

真双子叶基部类群十种植物的 CHS 基因家族成员分析

杨俊波, 杨汉奇, 李德铎, 李洪涛

(中国科学院昆明植物研究所生物多样性和生物地理学重点实验室, 云南 昆明 650204)

摘要: 用 PCR 方法从目前极少被报道的真双子叶基部类群 10 种植物的总 DNA 中扩增出 CHS 基因外显子 2 的部分序列进行分析, 经克隆后测序, 共得到 26 个不同的片段。分析表明, 所有序列具有较高的同源性 (> 70%), 每种植物有 2~4 个不同的拷贝, 在金鱼藻中发现有一个拷贝 (CedeCHS3) 含有一段长 29 个碱基的缺失, 可能已失去了该基因的功能, 而来自银桦的 GrrCHS8 拷贝具有较多活性位点的变异, 可能具有了新的功能。在对碱基含量的分析中表明, 只有来自雀舌黄杨的序列有一定的 GC 偏好, 特别是第三位的 GC 含量达 70% 以上。用贝叶斯法、最大简约法和邻接法构建的 CHS 基因的分子系统树均由 3 个主要分支构成, 其中一个分支由来自金鱼藻的 4 个序列聚在一起构成了 subfamily , 来自昆栏树的 2 个序列聚在一起网接于 subfamily 中; 来自独叶草的 3 个序列分散于同一亚家族 (subfamily) 的不同分支中; 而多数种的序列分散在相距较远的两个分支中, 构成了两个主要的亚家族 subfamily 、subfamily 。从所建分子系统树看真双子叶基部类群植物的 CHS 基因家族成员来源于该类群形成前的两个祖先拷贝, 在不同植物这两个祖先拷贝经历了不同的进化过程, 这些进化上的差异可能与不同植物的生活习性及其生存环境的多样性相关。

关键词: 真双子叶基部类群; 查尔酮合酶基因家族; 进化; 重复

中图分类号: Q 945

文献标识码: A

文章编号: 0253-2700(2006)04-352-07

An Analysis of CHS Gene Family for Some Basal Eudicots*

YANG Jun-Bo, YANG Han-Qi, LI De-Zhu**, LI Hong-Tao

(Laboratory of Biodiversity and Biogeography, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650204, China)

Abstract: Partial sequences of chalcone synthase (CHS) gene exon 2 were amplified by PCR method from genomic DNA of 10 species of the basal eudicots. After cloning and sequencing, 26 sequences were obtained. There were two to four copies of CHS gene family among various plants. All of sequences showed a similarity higher than 70% at the nucleotide level. A 29-base deletion was found in a copy (CedeCHS3) in *Ceratophyllum demersum*, which indicated that the function of the CHS enzyme in this species was lost. GrrCHS8, a copy from *Grevillea robusta*, had many mutations at active sites, implying the new function may be developed in this copy. The result of analyzing the base content revealed that only the sequences from *Buxus bobinieri* have GC bias, especially at the third base of codes, and the GC content is more than 70%. All of molecular phylogenetic trees, based on the Bayesian, MP, and NJ methods, consisted of three major clades. Four sequences from *Ceratophyllum demersum* were grouped into a monophyletic clade (subfamily). Two sequences out of *Trochodendron aralioides* nested in subfamily . Another three sequences from *Kingdonia uniflora* interspersed among different subclades of subfamily . Sequences of most species dispersed in subfamily and subfamily . And it revealed that CHS gene family derived from two ancestor copies, which had existed before eudicots were evolved. The two copies experienced different evolutionary process among different species. These differences related to the diversity

基金项目: 科技部国家科技基础条件平台工作项目 (项目编号 grant No: 2004DKA30430), 吴征镒先生云南省突出贡献奖 (2001 年度) 资助项目 (项目编号 grant No: KIB-WU-2001-03)

通信作者: Author for correspondence. E-mail: DZL@mail.kib.ac.cn

收稿日期: 2005-10-19, 2006-01-20 接受发表

作者简介: 杨俊波 (1970-) 男, 云南人, 高级工程师, 主要从事植物分子进化和分子系统学研究。

of habits and environments of these plants .

