATA box 和 CAAT box 在 -125 bp 和 -264 bp 处存在抑制子（Repressor），而在 -537 处和 -887 处存在增强子（Enhancer）。*IFS* 启动子受 UV、水杨酸、茉莉酸甲酯和固氮菌的诱导。大豆中的 *IFS* 基因有两种，分别命名为 *IFS1* 和 *IFS2*，二者高度同源，但表达上存在组织特异性。*IFS1* 含有一个 218 bp 的内含子，主要在大豆的根系中表达；*IFS2* 含有一个 135 bp 的内含子，主要在豆荚和胚芽中表达^[40-45]。目前，有关 *IFS* 的研究已经比较深入，且存在许多直接证据表明，*IFS* 与异黄酮的合成与含量密切相关，例如，Jung 等^[14] 在模式植物拟南芥（非豆科植物且不生成异黄酮）中表达 *IFS* 基因并发现染料木素的生成；Sohn 等^[46] 对水稻进行 *IFS* 基因的转化实验，结果表明水稻种子中染料木素含量最高可达 103 μg·g⁻¹。

2.2.2 类黄酮代谢途径中的相关基因。Sommer 等^[47] 于 1988 年获得 *CHS* 基因的序列，目前已获得超过 2 700 个编码 *CHS* 的基因序列，广泛存在于在拟南芥（同源基因为 *tt4*、*lap5*、*lap6* 等）^[48-49]、水稻（同源基因为 *LOC4334901*）^[50] 和大豆中。*CHS* 基因一般由 2 个外显子和一个内含子构成，内含子插入在同一个半胱氨酸位点密码子的前 2 位碱基之间。其同源基因的保守性较高，不同种类植物之间的氨基酸同源性一般在 80% 以上。Anguraj 等^[51] 在全基因组范围内定位了 21 个 *CHS* 位点，发现 *CHS* 基因共有 14 种，即 *CHS1*~14，其中，*CHS1* 主要在绿叶中表达，*CHS2* 主要在子叶中表达，种皮中主要表达的是 *CHS7* 和 *CHS8*。*CHS* 基因表达具有组织特异性和时间特异性，且受温度、光照及营养元素等条件影响^[47,51-56]。并且，也正是因为 *CHS* 这种表达的差异，大多数植物具有多拷贝的 *CHS* 序列^[31]。Dhaubhadel 等^[57] 利用 cDNA 芯片，研究大豆胚胎发育过程中 *CHS7*、*CHS8*、*IFS* 及 *PAL* 等一些与异黄酮合成相关基因的表达，在授粉后第 70 天呈现出较高的水平，这与异黄酮在胚胎中开始积累的时间是一致的，说明上述基因对大豆种子的异黄酮积累有着重大影响；且该研究还通过比较 RCAT（高异黄酮品种）与 Harovinton（低异黄酮品种）中 *CHS7*、*CHS8* 表达量的差异，发现 RCAT 与 Harovinton 的异黄酮含量差异很大可能是因为 *CHS7*、*CHS8* 在种子发育后期表达量差异引起的^[57]，这再次佐证了这一观点。

Li 等^[58] 在大豆中首次分离出 *GmCHR* 基因。在豆科植物中，*CHR* 基因也是以小基因家族及多拷贝形式存在。他们以大豆为材料，研究发现，*CHR* 基因在大豆的各个部位均有表达，而且其基因的表达量在花中的含量最高。而紫花苜蓿被克隆出 3 个 *CHR* 家族的基因，在根部中表达较多。这说明 *CHR* 基因的表达存在物种、器官和组织的差异性^[58-60]。吴楠等^[61] 通过 RNAi 抑制济农 28 中的 *CHR1* 的表达，发现大豆昔元等异黄酮前体物质的下降了 38.7%，这为 *CHR* 影响异黄酮的合成及含量提供了有力证据。

CHI 基因的序列由 Mehdy 等^[62] 从法国豌豆中首次获得，之后，陆续从菜豆及玉米等多种植物中分离出 *CHI* 的同源基因，现已发现超过 1 300 条 *CHI* 同源基因。如前文所述，*CHI* 酶分为 I 型、II 型以及 III 型（存在于真菌中），同样 *CHI* 基因家族成员也分为 I 型、II 型以及 III 型。*CHI* 基因家族间的序列同源性较

低, cDNA 的同源性一般在 42% ~65% 之间^[62-65]。

2.2.3 苯丙氨酸途径中的相关基因。*PAL* 基因序列已经在大豆、拟南芥和水稻等多种植物中被解析, 它由一个小基因编码。在菜豆中, *PAL* 的基因有 3 个, 分别命名为 *PAL1*、*PAL2* 和 *PAL3*。这些基因的表达存在组织特异性, 在叶片中, 仅 *PAL1* 表达; 在根系中 *PAL1*、*PAL2* 和 *PAL3* 均可大量表达; 在花瓣中 *PAL2* 可大量表达, *PAL1* 表达量很少, *PAL3* 几乎不表达。*PAL* 受环境影响较大, 是一种典型的诱导酶, 受光质(光的波长)、温度(低温)和机械损伤等诱导, 也有研究表明生长素、乙烯、激动素和赤霉素等植物激素也可诱导 *PALs* 基因表达^[66-69]。

如前文所述, *C4H* 酶是一种典型的 P450 单加氧酶, 而 P450 蛋白家族是一种常见于动物肝脏细胞膜上的血红素蛋白。目前, 约 20 种 *C4H* 基因已被解析。与 *PAL* 基因类似, 它由一个小基因编码, 在不同植物中拷贝数也不尽相同。这些基因启动子大约 1.1 kb 左右, 序列具有较高的同源性, 但也有例外, 例如, 来自玉米和法国菜豆的两个序列的同源性就相对较低。*C4H* 的表达一般与该部位的木质化程度相关, 如在欧芹中, *C4H* 在维管束发达的花梗中表达丰富, 在幼叶和老叶中不表达。在一项以拟南芥为对象的研究中发现, 当细胞正在发生木质化时, *C4H* 基因表达达到峰值。此外, *C4H* 还受机械损伤及真菌感染等外界环境的影响, 这点与 *PAL* 类似^[70-72]。

植物中, *4CL* 基因以小的基因家族及多拷贝的形式存在, 具有细胞、组织和器官特异性。该基因的表达受到发育的调控和环境因子的影响, 这些方面与 *PAL*、*C4H* 等酶系的基因别无二致。但由于 *4CL* 催化的反应位于苯丙氨酸代谢途径的分支处, 所以它的表达与 *PAL*、*C4H* 又有不同。Hu 等^[73]在研究关于美洲山杨 *4CL* 酶时发现, 该植物中存在两种 *4CL* 基因 (*Pt4CL1* 和 *Pt4CL2*), 它们分别参与不同的代谢过程, 二者差异表达可以调控不同的苯丙烷类衍生物的生物合成。其中, *Pt4CL1* 主要参与植物苯丙氨酸代谢途径中木质素生物合成这一分支, 在正在发育的木质部组织中表达量较高; 而 *Pt4CL2* 主要参与该途径中类黄酮的生物成分支途径(大豆异黄酮的合成便处于这一分支), 因此在表皮细胞中含量较高^[73-76]。

2.2.4 影响大豆异黄酮含量的其它酶系基因。除了直接催化大豆异黄酮及其前体物质合成的酶系外, 有些酶还可以通过抑制苯丙氨酸途径的其他分支, 使得代谢反应更倾向于大豆异黄酮合成的方向行进。例如, 黄烷酮-3-羟化酶(Flavanone 3-hydroxylase, F3H)对大豆异黄酮合成的影响就有相关报道。F3H 虽然不催化大豆异黄酮的合成途径中的任何反应, 但它的活性会影响苯丙氨酸途径中, 柚皮素生成二氢黄酮醇(Dihydroflavonol)的分支途径, 影响柚皮素的含量, 进而影响柚皮素生成 GEN 这一分支途径, 最终导致 TIF 变化。Yu 等^[77]报道了大豆中的玉米同源基因 *CI/R* 的表达, 再结合 F3H 阻断花青素的分支途径(The anthocyanin branch pathway), 导致 TIF 水平增加。

此外, 一些编码蛋白激酶的基因对异黄酮含量也有较大影响。Wu 等^[21]以 200 个大豆品种和 150 个重组自交系为基础, 通过全基因组关联分析和连锁定位后发现, qISO8-1 号染色体上存在一个主基因座。qISO8-1 在 BC2F5 群体中精细定位于 99.5 千碱基区域, 两侧分别为 SSR_08_1651 和 SSR_08_1656。qISO8-1 的致病基因为编码丝裂原活化蛋白激酶的 GmMPK1, 其中两个天然 GmMPK1 多态性与异黄酮含量显著相关^[21]。不仅如此, 这项研究还发现, 过表达 GmMPK1 后, 大豆毛状根中异黄酮浓度升高^[21]。

综上所述, 大豆异黄酮的合成受多个基因的调控, 其中, *IFS* 是影响异黄酮合成最主要的基因; 而 *CHS*、*CHR* 和 *CHI* 等 3 个基因的影响也十分巨大; *PAL*、*C4H*、*4CL* 等基因通过调控植物苯丙氨酸代谢的基因来影响大豆异黄酮的合成。此外, 还有一些基因, 如 *F3H*, 通过抑制苯丙氨酸途径的其他分支来间接影响异黄酮的合成。这些基因已经被克隆出来, 功能也被解析清楚。但该领域中, 仍有几点值得关注, 如苯丙氨酸途径中各调控基因在不同品种大豆中表达量的差异, 因为该代谢途径及其的分支途径对大豆的生长发育过程有很大影响, 并涉及到包括抗虫、抗病及抗倒伏等诸多农艺性状, 也和异黄酮含量及花青素

含量等性状密切相关。因此，可以通过分析各基因在不同地方品种（系）中表达量的差异，结合该品种的特点、种植地的环境条件等因素进行分析，进而为地方品种的分子设计育种提供参考。此外，同源基因的深入挖掘是另一个重要的研究方向，这些基因在大多数植物中都存在对应的同源基因，通过综合比较基因间序列的差异性，可以为物种间的亲缘关系比较提供一定的参考。*PAL*、*C4H*、*4CL* 等编码异黄酮合成途径上游酶系的基因在大豆中研究相对较少，在模式植物拟南芥中研究相对较多，这些基因在非模式植物中特殊功能的研究日后应有较大的发展。同样，随着测序技术和代谢组学的进一步发展，非模式植物的特殊代谢通路也是一个很有发展前景的领域。

3 影响大豆异黄酮含量的转录因子及相关基因

近期，一些转录因子对大豆异黄酮合成的影响也颇受研究者的关注。如前文所述的 *CI/R* 就是典型的编码转录因子的基因。大豆种子中的这 2 个转录因子，可以激活植物苯基丙酸合成途径相关的基因，从而引起染料木素的减少和大豆苷的增多，而 *TIF* 则少量增多^[77]。一般认为，参与植物苯丙氨酸代谢途径调控的转录因子，主要有 *MYB* 类蛋白和 *bHLH* 类蛋白。

3.1 MYB 转录因子及其相关基因

Klempnauer 等^[78]首次将来自髓细胞瘤病毒（Avian myeloblastosis virus）中一个癌基因（即 *V-MYB*）和它在细胞中的同源基因（即 *C-MYB*）成功克隆；之后，*MYB* 的同源基因又相继在脊椎动物、果蝇（*Drosophila melanogaster*）和玉米中被发现，如上文提及的 *CI* 基因就是 *MYB* 基因家族；Biedenkapp 等^[79]研究结果显示，蛋白质 *V-MYB* 的 DNA 结合结构域（DNA-binding domain）有一段非常保守的氨基末端（N 末端），这表明 *MYB*/DNA 互作是 *MYB* 蛋白的主要功能。此外，他们还发现了 *V-MYB* 可以特异性识别 5'-AACGG-3'和 5'-AACTG-3'序列（现一般认为 *MYB* 蛋白结合 5'-AACNG-3'序列），且这段序列在生物中十分保守；并且，与已发现的 DNA 结合蛋白 *HLH* 和 *Zinc-Finger* 不同，*MYB* 蛋白的 DNA 结构域位于 N 末端，这表明 *MYB* 又是一种新类型的 DNA 结合蛋白^[80-83]。*MYB* 蛋白包含 1~4 个 R 结构域，而每个 R 结构域包含 1 个 *HTH*（Helix-turn-helix）结构以及 1 个作为疏水核心（Hydrophobic core）的色氨酸残基（Tryptophan residues），*HTH* 结构一般参与 DNA 合成过程，而疏水核心主要功能为保持 *HTH* 结构的稳定^[84-85]。

根据 R 结构域数目，*MYB* 蛋白被分为 4 类，1R-*MYB*（含 1 个 R 结构域）主要功能为稳定染色体结构，是一类重要的端粒结合蛋白；含有 3 个和 4 个 R 结构域的 *MYB* 蛋白分别被称为 R1R2R3-*MYB* 和 4R-*MYB*，前者在植物和真菌中发现，而后者只在拟南芥、杨树等少数植物中存在，且研究都不深入；含 2 个 R 结构域的 *MYB* 蛋白称为 R2R3-*MYB*，是植物细胞中含量最丰富、应用最广泛的 *MYB* 蛋白，主要参与激素应答、细胞分化、抗病抗虫和环境胁迫等过程，它们也是植物苯丙氨酸途径中最重要的转录因子之一^[85-88]。

MYB 转录因子通过调节植物苯丙氨酸代谢途径中的各个酶基因的转录水平，来影响代谢过程行进的方向和各次级代谢产物的含量。例如，大豆中的 *GmMYB12B12*（*MYB* 同源基因）就是典型的 *CHI*、*CHS*、*IFS* 正调控基因，玉米中的 *CI* 基因就是 *F3H* 的负调控基因。由此确保植物在不同时期、不同环境、不同生理状态下稳定生长和发育。Pei 等^[19]利用大豆种质 LHD2 和 NHZ 构建的群体，定位与大豆异黄酮含量相关的 24 个 QTL 中，有 13 个编码 *MYB* 转录因子的基因座，这表明许多 *MYB* 转录因子对异黄酮含量有着重要的影响。因此，我们也可以在基因水平增加或者减少一些转录因子的拷贝数，从而达到增加目的产物的效果。

在大豆中，*MYB* 的同源基因被命名 *GmMYBs*，如 *GmMYB12d*、*GmMYB12a* 和 *GmMYB12B2* 等。Du

等^[87]定位了大约 700 个 *MYB* 相关的 DNA 序列,经过筛选、比对和分析后,得到了 2 个 4R - *MYB* 蛋白、6 个 R1R2R3 - *MYB* 蛋白和 244 个 R2R3 - *MYB* 蛋白的基因序列。此外,还有 26 个与 R2R3 - *MYB* 相关的基因序列,但是它们所编码的蛋白质在 N 末端都不完整;同时,基因 *Glyma09g37340* 虽然拥有 *MYB* 基序和完整的开放阅读框 (Open reading frame, ORF),但是置信区间较低,因此被排除在 *MYB* 基因家族之外。这 244 种 R2R3 - *MYB* 的 5' - 非翻译区 (5' - Untranslated Regions, 5' - UTR) 非常短,与 3' - UTR 一样保守性较低,而它们的 DNA 结合结构域的保守性却非常高^[87]。

除此之外,Du 等^[87]还发现大豆中有 10 种 R2R3 - *MYB* 的 mRNA 前体 (Precursor mRNA, pre - mRNA) 存在自我剪接 (Alternative splicing, AS) 现象,这会造成一段 pre - mRNA 发生序列的插入和删除,从而生成多种不同的 mRNA,进而得到多种不同的 GmMYBs 蛋白,造成 GmMYBs 的类型发生改变,如 GmMYB082 的一个 189 bp 的可变启动子位点 (Alternative promoter site) 变异导致了移码 (Frame shift),将一个典型的 R2R3 - *MYB* 转变为信号重复型 *MYB* (Signal - repeat *MYB*)。