Key words: Basal eudicots; CHS gene family; Evolution; Duplication

查尔酮合酶 (Chalcone Synthase, CHS) 是类黄酮物质生物合成途径的关键酶, 它催化 3 个分子的丙酰 CoA (Malonyl CoA) 与 1 个分子的香豆酰 CoA (Coumaryl CoA) 生成查尔酮。CHS 基因是一个较大的多基因家族, 其编码区比较保守, 长约 1.2 kb, 它的家族成员有查尔酮合酶基因, 1, 2-二苯乙烯合酶基因及参与雄蕊发育的基因 (Helasiutta 等, 1996; Qu 等, 1997), 科间的氨基酸同源性在 70% ~ 90%。CHS 基因编码区被不同长度的内含子分隔成外显子 1 和外显子 2, 其中外显子 1 较短, 只编码约 60 个氨基酸, 且长度变异较大; 外显子 2 编码约 340 个氨基酸, 在进化中较为保守, 易于排序, 提供的进化信息较多。Ursula 等 (1987) 的研究认为可用 CHS 基因外显子 2 代表全基因进行研究, Wang 等 (2000) 的研究得出了同样结果。

由于类黄酮物质在植物中具有很多重要的生物功能, 例如与花色的形成, 防 UV 损伤, 抗病, 花粉育性等有关, 在豆科植物中与根瘤的形成有关, 等等。在上个世纪 60 ~ 70 年代就有人开始有对该基因家族进行研究, 但绝大多数研究限于植物生理、生化及亲缘关系的探讨, 且多集中于个别科属中 (Durbin, 1995; Helariutta, 1996; Tropf, 1994; Ferrer, 1999), 这些研究揭示了该基因的结构、化学性质、功能位点等, 为 CHS 基因家族的分子进化及系统发育研究打下了坚实基础, 到上个世纪末本世纪初, 随着分子系统学的发展和被报道的 CHS 基因家族序列的增加, 对该基因家族的进化及系统发育研究也不断出现 (Yang 等, 2004; Yang 等, 2003; Yang 等, 2002; Wang 等, 2000), 这些研究在一定程度上揭示了该基因家族的一些进化特点和规律, 对其在系统发育学应用中的意义也有所讨论。但目前为止只有不到 200 个 CHS 基因家族序列被报道, 且在植物类群中的分布极不均匀。从 APG 系统 (1998; 2003) 看, 目前研究涉及到的该基因家族序列集中于 APG 系统中的 core eudicots 的 rosids 和 asterids 及 monocots 的 commelinids 分枝的少数科的少数种内。在整个被子植物的基部类群

(Amborellaceae、Nymphaeaceae、Austrobaileyales、Chloranthaceae、Magnoliids) 中只有 2 个种的 5 条序列, 而在真双子叶的基部类群 (Ranunculales、Sabiaceae、Proteales、Buxaceae、Trochodendraceae) 中则只有 1 个种的 2 条序列被报道。而这些类群都是被子植物研究中极其重要和倍受关注的。由于这些类群资料的缺少, 我们对该基因家族在整个被子植物中的进化和系统发育的认识还是很有限的。

本文用 PCR 方法, 从目前研究极少涉及的真双子叶的基部类群 8 科 9 种植物以及系统位置尚不确定的基部被子植物金鱼藻的总 DNA 中扩增出 CHS 基因外显子 2 的部分序列, 经克隆后测序, 对不同植物的 CHS 基因的拷贝数、不同成员的同源性、进化方式等进行分子系统学分析, 并与已有的研究进行对比, 为全面揭示该基因家族在整个被子植物中的进化全貌积累资料。