许多 *MYB* 基因的功能也被鉴定出来,如 *CI* 亚家族 (如 *GmMYB096*、*GmMYB080* 等 7 个) 包含了构成 GmMYBs 蛋白 C 末端 (C - terminal) 的 2 个基序^[88-89]; *AtMYB12* (拟南芥中的 *MYB* 基因) 会激活植物苯丙氨酸途径中 *CHS* 和 *F3H* 的表达,并对 *CHI* 的表达有一定促进作用^[88];前文所述的玉米中 *CI* 基因,及其同源基因 *PcMYB1* (欧芹中的 *MYB* 基因) 和 *AtMYB111* 可以激活 *CHS* 的表达^[88,90-91]。*MYB* 家族中除了有转录的激活蛋白 (Activator),同样也存在阻遏蛋白 (Repressor)。*AtMYB4* 通过抑制 *C4H* 的表达,来提高芥子酸酯 (一类具有紫外吸收作用的羟基肉桂酸衍生物) 的积累量^[91]; *FaMYB1* (草莓中的 *MYB* 基因) 可以抑制烟草中花青素和黄酮的生成^[92]。Yi 等^[93]和 Li 等^[94]报道了 *GmMYB12B12* 通过调控大豆中 *CHS8* 基因的表达来影响大豆异黄酮的生成;Yi 等^[95]通过 RNA_i (RNA interference) 抑制 *GmMYB176* (一种 R1 - *MYB* 基因) 表达,影响了大豆根系细胞中的 *CHS* 的表达,从而使得大豆异黄酮含量增加,证明了 *GmMYB176* 可调控 *CHS* 的表达。

3.2 bHLH 转录因子及其相关基因

Murre 等^[96]以小鼠为材料,报道了 2 段 cDNA 序列 (E12 和 E47),它们及其同源序列可以合成 1 种具有两亲活性 (Amphipathic activity) 的蛋白质,该蛋白质由 2 个螺旋结构 (Helix) 通过 1 个环状结构 (Loop) 相连组成 (即螺旋 - 环 - 螺旋, Helix - loop - helix, HLH),此蛋白具有二聚化 (Dimerization) 和 DNA 结合两种功能。之后,越来越多的 bHLH 蛋白及相关基因在线虫、果蝇、小鼠和人类中被分离出来。大多 bHLH 蛋白与序列 5' - CANNTG - 3' (E - box) 相结合,有少部分这类转录因子在 bHLH 基序后还有一个亮氨酸拉链 (Leucine zipper, LZ)。Atchley 等^[97]通过对 242 条 bHLH 蛋白序列的系统发生分析,提出了 bHLH 蛋白的自然分类法 (Natural classification),并将 bHLH 按照结合位点、是否含有亮氨酸拉链以及其他基序氨基酸构成的不同分成了 5 个家族,这表明 bHLH 蛋白家族可能起源不是单一的^[98],又对 bHLH 蛋白进一步解构,发现 bHLH 包含了 19 个保守的氨基酸残基:5 个氨基酸残基存在于碱性序列中,5 个氨基酸残基存在于第一个 α - Helix 中,1 个氨基酸残基存在于 Loop 中,8 个氨基酸残基存在于第一个 α - Helix 中。

bHLH 蛋白具有一段非常保守的结构域,可以同时结合 DNA,并参与蛋白质互作。这些功能主要通过几个基序实现,其中一个是由碱性氨基酸残基构成的基序,这保证了 bHLH 蛋白可以特异结合 E - BOX;另一个基序是由疏水残基构成,这保证了 bHLH 蛋白可以与蛋白质互作并构成二聚体^[97,99]。

已发现的 bHLH 转录因子主要有两种功能:DNA 结合和二聚化^[96]。其 DNA 结合功能取决于 bHLH 的碱性氨基酸序列,因此在氨基酸序列的前 17 位中是否富含碱性氨基酸残基决定了这个蛋白是 DNA 结合的 bHLHs (DNA binding bHLHs),还是非 DNA 结合的 bHLHs (non - DNA binding bHLHs)。在拟南芥中,

DNA 结合的 bHLHs 的氨基酸序列前 17 位, 平均含有 6 个碱性氨基酸残基, 而非 DNA 结合的 bHLHs 则平均含有 3.8 个。DNA 结合的 bHLHs 又可根据其结合位点的不同, 分为 E-Box 结合型 (E-box binder) 和非 E-Box 结合型 (non E-box binder)。根据所识别的 E-Box 类型, E-Box 结合型又分为 G-Box 结合型 (G-box binder) 和非 G-Box 结合型 (non G-box binder)^[100]。

植物中, bHLH 转录因子及其同源基因也已被发现。Li 等^[101]以模式植物拟南芥和水稻为材料, 进行全基因组分析, 鉴定出 167 个 bHLH 转录因子相关基因, 其中, 84.7% 含有内含子, 数目在 0~4 个范围内。前文提及的玉米中 *R* 基因, 便是一种研究相对较多的 bHLH 转录因子编码基因^[77], 它可以与顺式调控元件 (Cis-regulatory elements) 结合, 并通过招募其他的因子, 调控一些参与花青素途径的基因。bHLH 蛋白质被认为无法单独结合 DNA, 但越来越多的证据表明 bHLH 蛋白单独存在时, 也具有结合 DNA 的活性。不过, 主流观点依旧认为 bHLH 转录因子直接结合 DNA 的活性是可有可无的, 因为绝大多数的情况下, bHLH 会与 MYB 结合后共同调控 DNA 的转录过程; 而且, 有研究表明 MYB 互作区域的突变导致 MYB 激活 bHLH 的表达, 这说明 MYB 转录因子可能可以调控 *bHLH* 基因的转录^[102-103]。