1 材料与方法

1.1 材料

用于 PCR 扩增的 10 种植物, 代表了 APGII 系统中除 Sabiaceae 外的真双子叶植物基部类群的每一个大分支的植物, 其中金鱼藻为基部被子植物, 系统位置尚待进一步研究确定。其采集点及存放处见表 1。

1.2 方法

1.2.1 DNA 提取和 PCR 扩增 采用改良的 CTAB 法 (Doyle and Doyle, 1987) 从新鲜的或硅胶干燥的叶片中提取植物的总 DNA 用于 PCR 扩增。PCR 反应采用 Wang 等 (2000) 设计的兼并引物在 Perkin-Elmer 公司的 9600 型 PCR 仪上进行, 总反应体积为 50 μ l。由于引物的兼并性和植物中类 CHS 基因的存在, 用 PCR 方法从植物总 DNA 中往往扩增到多带产物, 并有较多的引物二聚体。通过增加反应循环数, 调整引物用量, 采用热启动 PCR 反应等办法, 能较大改善扩增效果, 得到以目的产物为主的 PCR 产物。PCR 产物采用华舜公司的纯化试剂盒纯化后, 克隆到 pGEM-T 克隆载体系统中, 转入 DH5 菌培养后进行阳性克隆的鉴定和筛选。

1.2.2 阳性克隆的鉴定和筛选 利用一种基于 PCR 技术的快速筛选法对阳性克隆斑进行鉴定和筛选 (Yang 等, 2003), 为了得到尽可能多的 CHS 家族成员, 每种植物至少进行 24 个白斑的筛选, 其中至少对 12 个阳性克隆

进行序列测定, 每一个拷贝至少有一个重复序列, 在分析时序列相似性在 99% 以上且来自同一植物的聚在一起的序列只用其一作代表。每一个序列的名称由该植物的属名的前两个字母加上种加词的前两个字母组成, 末尾的数字为克隆号。

1.2.3 DNA 序列测定 经过鉴定和筛选的长度为 860bp 左右的 PCR 产物, 用华舜公司的纯化试剂盒纯化后, 用 Perkin-Elmer 公司的 ABI PRISM^(tm) Bigdye Terminator Cycle Sequencing Ready Reaction Kit 进行测序反应, 测序反应物经纯化后在 ABI310 或 ABI3700 上检测。

1.2.4 序列处理和系统进化分析 测序得到的双向序列用 DNASTAR 软件包的 Seqman 程序组装和校对后, 用 CLUSTAL W (Version 1.81) (Thompson 等, 1994) 软件进行排序, 并根据氨基酸序列进行适当调整。

分别用贝叶斯法 (Bayesian) (Ramala and Yang 1996;

Huelsenbeck and Ronquist, 2001), 最大简约法 (MP) 和基于碱基距离的邻接法 (NJ) (Saitou and Nei, 1987) 构建 CHS 基因的系统进化树。贝叶斯法采用最新版本的 Mr-Bayesian V 3.01, 使用 GTR + I + 为替换模型, 位点间变异模型 rates = invgamma, 两次独立的运算同时进行, 马尔科夫链的蒙特卡洛方法 (Markov chain Monte Carlo process) 设置为四条链同时进行, 3 条热链和 1 条冷链, 以随机树为起始树, 运行 2 000 000 代, 每 100 代取一棵树, 在舍去了前 2000 棵树后, 根据剩余的样本构建一致树 (Consensus tree) 并计算相关参数。MP 和 NJ 法的系统进化树在 PAUP 软件 (Version 4.0b8) (Swofford, 2001) 上完成。MP 树的构建采用 TBR 模型用渐进式方法进行搜索 (Heuristic search procedures), NJ 树使用 HK Y85 距离法构建。系统树各分支的置信度以 1000 次自引导法 (Bootstrap) 重复检测。

表 1 本研究的材料来源和 CHS 基因的 GC 含量及 GenBank 接收号

Table 1 Plant materials investigated, GC content and GenBank accession numbers used in this study