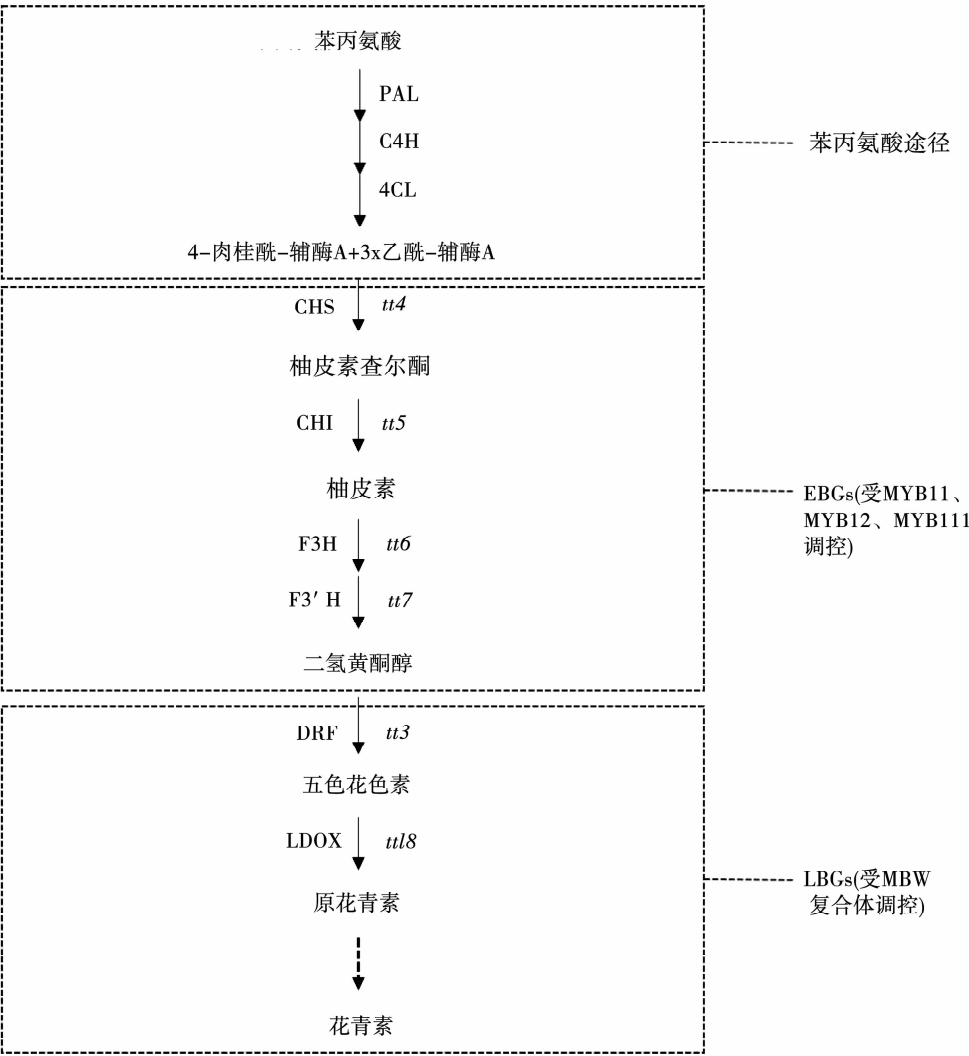
3.3 MYB 与 bHLH 互作

上述的 R2R3-MYB 可以直接与 DNA 连接, 有些也需要在 bHLH 蛋白的参与下与 DNA 相连。已有实验证明部分 bHLH 和 MYB 蛋白会相互作用^[88-91]。通常情况下, 二者互作的结构域分别是 R2R3-MYB 的 R3 结构域和 bHLH 的 N-末端^[104-109]。植物中, 两种转录因子互作的情况也是十分常见的, 且在苯丙氨酸代谢途径中存在 MYB 与 bHLH 两种蛋白结合后共同调控的情况, 如前文所述玉米中 *C1* 基因的产物 (R2R3-MYB), 便会与 *R* 基因产物 (bHLH) 互作, 且只有在 *R* 的参与下 *C1* 才能发挥转录因子的作用, 即 R-依赖型。Hernandez 等^[110]以玉米为材料, 通过 *C1*、*R*、*P1* (一种 R2R3-MYB 转录因子, 并可以在无 *R* 蛋白的参与下调控部分 *C1* 调控的基因 (*C1*-regulated gene)、*P1** (将 *P1* 中 6 个氨基酸替换成 *C1* 中对应位置中的 6 个, 其在调控 *C1*-调控基因时, 表现为 R-依赖型; 调控 *P1*-调控基因时, 表现为 R-增强型) 等 4 个基因, 研究了 *C1* 基因与 *R* 基因的互作关系, 发现 *R* 基因其中一个功能是减轻植物中的一种抑制子对 *C1* 基因的抑制作用, 从而促进 *C1* 的表达。又比如矮牵牛花中 *PhAN2* (*MYB* 同源基因) 与 *PhJAF13* (*bHLH* 同源基因) 的结合; 以及前文中所提及的草莓中 *FaMYB1*, 也会与 bHLH 蛋白结合, 且该 bHLH 与玉米中的 *R* 基因同源^[92,110]。

除了 MYB 与 bHLH 直接结合外, 二者还会与第三个蛋白质 WDR (WD repeat) 结合形成 MYB-bHLH-WDR 三聚体 (MBW)。WDR 蛋白又称 WD40 蛋白, 是真核生物体中最丰富的蛋白质家族之一, 通常会与其他生物大分子互作, 并作为支架参与蛋白质复合体的装配过程^[94]。WDR 大约由 40~60 个氨基酸残基构成, 是一类含 4~10 个随机 WD 重复结构域 (WD repeat domain) 的蛋白质家族, 而 WDR 结构域是由以色氨酸和天冬氨酸结尾的氨基酸序列组成的^[111]。因此, WDR 又称为 Trp-asp。此外, WDR 蛋白的 N-末端, 还含有一个甘氨酸-组氨酸二肽^[112-113]。WDR 蛋白本身不具有催化功能, 但它能够与多种蛋白质相互作用, 进而为大分子蛋白复合体的组合起到连接作用^[112-117]。WDR 蛋白在真核生物中十分常见, 在许多生命活动中都承担着重要的角色, 包括 G 蛋白介导的信号转导、转录调控、泛素依赖型蛋白质降解、染色质修饰、囊泡运输、细胞骨架组装和细胞周期调控等^[114-118]。并且, WDR 在花青素合成、分生组织形成、幼苗发育、花发育、光信号传递和感知等方面也发挥着重要作用^[116-119]。拟南芥中的 *TTG1* 基因就是一种典型的 WDR 蛋白编码基因, 它会影响黄酮类化合物的生物合成途径 (植物苯丙氨酸途径的一个分支)^[120]。

MBW 三聚体在植物中普遍存在。例如, 在拟南芥的种皮内细胞层的原花青素积累过程, 就涉及至少 4 种 MBW 复合物, 包括 TT2-TT8/GL3/EGL3-TTG1 (依次对应为 MYB-bHLH-WDR 的基因), 以及

MYB5 – TT8 – TTG1。TT2 – TT8 – TTG1 在调控后期生物合成基因中承担重要角色，包括无花色素双加氧酶、黄烷酮醇还原酶、*TT12* (Transparent Testa 12)、*TT19* (Transparent Testa 19)、*AHA10* (Autoinhibited H⁺ – ATPase Isoform 10) 等基因；TT2 – EGL3/GL3 – TTG1 主要调控 *BAN*、*LODX*、*TT12*、*AHA10* 等基因；MYB5 – TT8 – TTG1 主要参与 *DFR*、*LODX* 的调控。此外，花青素途径也受 MBW 复合物的调控，如 *PAP* (Production of anthocyanin pigment) 1/2/3/4 (*MYB* 同源基因)，*GL3/EGL3/TT8* (*bHLH* 同源基因)，*TTG1* (*WDR* 同源基因) (图 1)^[121]。



注：图为拟南芥中黄酮类、花青素、原花青素的生物合成途径。其中，前期生物合成基因（EBGs）受到 MYB11、MYB12、MYB111 等 3 个转录因子调控；后期生物合成基因（LBGs）受 MBW 复合物调控。流程图中大写字母表示相应的酶，如 CHS，CHI 等；小写字母表示相应的基因座，如 *u* 等。

Note: The biosynthetic pathway for flavanols, anthocyanins, and PAs in *Arabidopsis*. The early biosynthetic genes (EBGs) are activated by 3 functionally redundant R2R3 – MYB proteins (MYB11, MYB12, and MYB111), whereas the expression of the late biosynthetic genes (LBGs) requires the transcriptional activation activity of the R2R3 – MYB/bHLH/wD40 (MBW) complex. Enzymes are denoted in uppercase (CHS, CHI) and corresponding genetic loci are indicated in italic lowercase letters (*u*).

图 1 拟南芥中转录因子调控植物苯丙氨酸途径（修改自 Li 等^[121]）

Fig. 1 Regulation of phenylpropanoid biosynthetic pathway by transcription factors in *Arabidopsis*

综上所述，参与调控大豆异黄酮合成的转录因子主要有 R2R3 - MYB 类、bHLH 类以及二者同 WD40 的结合体 MBW 复合物。其中，MYB 类蛋白可以直接调控合成代谢过程，而 bHLH 蛋白通常需要和 MYB 蛋白结合后发挥作用，bHLH 能否单独结合 DNA 仍存在一定争议。虽然有关转录因子的研究已经开展了近 40 年，但该领域仍有许多问题尚未解决。如 bHLH 能否单独结合 DNA 并发挥作用的研究已经开展多年，但进展依旧缓慢。此外，有关 1R - MYB，3R - MYB 和 4R - MYB 的研究也相对较少，前者作为端粒结合蛋白，起到了稳定染色体结构的重要作用，这一领域的研究无论是在动物、植物还是微生物中的研究都十分缺乏；而后二者在植物和真菌中较多，在其他生物中鲜有发现，甚至 4R - MYB 只在拟南芥、杨树等少数植物中存在，这表明这些蛋白可能承担着部分物种中的特殊功能。此外，关于 R2R3 - MYB 对异黄酮合成相关基因的调控也只停留在部分异黄酮合成阶段的早期，其对糖苷类等其他种类异黄酮的调控过程还不甚了解，这些都将会成为未来的研究热点。

4 展 望

大豆异黄酮的研究虽已取得一定进展，但仍有许多重要问题亟待解决。其中有几项，在未来一段时间内可能会成为热点。

4.1 对大豆异黄酮合成机制的进一步解析

大豆异黄酮有 12 种，其中苷元类在植物苯丙氨酸途径和类黄酮途径中合成，该过程的解析仍然不够透彻，许多问题仍有争议，包括苯丙氨酸途径限速酶是否是 PAL 酶，IFS 酶催化底物的芳香基团重排的机制等，这些都将是未来人们关注的重点。此外，还有一些含有糖苷配基、丙二酰配基的大豆异黄酮会由前体迁移至其他部位合成加工，目前的研究中，对这些过程的分子机制了解不够深入。加强这些方面的研究，可以对大豆异黄酮乃至其他诸多次级代谢产物合成方面的研究都大有裨益。

4.2 对相关主效 QTLs 进行深入挖掘

由于大豆异黄酮合成受环境因素影响很大，这给控制大豆异黄酮合成的相关 QTLs 的解析带来了很大的困扰。因此，对 QTLs 定位方法的改进也是一项重要的研究工作。不仅如此，不同品种的 QTLs 定位结果也存在差异，因此在多个品种当中共同存在的 QTLs，以及存在于不同品种中的 QTLs 进行解析，也是一项值得研究的工作。

4.3 QTL 位点与基因间的关系

随着测序技术与生物信息学的发展，QTL 定位工作越来越精细准确，但 QTL 终究只是一个区间，上面可能存在着多个基因。因此，寻找 QTL 位点与对应基因之间的关系，将是未来一项很重要的工作。这需要更加庞大的计算量，也需要生物信息学的继续发展，同时要结合遗传转化等分子与遗传学手段，当基因达到一定数目时，需要借助“正交试验”的思路来设计实验，以便研究者更加高效地进行探究与分析。

4.4 高异黄酮大豆品种的分子设计育种

就目前已知的栽培大豆而言，虽然国内已有 10 余个高异黄酮大豆品种，但高异黄酮特用大豆的选育工作仍有很大发展空间。因为部分地方大豆品种具备高异黄酮的品质，但仅适合该地方种植，而在其它重要的农艺指标中呈现的水平相对较低。因此，我们可以建立高异黄酮分子模块，并以影响大豆品质性状的遗传改良为切入点，向目标受体亲本中分别导入相应分子模块。在现有大豆种质基础之上，采用分子模块设计育种与传统育种相结合的方式，将高异黄酮模块精准设计组装到现有的品种中，培育出富含大豆异黄酮、适应性强且具备高产、抗病、高蛋白等多个优良性状的特色大豆新品种。此外，由于一些基因，如 IFS，在其他植物（如玉米、拟南芥等）中表达时也可生成异黄酮，因此，非豆科植物中的异黄酮分子（模块）设计育种也是一个非常有前景的领域。

参考文献 (References):

- [1] BENNETT J O, YU O, HEATHERLY L G, et al. Accumulation of genistein and daidzein, soybean isoflavones implicated in promoting human health, is significantly elevated by irrigation [J]. Journal of Agricultural and Food Chemistry, 2004, 52 (25): 7574–7579.
- [2] BENNETTS H W, UNDERWOOD E J, SHIER F L. A specific breeding problem of sheep on subterranean clover pastures in western Australia [J]. The Australian Veterinary Journal, 1946, 22 (11): 2–12.
- [3] LAMARTINIERE C A, ZHAO Y X, FRITZ W A. Genistein: mammary cancer chemoprevention, in vivo mechanisms of action, potential for toxicity and bioavailability in rats [J]. Journal of Women's Cancer, 2000, 2: 11–19.
- [4] FERREIRA M P, OLIVEIRA M, MANDARINO J, et al. Changes in the isoflavone profile and in the chemical composition of tempeh during processing and refrigeration [J]. Pesquisa Agropecuária Brasileira, 2011, 46 (11): 1555–1561.
- [5] AZAM M, ZHANG S D, ABDELGHANY A M, et al. Seed isoflavone profiling of 1168 soybean accessions from major growing ecoregions in China [J]. Food Research International, 2020, 130: 108957.
- [6] SUN J M, SUN B L, HAN F X, et al. Rapid HPLC method for determination of 12 isoflavone components in soybean seeds [J]. Agricultural Sciences in China, 2011, 10 (1): 70–77.
- [7] LIU J, YANG C Q, ZHANG Q, et al. Partial improvements in the flavor quality of soybean seeds using intercropping systems with appropriate shading [J]. Food Chemistry, 2016, 207: 107–114.
- [8] WU H J, DENG J C, YANG C Q, et al. Metabolite profiling of isoflavones and anthocyanins in black soybean [*Glycine max* (L.) Merr.] seeds by HPLC–MS and geographical differentiation analysis in Southwest China [J]. Analytical Methods, 2017, 9 (5): 792–802.
- [9] YU O, JUNG W, SHI J, et al. Production of the isoflavones genistein and daidzein in nonlegume dicot and monocot tissues [J]. Plant Physiology, 2000, 124 (2): 781–794.
- [10] 李 莉, 赵 越, 马君兰. 苯丙氨酸代谢途径关键酶: PAL、C4H、4CL 研究新进展 [J]. 生物信息学, 2007, 5 (4): 187–189.
- LI L, ZHAO Y, MA J L. Recent progress on key enzymes: PAL, C4H, 4CL of phenylalanine metabolism pathway [J]. China Journal of Bioinformatics, 2007, 5 (4): 187–189.
- [11] GRAHAM T L. Flavonoid and isoflavonoid distribution in developing soybean seedling tissues and in seed and root exudates [J]. Plant Physiology, 1991, 95 (2): 594–603.
- [12] LATUNDE–DADA A O, CABELLO–HURTADO F, CZITTRICH N, et al. Flavonoid 6–hydroxylase from soybean (*Glycine max* L.), a novel plant P–450 monooxygenase [J]. The Journal of Biological Chemistry, 2001, 276 (3): 1688–1695.
- [13] LATUNDE D A O, CABELLO H F, CZITTRICH N, et al. Cloning and functional expression of a cytochrome P450 cDNA encoding 2–hydroxyisoflavanone synthase involved in biosynthesis of the isoflavonoid skeleton in licorice [J]. Plant Physiology, 1999, 121 (3): 821–828.
- [14] JUNG W, YU O, LAU S M C, et al. Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes [J]. Nature Biotechnology, 2000, 18 (2): 208–212.
- [15] MEKSEM K, NJITI V N, BANZ W J, et al. Genomic regions that underlie soybean seed isoflavone content [J]. Journal of Biomedicine and Biotechnology, 2001, 1 (1): 38–34.
- [16] KASSEM M A, SHULTZ J, MEKSEM K, et al. An updated 'Essex' by 'Forrest' linkage map and first composite interval map of QTL underlying six soybean traits [J]. Theoretical and Applied Genetics, 2006, 113: 1015–1026.
- [17] PRIMOMO V, POYSA V, ABLETT G R, et al. Mapping QTL for individual and total isoflavone content in soybean seeds [J]. Crop Science, 2005, 45 (6): 2454–2464.
- [18] ZHANG H J, LI J W, LIU Y J, et al. Quantitative trait loci analysis of individual and total isoflavone contents in soybean seeds [J]. Journal of Genetics, 2014, 93 (2): 331–338.
- [19] PEI R, ZHANG J Y, TIAN L, et al. Identification of novel QTL associated with soybean isoflavone content [J]. The Crop Journal, 2018, 6 (3): 244–252.
- [20] CAI Z D, CHENG Y B, MA Z W, et al. Fine–mapping of QTLs for individual and total isoflavone content in soybean (*Glycine max* L.) using a high–density genetic map [J]. Theoretical and Applied Genetics, 2018, 131 (3): 555–568.
- [21] WU D P, LI D M, ZHAO X, et al. Identification of a candidate gene associated with isoflavone content in soybean seeds using genome–wide