Family	Species	Voucher	Locality	GenBank ID	GenBank AC number	G + C content
Ceratophyllaceae	<i>Ceratophyllum demersum</i>	Yang2005011	Kunming, Yunnan	CedeCHS1	DQ366590	49.65
				CedeCHS2	DQ366589	51.16
				CedeCHS3	DQ366588	51.67
				CedeCHS5	DQ366587	51.05
Lardizabalaceae	<i>Akebia quinata</i>	WF094	Zhangjiajie, Hunan	AkquCHS1	DQ366570	47.32
				AkquCHS4	DQ366595	47.09
	<i>Holboellia latifolia</i>	Yang2005012	Kunming, Yunnan	HolaCHS2	DQ366582	46.50
				HolaCHS3	DQ366581	46.21
Sargentodoxaceae	<i>Sargentodoxa cuneata</i>	WF090	Pingbian, Yunnan	SacuCHS1	DQ366576	51.99
				SacuCHS4	DQ366575	50.93
Ranunculaceae	<i>Kingdonia uniflora</i>	Liu Xiao (PE)	Mt. Emei, Sichuan	KiunCHS4	DQ366579	48.36
				KiunCHS5	DQ366578	49.07
				KiunCHS7	DQ366577	49.89
Eupteleaceae	<i>Euptelea pleiosperma</i>	Yang2005013	Kunming, Yunnan	EuplCHS7	DQ366586	51.87
				EuplCHS8	DQ366585	47.20
Proteaceae	<i>Grevillea robusta</i>	Yang2005014	Kunming, Yunnan	GrrroCHS5	DQ366584	51.89
				GrrroCHS8	DQ366583	46.46
Buxaceae	<i>Buxus bobinieri</i>	Yang2005015	Kunming, Yunnan	BuboCHS1	DQ366593	58.39
				BuboCHS2	DQ366592	61.66
				BuboCHS5	DQ366591	58.74
Trochodendraceae	<i>Trochodendron aralioides</i>	Yang2005016	Kunming, Yunnan	TrarCHS1	DQ366572	54.66
				TrarCHS10	DQ366571	54.78
Tetracentraceae	<i>Tetracentron sinense</i>			TesiCHS6	DQ366574	52.10
				TesiCHS10	DQ366573	53.96
Ranunculaceae	<i>Ranunculus acer</i>	Ursula (1987)		RA1	Ursula (1987)	46.78
				RA2		45.92
Pinaceae	<i>Pinus sylvestris</i>	Fliegmann <i>et al</i> (1992)		PSCHS	X60754	55.47

2 结果

2.1 PCR 扩增及克隆结果

利用 Wang 等 (2000) 设计的兼并引物用 PCR 的方法从真双子叶基部类群 10 种植物中扩增到了 CHS 基因外显子 2 的部分片段, 经克隆

后, 每种植物至少测定 12 个阳性克隆斑, 共得到 26 个不同的 CHS 基因外显子 2 的序列, 且每个序列至少有一个相同的片段。所有被测定的序列都在 GenBank 中注册 (表 1)。

2.2 核苷酸及其推定的氨基酸序列的分析

26 个不同的 CHS 基因外显子 2 的序列经 Clustal W 排序，表明所有的序列有较高的相似性 (>70%)。序列 AkquCHS6 含一个连续的 6 碱基的缺失，序列 CedeCHS1、CedeCHS2、CedeCHS5 在相同位置有一个连续的 3 碱基缺失，GrrOCHS8 含一个 3 碱基的插入，而 HolaCHS3 在序列的开始处有一个连续的 27 个碱基的缺失，只有 CedeCHS3 由于较长碱基的缺失 (29 个)，可能已失去了该基因的功能。其余序列长 861 bp，所有序列排序后长 864 bp。这些序列占整个 CHS 基因外显子 2 的 90%，并包含了所有外显子 2 的活性位点。来自不同植物的不同序列的碱基使用有一定的差异，其中来自 *Buxus bobinieri* 的 BuboCHS2 序列明显偏好 GC (61.66%)，特别是第三位 GC 含量高达 70% 以上。而来自 *Holboellia latifolia* 的 HolaCHS2、