- association and linkage mapping [J]. The Plant Journal, 2020, 104 (4): 950 – 963.
- [22] ZHAO K Y, TUNG C W, EIZENGA G C, et al. Genome – wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa* [J]. Nature Communications, 2011, 2: 467.
- [23] MENG S, HE J B, ZHAO T J, et al. Detecting the QTL – allele system of seed isoflavone content in Chinese soybean landrace population for optimal cross design and gene system exploration [J]. Theoretical and Applied Genetics, 2016, 129 (8): 1557 – 1576.
- [24] KOUKOL J, CONN E E. The metabolism of aromatic compounds in higher plants. IV. Purification and properties of the phenylalanine deaminase of *Hordeum vulgare* [J]. The Journal of Biological Chemistry, 1961, 236 (10): 2692 – 2698.
- [25] YOUNG M R, NEISH A C. Properties of the ammonia – lyases deaminating phenylalanine and related compounds in *Triticum aestivum* and *Pteridium aquilinum* [J]. Phytochemistry, 1966, 5 (6): 1121 – 1132.
- [26] NAIR P M, VINING L C. Cinnamic acid hydroxylase in spinach [J]. Phytochemistry, 1965, 4 (1): 161 – 168.
- [27] FAHRENDORF T, DIXON R A. Stress responses in alfalfa (*Medicago sativa* L.). XVIII: molecular cloning and expression of the elicitor – inducible cinnamic acid 4 – hydroxylase cytochrome P450 [J]. Archives of Biochemistry and Biophysics, 1993, 305 (2): 509 – 515.
- [28] GARRY D, ALAN C. Plant – derived phenolic antioxidants [J]. Current Opinion in Clinical Nutrition and Metabolic Care, 2000, 3: 447 – 451.
- [29] HU W J, KAWAOKA A, TSAI C J, et al. Compartmentalized expression of two structurally and functionally distinct 4 – coumarate: CoA ligase genes in aspen (*Populus tremuloides*) [J]. Proceedings of the National Academy of the Sciences of the United States of America, 1998, 95 (9): 5407 – 5412.
- [30] LEE D, ELLARD M, WANNER L A, et al. The *Arabidopsis thaliana* 4 – coumarate: CoA ligase (*4CL*) gene: stress and developmentally regulated expression and nucleotide sequence of its cDNA [J]. Plant Molecular Biology, 1995, 28 (5): 871 – 884.
- [31] AUSTIN M B, NOEL J P. The chalcone synthase superfamily of type III polyketide synthases [J]. Natural Product Reports, 2003, 20 (1): 79 – 110.
- [32] NAPOLI C, LEMIEUX C, JORGENSEN R, et al. Introduction of a chimeric chalcone synthase gene into *Petunia* results in reversible co – suppression of homologous genes in trans [J]. The Plant Cell, 1990, 2 (4): 279 – 289.
- [33] MATHIEU M, ZEELEN J, PAUPTIT R A, et al. The 2.8 crystal structure of peroxisomal 3 – ketoacyl – CoA thiolase of *Saccharomyces cerevisiae*: a five – layered $\alpha\beta\alpha\beta\alpha$ structure constructed from two core domains of identical topology [J]. Structure, 1994, 2 (9): 797 – 808.
- [34] WELLE R, SCHRÖDER G, SCHILTZ E, et al. Induced plant responses to pathogen attack. Analysis and heterologous expression of the key enzyme in the biosynthesis of phytoalexins in soybean (*Glycine max* L. Merr. cv. Harosoy 63) [J]. European Journal of Biochemistry, 1991, 196 (2): 423 – 430.
- [35] XU R Y, NAN P, PAN H Y, et al. Molecular cloning, characterization and expression of a chalcone reductase gene from *Astragalus membranaceus* Bge. var. *mongholicus* (Bge.) Hsiao [J]. Molecular Biology Reports, 2012, 39 (3): 2275 – 2283.
- [36] 杨宏霞, 刘冰雁, 刘振坤, 等. ‘苹果梨’果实着色过程中 PyCHI 和 PyF3H 的克隆与表达分析 [J]. 果树学报, 2015, 32 (3): 359 – 365 + 521.
- YANG H X, LIU B Y, LIU Z K, et al. Molecular cloning and expression of PyCHI and PyF3H in Pingguoli during the pigmentation [J]. Journal of Fruit Science, 2015, 32 (3): 359 – 365 + 521.
- [37] WAKI T, MAMEDA R, NAKANO T, et al. A conserved strategy of chalcone isomerase – like protein to rectify promiscuous chalcone synthase specificity [J]. Nature Communications, 2020, 11 (1) 870.
- [38] PIČMANOVÁ M, REŇÁK D, FECIKOVÁ J, et al. Functional expression and subcellular localization of pea polymorphic isoflavone synthase CYP93C18 [J]. Biologia Plantarum, 2013, 57 (4): 635 – 645.
- [39] AKASHI T, AOKI T, AYABE S I. Cloning and functional expression of a cytochrome P450 cDNA encoding 2 – hydroxyisoflavanone synthase involved in biosynthesis of the isoflavonoid skeleton in licorice [J]. Plant Physiology, 1999, 121 (3): 821 – 828.
- [40] SUBRAMANIAN S, HU X, LU G H, et al. The promoters of two isoflavone synthase genes respond differentially to nodulation and defense signals in transgenic soybean roots [J]. Plant Molecular Biology, 2004, 54 (5): 623 – 639.
- [41] HOLLAND R, LEFF A P, PENNY W D, et al. Modulation of frontal effective connectivity during speech [J]. NeuroImage, 2016, 140: 126 – 133.
- [42] LIU C J, BLOUNT J W, STEELE C L, et al. Bottlenecks for metabolic engineering of isoflavone glycoconjugates in *Arabidopsis* [J]. Pro-

- ceedings of the National Academy of Sciences, 2002, 99 (22): 14578–14583.
- [43] KIM B G, KIM S Y, SONG H S, et al. Cloning and expression of the isoflavone synthase gene (*IFS-Tp*) from *Trifolium pratense* [J]. *Molecules and Cells*, 2003, 15 (3): 301–306.
- [44] STEELE C L, GIJZEN M, QUTOB D, et al. Molecular characterization of the enzyme catalyzing the aryl migration reaction of isoflavonoid biosynthesis in soybean [J]. *Archives of Biochemistry and Biophysics*, 1999, 367 (1): 146–150.
- [45] GUPTA O P, DAHUJA A, SACHDEV A, et al. Cytosine methylation of isoflavone synthase gene in the genic region positively regulates its expression and isoflavone biosynthesis in soybean seeds [J]. *DNA and Cell Biology*, 2019, 38 (6): 510–520.
- [46] SOHN S I, KIM Y H, KIM S L, et al. Genistein production in rice seed via transformation with soybean *IFS* genes [J]. *Plant Science*, 2014, 217–218: 27–35.
- [47] SOMMER H, BONAS U, SAEDLER H. Transposon-induced alterations in the promoter region affect transcription of the chalcone synthase gene of *Antirrhinum majus* [J]. *Molecular and General Genetics*, 1988, 211 (1): 49–55.
- [48] KIM S S, GRIENENBERGER E, LALLEMAND B, et al. *LAP6/POLYKETIDE SYNTHASE A* and *LAP5/POLYKETIDE SYNTHASE B* encode hydroxyalkyl α -pyrone synthases required for pollen development and sporopollenin biosynthesis in *Arabidopsis thaliana* [J]. *The Plant Cell*, 2010, 22 (12): 4045–4066.
- [49] XUAN L J, ZHANG C C, YAN T, et al. *TRANSPARENT TESTA 4*-mediated flavonoids negatively affect embryonic fatty acid biosynthesis in *Arabidopsis* [J]. *Plant, Cell and Environment*, 2018, 41 (12): 2773–2790.
- [50] KIKUCHI, SHOSHI, SATOH, et al. Collection, Mapping, and Annotation of Over 28, 000 cDNA Clones from Japonica rice [J]. *Science*, 2003, 301 (5631): 376–379.
- [51] ANGURAJVADIVEL A K, KRYSIAK K, TIAN G, et al. Genome-wide identification and localization of chalcone synthase family in soybean (*Glycine max* [L] Merr) [J]. *BMC Plant Biology*, 2018, 18 (1): 325.
- [52] LANZ T, TROPF S, MARNER F J, et al. The role of cysteines in polyketide synthases. site-directed mutagenesis of resveratrol and chalcone synthases, two key enzymes in different plant-specific pathways [J]. *Journal of Biological Chemistry*, 1991, 266 (15): 9971–9976.
- [53] DURBIN M L, MCCAIG B, CLEGG M T. Molecular evolution of the chalcone synthase multigene family in the morning glory genome [J]. *Plant Molecular Biology*, 2000, 42 (1): 79–92.
- [54] HARASHIMA S, TAKANO H, ONO K, et al. Chalcone synthase-like gene in the liverwort, *Marchantia paleacea* var. *diptera* [J]. *Plant Cell Reports*, 2004, 23 (3): 167–173.
- [55] REIMOLD U, KRÖGER M, KREUZALER F, et al. Coding and 3′non-coding nucleotide sequence of chalcone synthase mRNA and assignment of amino acid sequence of the enzyme [J]. *The EMBO Journal*, 1983, 2 (10): 1801–1805.
- [56] HELARIUTTA Y, KOTILAINEN M, ELOMAA P, et al. Duplication and functional divergence in the chalcone synthase gene family of *Asteraceae*: evolution with substrate change and catalytic simplification [J]. *Proceedings of the National Academy of Sciences*, 1996, 93 (17): 9033–9038.
- [57] DHAUBHADEL S, GIJZEN M, MOY P, et al. Transcriptome analysis reveals a critical role of *CHS7* and *CHS8* genes for isoflavonoid synthesis in soybean seeds [J]. *Plant Physiology*, 2007, 143 (1): 326–338.
- [58] LI J R, TODD T C, TRICK H N. Rapid in planta evaluation of root expressed transgenes in chimeric soybean plants [J]. *Plant Cell Reports*, 2010, 29 (2): 113–123.
- [59] KAO T H, CHEN B H. Functional components in soybean cake and their effects on antioxidant activity [J]. *Journal of Agricultural and Food Chemistry*, 2006, 54 (20): 7544–7555.
- [60] WANG X Q. Structure, function, and engineering of enzymes in isoflavonoid biosynthesis [J]. *Functional and Integrative Genomics*, 2011, 11 (1): 13–22.
- [61] 吴楠, 王丕武, 李丹, 等. 大豆查尔酮还原酶基因 *CHRI* 的功能研究 [J]. *遗传*, 2014, 36 (7): 707–712.
- WU N, WANG P W, LI D, et al. Function of chalcone reductase gene *CHRI* in soybean [J]. *Hereditas*, 2014, 36 (7): 707–712.
- [62] MEHDY M C, LAMB C J. Chalcone isomerase cDNA cloning and mRNA induction by fungal elicitor, wounding and infection [J]. *The EMBO Journal*, 1987, 6 (6): 1527–1533.
- [63] DRUKA A, KUDRNA D, ROSTOKS N, et al. Chalcone isomerase gene from rice (*Oryza sativa*) and barley (*Hordeum vulgare*): physical,

- genetic and mutation mapping [J]. *Gene*, 2003, 302 (1/2): 171–178.
- [64] SCHIJLEN E, VOS C H R D, JONKER H, et al. Pathway engineering for healthy phytochemicals leading to the production of novel flavonoids in tomato fruit [J]. *Plant Biotechnology Journal*, 2006, 4 (4): 433–444.
- [65] WOOD A J, DAVIES E. A cDNA encoding chalcone isomerase from aged pea epicotyls [J]. *Plant Physiology*, 1994, 104 (4): 1465–1466.
- [66] CRAMER C L, EDWARDS K, DRON M, et al. Phenylalanine ammonia-lyase gene organization and structure [J]. *Plant Molecular Biology*, 1989, 12 (4): 367–383.
- [67] RICKEY T M, BELKNAP W R. Comparison of the expression of several stress-responsive genes in potato tubers [J]. *Plant Molecular Biology*, 1991, 16 (6): 1009–1018.
- [68] GOWRI G, PAIVA N L, DIXON R A. Stress responses in alfalfa (*Medicago sativa* L.) 12. Sequence analysis of phenylalanine ammonia-lyase (PAL) cDNA clones and appearance of PAL transcripts in elicitor-treated cell cultures and developing plants [J]. *Plant Molecular Biology*, 1991, 17 (3): 415–429.
- [69] DIALLINAS G, KANELIS A K. A phenylalanine ammonia-lyase gene from melon fruit: cDNA cloning, sequence and expression in response to development and wounding [J]. *Plant Molecular Biology*, 1994, 26 (1): 473–479.
- [70] DAVID J P, BOYER S, MESNEAU A, et al. Involvement of cytochrome P450 monooxygenases in the response of mosquito larvae to dietary plant xenobiotics [J]. *Insect Biochemistry and Molecular Biology*, 2006, 36 (5): 410–420.
- [71] NOMURA T, BISHOP G J. Cytochrome P450s in plant steroid hormone synthesis and metabolism [J]. *Phytochemistry Reviews*, 2006, 5 (2/3): 421–432.
- [72] BOURGAUD F, HEHN A, LARBAT R, et al. Biosynthesis of coumarins in plants: a major pathway still to be unravelled for cytochrome P450 enzymes [J]. *Phytochemistry Reviews*, 2006, 5 (2/3): 293–308.
- [73] HU W J, KAWAOKA A, TSAI C J, et al. Compartmentalized expression of two structurally and functionally distinct 4-coumarate: CoA ligase genes in aspen (*Populus tremuloides*) [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1998, 95 (9): 5407–5412.
- [74] SCHNEIDER K, HÖVEL K, WITZEL K, et al. The substrate specificity – determining amino acid code of 4-coumarate: CoA ligase [J]. *Proceedings of the National Academy of Sciences*, 2003, 100 (14): 8601–8606.
- [75] AHMAD I, MENG X P, KAMRAN M, et al. Effects of uniconazole with or without micronutrient on the lignin biosynthesis, lodging resistance, and winter wheat production in semiarid regions [J]. *Journal of Integrative Agriculture*, 2020, 19 (1): 62–77.
- [76] TOMBULOGLU H. Genome-wide identification and expression analysis of R2R3, 3R- and 4R-MYB transcription factors during lignin biosynthesis in flax (*Linum usitatissimum*) [J]. *Genomics*, 2020, 112 (1): 782–795.
- [77] YU O, SHI J, HESSION A O, et al. Metabolic engineering to increase isoflavone biosynthesis in soybean seed [J]. *Phytochemistry*, 2003, 63 (7): 753–763.
- [78] KLEMPNAUER K H, GONDA T J, BISHOP M J. Nucleotide sequence of the retroviral leukemia gene *v-myb* and its cellular progenitor *c-myb*: the architecture of a transduced oncogene [J]. *Cell*, 1982, 31: 453–463.
- [79] BIEDENKAPP H, BORGMAYER U, SIPPEL A E, et al. Viral *myb* oncogene encodes a sequence-specific DNA-binding activity [J]. *Nature*, 1988, 335 (6193): 835–837.
- [80] COLBY W W, CHEN E Y, SMITH D H, et al. Identification and nucleotide sequence of a human locus homologous to the *v-myc* oncogene of avian myelocytomatosis virus MC29 [J]. *Nature*, 1983, 301 (5902): 722–725.
- [81] KLEMPNAUER K H, RAMSAY G, BISHOP J M, et al. The product of the retroviral transforming gene *v-myb* is a truncated version of the protein encoded by the cellular oncogene *c-myb* [J]. *Cell*, 1983, 33 (2): 345–55.
- [82] LEPINIEC L, DEBEAUJON I, ROUTABOUL J M, et al. Genetics and biochemistry of seed flavonoids [J]. *Annual Review of Plant Biology*, 2006, 57 (1): 405–430.
- [83] ALITALO K, RAMSAY G, BISHOP J M, et al. Identification of nuclear proteins encoded by viral and cellular *myc* oncogenes [J]. *Nature*, 1983, 306 (5940): 274–277.
- [84] LIPSICK J S. One billion years of Myb [J]. *Oncogene*, 1996, 13 (2): 223–235.
- [85] WANG B H, LUO Q, LI Y P, et al. Structural insights into target DNA recognition by R2R3-MYB transcription factors [J]. *Nucleic*