HolaCHS3 等序列具有较低的 GC 含量 (表 1)。

为了检测 CHS 基因的功能，所有被测定的序列用 DNASTAR 软件中的 EditSeq 程序转化为氨基酸序列。其中序列 CedeCHS3 由于较长的非 3 的倍数的碱基缺失使得读码框发生了变化而没有氨基酸序列进行分析，其余绝大多数序列被转化为 287 个氨基酸残基组成的多肽序列，对应于核苷酸的缺失，序列 CedeCHS1、CedeCHS2、CedeCHS5 有 286 个氨基酸残基 GrrOCHS8 有 288 个，HolaCHS3 有 277 个氨基酸残基组成。所有被测定的序列具有相似的氨基酸组成，其中 leucine、glycine、alanine 含量最高，cysteine、tyrosine、tryptophan 含量最低。CHS 基因外显子 2 上的 4 个催化位点在所有的序列中均未发生变异，但一些序列的活性位点有不同程度的改变 (图 1, 2)。

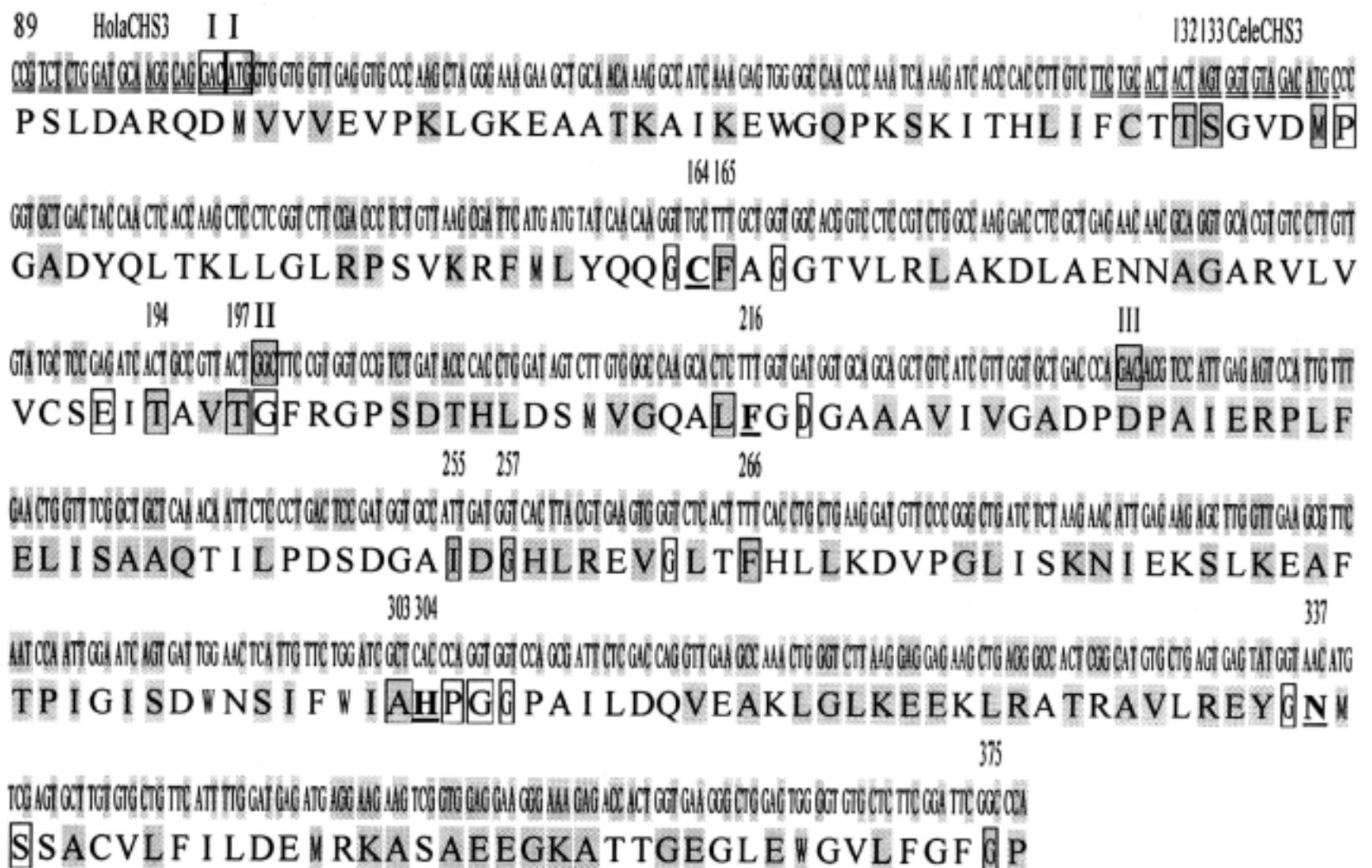


图 1 9 种真双子叶植物基部类群植物及金鱼藻 CHS 基因外显子 2 的核苷酸一致序列及推导的氨基酸一致序列。阴影表示有变异的位点，框住的位点表示 CHS 基因的活性位点，有下划线的位点是该酶的催化位点，双下划线的位点是该序列 (序列号标于其上方) 缺失的位点。序列上方的阿拉伯数字表示该位点在整个酶的一级结构中的位置，罗马数字表示有核苷酸插入 缺失的位点。其中 表示的是序列 AkquCHS6 的缺失位点； 表示的是序列 GrrOCHS8 的插入位点； 表示的是序列 CeleCHS1、2、5 共同的缺失位点。

Fig. 1 The consensus sequence of CHS exon 2 from nine basal eudicots and *Ceratophyllum demersum* and reference amino acid sequences. Variable sites are in the shadow; and active sites and catalytic sites are boxed and underlined, respectively. The longer deleted sites are double underline with sequence number above the sequence. Arabic numerals above the sequence denote the position of certain amino acids in primary structure of the CHS enzyme. Roman numerals denote the positions of single indel site. Among them denotes deleted sites of AkquCHS6, denotes inserted sites of GrrOCHS8, denotes deleted sites of CeleCHS1, 2, 5, respectively.

少数几个序列如：KiunCHS4、SacuCHS4 的位置不是很稳定，且利用贝叶斯法得到的基因树具有较高的支持率。来自不同植物的序列在基因树上的分布有较大的差异，如：来自金鱼藻 (*Ceratophyllum demersum*) 的 4 个序列聚为一支构成了 subfamily ，来自昆栏树 (*Trochodendron aralioides*) 的 2 个序列聚在和和其它种的序列一起网接于 subfamily 中；来自独叶草 (*Kingdonia uniflora*) 的 3 个序列分散于同一亚家族的不同分支中；而多数种的序列分散在相距较远的两个分支中，构成了两个主要的亚家族 subfamily 、subfamily 。这些差异在一定程度上预示了不同植物的 CHS 基因在进化中有很大大差异。