- Acids Research, 2020, 48 (1): 460–471.
- [86] ZHONG R Q, RICHARDSON E A, YE Z H. The MYB46 transcription factor Is a direct target of SND1 and regulates secondary wall biosynthesis in *Arabidopsis* [J]. The Plant Cell, 2007, 19 (9): 2776–2792.
- [87] DU H, YANG S S, LIANG Z, et al. Genome – wide analysis of the MYB transcription factor superfamily in soybean [J]. BMC Plant Biology, 2012, 12 (1): 106–127.
- [88] STRACKE R, ISHIHARA H, HUEP G, et al. Differential regulation of closely related R2R3 - MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling [J]. The Plant Journal, 2007, 50 (4): 660–677.
- [89] FRANK B, JACQUELINE G, JOHN M. Conifer R2R3 – MYB transcription factors: sequence analyses and gene expression in wood – forming tissues of white spruce (*Picea glauca*) [J]. BMC Plant Biology, 2007, 7 (1): 17–33.
- [90] PAZ – ARES J, GHOSAL D, WIENAND U, et al. The regulatory *c1* locus of *Zea mays* encodes a protein with homology to *myb* proto – onco-gene products and with structural similarities to transcriptional activators [J]. The EMBO Journal, 1987, 6 (12): 3553–3558.
- [91] JIN H, COMINELLI E, BAILEY P, et al. Transcriptional repression by *AtMYB4* controls production of UV – protecting sunscreens in *Arabi-dopsis* [J]. The EMBO Journal, 2000, 19 (22): 6150–6161.
- [92] ASAPH A, RIC DE VOS C H, MARTINA W, et al. The strawberry *FaMYB1* transcription factor suppresses anthocyanin and flavonol accu-mulation in transgenic tobacco [J]. The Plant Journal, 2001, 28 (3): 319–332.
- [93] YI J, DERYNCK M R, CHEN L, et al. Differential expression of *CHS7* and *CHS8* genes in soybean [J]. Planta, 2010, 231 (3): 741–753.
- [94] LI X W, LI J W, ZHAI Y, et al. A R2R3 – MYB transcription factor, *GmMYB12B2*, affects the expression levels of flavonoid biosynthesis genes encoding key enzymes in transgenic *Arabidopsis* plants [J]. Gene, 2013, 532 (1): 72–79.
- [95] YI J X, Derynck M R, Li X Y, et al. A single – repeat MYB transcription factor, *GmMYB176*, regulates *CHS8* gene expression and affects isoflavonoid biosynthesis in soybean [J]. The Plant Journal, 2010, 62 (6): 1019–1034.
- [96] MURRE C, MCCA W P S, BALTIMORE D. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, *daughterless*, *MyoD*, and *myc* proteins [J]. Cell, 1989, 56 (5): 777–783.
- [97] ATCHLEY W R, FITCH W M. A natural classification of the basic Helix – Loop – Helix class of transcription factors [J]. Proceedings of the National Academy of Sciences of the United States of America, 1997, 94 (10): 5172–5176.
- [98] ATCHLEY W R, TERHALLE W, DRESS A. Positional dependence, cliques, and predictive motifs in the bHLH protein domain [J]. Journal of Molecular Evolution, 1999, 48 (5): 501–516.
- [99] SAMMAK S, HAMDANI N, GORREC F, et al. Crystal structures and nuclear magnetic resonance studies of the apo form of the c – MYC: MAX bHLHZip complex reveal a helical basic region in the absence of DNA [J]. Biochemistry, 2019, 58 (29): 3144–3154.
- [100] TOLEDO – ORTIZ G, HUQ E, QUAIL P H. The *Arabidopsis* basic/Helix – Loop – Helix transcription factor family [J]. The Plant Cell, 2003, 15 (8): 1749–1770.
- [101] LI X X, DUAN X P, JIANG H X, et al. Genome – wide analysis of basic/Helix – Loop – Helix transcription factor family in rice and *Arabi-dopsis* [J]. Plant Physiology, 2006, 141 (4): 1167–1184.
- [102] FELLER A, MACHEMER K, BRAUN E L, et al. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors [J]. Plant Journal, 2011, 66 (1): 94–116.
- [103] PATTANAIK S, XIE C H, YUAN L. The interaction domains of the plant Myc – like bHLH transcription factors can regulate the transactiva-tion strength [J]. Planta, 2008, 227 (3): 707–715.
- [104] ZIMMERMANN I M, HEIM MA, WEISSHAAR B, et al. Comprehensive identification of *Arabidopsis thaliana* MYB transcription factors in-teracting with R/B - like BHLH proteins [J]. The Plant Journal, 2004, 40 (1): 22–34.
- [105] ZHANG F, GONZALEZ A, ZHAO M Z. A network of redundant bHLH proteins functions in all TTG1 – dependent pathways of *Arabidopsis* [J]. Development, 2003, 130 (20): 4859–4869.
- [106] PAYNE C T, ZHANG F, LLOYD A M. *GL3* encodes a bHLH protein that regulates trichome development in *Arabidopsis* through interaction with GL1 and TTG1 [J]. Genetics, 2000, 156 (3): 1349–1362.
- [107] LEE M M, SCHIEFELBEIN J. WEREWOLF, a MYB – related protein in *Arabidopsis*, is a position – dependent regulator of epidermal cell patterning. [J]. Cell, 1999, 99 (5): 473–483.

[108] BAUDRY A, HEIM M A, DUBREUCQ B, et al. TT2, TT8, and TTG1 synergistically specify the expression of *BANYULS* and proanthocyanidin biosynthesis in *Arabidopsis thaliana* [J]. The Plant Journal, 2004, 39 (3): 366 – 380.

[109] GOFF S A, CONE K C, CHANDLER V L. Functional analysis of the transcriptional activator encoded by the maize B gene: evidence for a direct functional interaction between two classes of regulatory proteins. [J]. Genes and Development, 1992, 6 (5): 864 – 875.

[110] HERNANDEZ J M, HEINE G F, IRANI N G, et al. Different mechanisms participate in the R – dependent activity of the R2R3 MYB transcription factor C1. [J]. Journal of Biological Chemistry, 2004, 279 (46): 48205 – 48213.

[111] FONG H K W, HURLEY J B, HOPKINS R S, et al. Repetitive segmental structure of the transducin beta subunit: homology with the *CDC4* gene and identification of related mRNAs. [J]. Proceedings of the National Academy of Sciences, 1986, 83 (7): 2162 – 2166.

[112] STIRNIMANN C U, PETSALAKI E, RUSSELL R B, et al. WD40 proteins propel cellular networks [J]. Trends in Biochemical Sciences, 2010, 35 (10): 565 – 574.

[113] VAN NOCKER S, LUDWIG P. The WD – repeat protein superfamily in *Arabidopsis*: conservation and divergence in structure and function [J]. BMC Genomics, 2003, 4: 50.

[114] OUYANG Y, HUANG X, LU Z, et al. Genomic survey, expression profile and co – expression network analysis of *OsWD40* family in rice [J]. BMC Genomics, 2012, 13 (1): 100 – 114.

[115] WALL M A, COLEMAN D E, LEE E, et al. The structure of the G protein heterotrimer $G\alpha_1\beta_1\gamma_2$ [J]. Cell, 1995, 83 (6): 1047 – 1058.

[116] MYLONA A, FERNANDEZ – TORNERO C, LEGRAND P, et al. Structure of the $\tau 60/\Delta\tau 91$ subcomplex of yeast transcription factor iiic: insights into preinitiation complex assembly [J]. Molecular Cell, 2006, 24 (2): 221 – 232.

[117] RUTHENBURG A J, WANG W, GRAYBOSCH D M, et al. Histone H3 recognition and presentation by the WDR5 module of the MLL1 complex [J]. Nature Structural and Molecular Biology, 2006, 13 (8): 704 – 712.

[118] HU X J, LI T, WANG Y, et al. Prokaryotic and highly – repetitive WD40 proteins: a systematic study [J]. Scientific Reports, 2017, 7 (1): 10585 – 10597.

[119] SENISTERRA G, WU H, ALLALI – HASSANI A, et al. Small – molecule inhibition of MLL activity by disruption of its interaction with WDR5 [J]. Biochemical Journal, 2013, 449 (1): 151 – 159.

[120] LI Q, ZHAO P P, LI J, et al. Genome – wide analysis of the WD – repeat protein family in cucumber and *Arabidopsis* [J]. Molecular Genetics and Genomics, 2014, 289 (1): 103 – 124.

[121] LI S T. Transcriptional control of flavonoid biosynthesis: fine – tuning of the MYB – bHLH – WD40 (MBW) complex [J]. Plant Signaling and Behavior, 2014, 9 (1): e27522.