3 讨论

已有的研究表明，最早的 CHS 基因可能在苔藓类植物出现前就已经存在，其重复事件可能在原始的维管植物中就已经发生 (Huang 等, 2004)，形成了一个多基因家族。基因重复被认为是 CHS 基因家族进化和产生新功能基因的主要机制 (Ohta, 1993; Force 等, 1999)，通过基因的重复大多数植物中均有多个拷贝，其中豆科植物的拷贝最多，但在欧芹 (Herrmann 等, 1988)、拟南芥 (Feinbaum 等, 1988)、金鱼草 (Wienand 等, 1982) 中似乎只有一个拷贝。我们的研究表明，在真双子叶基部类群中每种植物有 2~4 个不同的拷贝，其中至少有一个拷贝具有典型的 CHS 基因功能，即所有的活性位点和催化位点都没有变异，绝大多数序列属于这一类型。在 subfamily 中发现了一些功能可能已经改变的序列。如：来自金鱼藻 (*Ceratophyllum demersum*) 的 CedeCHS3 由于较长碱基的缺失 (29 个) 导致了读码框改变而成为假基因，这种含较长碱基缺失的类 CHS 序列在以往的研究中很少有报道 (Yang 等, 2002)；来自银桦 (*Grevillea robusta*) 的 GrroCHS8 由于较多活性位点的改变，可能导致了类似 CHS 基因的新基因的产生；而来自木通 (*Akebia quinata*) 的 AkquCHS4 和五风藤 (*Holboellia latifolia*) 的 HolaCHS2 序列分别只有 1 个和 2 个活性位点发生改变，它们对该基因功能的影响较难确定。

由于被子植物和 CHS 基因家族功能的多样

性以及植物所经历的漫长演化历史使得几乎每一个已被研究的植物都有其独特的进化式样。Huang 等 (2004) 的研究把被子植物 CHS 基因家族的进化分为 4 种式样两个趋势。与已有的研究相似，在真双子叶基部类群不同的植物中 CHS 基因的拷贝数、碱基偏好、进化式样等都有较大差异。表现出了 3 种进化式样和 2 个趋势。第一种趋势是同一种植物的序列有很高的相似性和相同的碱基替换率，它们紧密地聚为一个单系，并有很高的支持率，如：来自金鱼藻 (*Ceratophyllum demersum*) 的 4 个序列聚为一支构成了 subfamily ，来自昆栏树 (*Trochodendron aralioides*) 的 2 个序列聚在一起网接于 subfamily 中；另一种趋势是同一种植物的不同序列的相似性较低，并不聚在一起形成一个单系，而是与其他种的序列一起形成多个分支，如：来自独叶草 (*Kingdonia uniflora*) 的 3 个序列分散于同一亚家族的不同分支中；而多数种的序列分散在相距较远的两个分支中，构成了两个主要的亚家族 subfamily 、subfamily 。这些差异在一定程度上反映了不同植物的 CHS 基因在进化中差异。

从所建的系统发育树上，真双子叶基部类群植物的 CHS 基因家族成员来源于该类群形成前的两个祖先拷贝，在不同植物这两个祖先拷贝经历了不同的进化过程，在多数物种内两个拷贝都被保留了下来，并各自经历新的重复和丢失过程，有的功能可能发生了改变；有的植物丢失了其中一个拷贝，另一个拷贝经历了至少一次重复，在不同植物中重复发生的时间和拷贝的分化也有一定差异，这种差异可能与不同植物的生活习性及其生存环境的多样性相关。

致谢 感谢本所龚洵副研究员提供部分材料。

【参 考 文 献】

- Doyle JJ, Doyle JL, 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue [J]. *Phytochem Bull*, 19: 11—15
- Durbin ML, Learn Jr GH, Huttley GA, et al, 1995. Evolution of the chalcone synthase gene family in the genus *Ipomoea* [J]. *Proc Natl Acad Sci*, 92: 3338—3342
- Feinbaum RL, Ausubel FM, 1998. Transcriptional regulation of the *Arabidopsis thaliana* chalcone synthase gene [J]. *Mol Cell Biol*, 8:

1985—1992

- Ferrer JL, Joseph M, Jez, Marianne E, Bowman, *et al*, 1999. Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis [J]. *Nature Structure Biology*, 6: 775—784
- Force A, Lynch M, Pickett FB, *et al*, 1999. Preservation of duplication genes by complementary, degenerative mutations [J]. *Genetics*, 151: 1531—1545
- Helasiutta Y, Kotilainen M, Elomaa P, *et al*, 1996. Duplication and functional divergence in the chalcone synthase gene family of Asteraceae: Evolution with substrate change and catalytic simplification [J]. *Proc Natl Acad Sci*, 93: 9033—9038
- Hermann A, Schulz W, Hahlbrock K, 1998. Two alleles of the single-copy chalcone synthase gene in parsley differ by a transposon-like element [J]. *Mol Gen Genet*, 12 (1): 93—98
- Huelsenbeck JP, Ronquist R, 2001. Bayesian inference of phylogeny and its impact on evolutionary biology [J]. *Science*, 294: 2310—2314
- Huang JX, Qu LJ, Yang J, *et al*, 2004. A preliminary study on the origin and evolution of Chalcone Synthase (CHS) gene in angiosperms [J]. *Acta Bot Sin*, 46 (1): 10—19
- Ohta T, 1993. Pattern of nucleotide substitution in growth hormone-prolactin gene family: a paradigm for evolution by gene duplication [J]. *Genetics*, 134: 1271—1276
- Qu LJ, Zhang Y, Xie M, *et al*, 1997. A chalcone synthase-like cDNA from rice anther [J]. *Sex Plant Reprod*, 10: 181—183
- Rannala B, Yang ZH, 1996. Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference [J]. *J Mol Evol*, 43: 304—311
- Saitou N, Nei M, 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees [J]. *Mol Biol Evol*, 4: 406—425
- Swofford DL, 2001. Paup: Phylogenetic Analysis Using Parsimony, ver. 4.0b8 [M]. Massachusetts: Sinauer, Associates, USA
- The Angiosperm Phylogeny Group, 1998. An ordinal classification for the families of flowering plants [J]. *Ann MO Bot Gard*, 85: 531—553
- The Angiosperm Phylogeny Group, 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II [J]. *Botanical Journal of the Linnean Society*, 141: 399—436
- Thompson JD, Higgins DG, Gibson TJ, 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice [J]. *Nucleic Acids Res*, 22: 4673—4680
- Tropf S, Lanz T, Rensing SA, *et al*, 1994. Evidence that stilbene synthase have developed from chalcone synthase several times in the course of evolution [J]. *J Mol Evol*, 38: 610—618
- Ursula NK, Barzen E, Bernhardt J, *et al*, 1987. Chalcone Synthase Genes in Plants: A Tool to Study Evolutionary Relationships [J]. *J Mol Evol*, 26: 213—225
- Wang JL, Qu LJ, Chen J, *et al*, 2000. Molecular evolution of the exon 2 of CHS gene and the possibility of its application to plant phylogenetic analysis [J]. *Chin Sci Bull*, 45 (19): 1735—1742
- Wienand U, Sommer H, Schwarz U, 1982. A general method to identify plant structural genes among genomic DNA clones using transposable element induced mutations [J]. *Mol Gen Genet*, 187: 195—201
- Yang J, Huang JX, Gu HY, *et al*, 2002. Duplication and adaptive evolution of the Chalcone Synthase genes of *Dendranthema* (Asteraceae) [J]. *Mol Biol Evol*, 19 (10): 1752—1759
- Yang J, Gu HY, Yang ZH, 2004. Likelihood analysis of the Chalcone Synthase genes suggests the role of positive selection in morning glories (*Ipomoea*) [J]. *Mol Biol Evol*, 19 (10): 1752—1759
- Yang JB (杨俊波), Tian X (田欣), Li DZ (李德铎), *et al*, 2002. An analysis of CHS gene family based on *Tutcheria spectabilis* (Theaceae) [J]. *Acta Bot Yunnan* (云南植物研究), 24: 209—214
- Yang JB, Tian X, Li DZ, *et al*, 2003. Molecular composition and evolution of the Chalcone Synthase (CHS) gene family in five species of *Camellia* (Theaceae) [J]. *Acta Bot Sin*, 45 (6): 659—666