

郭书巧, 束红梅, 巩元勇, 等. 菊科植物青蒿细胞色素 P450 基因家族分析[J]. 江苏农业学报, 2015, 31(6): 1232-1241.
doi: 10.3969/j.issn.1000-4440.2015.06.007

菊科植物青蒿细胞色素 P450 基因家族分析

郭书巧, 束红梅, 巩元勇, 蒋璐, 朱静雯, 倪万潮

(江苏省农业科学院经济作物研究所, 江苏 南京 210095)

摘要: 从 NCBI 中下载了青蒿 P450 全部全长序列, 经过一致性鉴定, 共获得 41 个 P450 单加氧酶基因。根据 P450 标准命名规则, 这 41 个基因分布在 3 个基因簇的 11 个家族中。与水稻、拟南芥 P450 氨基酸序列比对发现, 拟南芥中所特有的 CYP82、CYP83 和 CYP716 3 个家族在青蒿中出现, 而水稻特有的 CYP92 家族在双子叶植物青蒿中也存在。对青蒿 P450 蛋白质的理化性质和结构特性进行了分析, 二级结构分析结果表明, 青蒿 P450 蛋白质的二级结构以 α -螺旋为主。

关键词: 青蒿; 细胞色素 P450; 生物信息学

中图分类号: S184 **文献标识码:** A **文章编号:** 1000-4440(2015)06-1232-10

Bioinformatics analysis of cytochrome P450 monooxygenases gene family from *Artemisia annua*

GUO Shu-qiao, SHU Hong-mei, GONG Yuan-yong, JIANG Lu, ZHU Jing-wen, NI Wan-chao

(*Institute of Industrial Crops, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China*)

Abstract: In the study, a total of 41 genes encoding cytochrome P450 monooxygenases were identified based on the sequence downloaded from NCBI. These genes were classified into 3 clans and 11 families by sequence similarity. One of the first interesting features of the *Arabidopsis* species-specific families CYP82, CYP83 and CYP716, and family CYP92 specifically in rice are present in *Artemisia annua*. The secondary structure of putative protein P450 is characterized by α -helix.

Key words: *Artemisia annua*; cytochrome P450; bioinformatics

细胞色素 P450 (P450) 广泛存在于各种形式的生命体中, 它是自然界中最大且最古老的家族^[1-2]。在植物中, P450 主要参与植物体内初级和

次级代谢反应, 并发挥重要的催化作用。其催化反应的底物包括内源性物质, 如脂类、黄酮类、萜类、生物碱类等以及外源性物质, 如农药、除草剂等^[3]。P450 底物与功能的多样性是为了适应各种组织中不同的代谢途径和环境胁迫而进化的结果^[4]。

菊科中含有多种芳香族植物, 其中青蒿属 (*Artemisia annua*) 植物能产生一种倍半萜内酯过氧化物—青蒿素, 青蒿素是目前世界公认的抗疟疾首选药物。紫穗槐-4, 11-二烯 (Amorphadiene) 合酶是一种来源于青蒿的细胞色素 P450 单加氧 CYP71AV1, 它是一种多功能倍半萜内酯环化酶, 能将萜类前体

收稿日期: 2015-05-18

基金项目: 国家自然科学基金项目 (31101452); 江苏省农业三新工程项目 [SXGC (2015) 213]; 江苏省自主创新基金项目 [CX(13)5011]

作者简介: 郭书巧 (1972-), 山西洪洞人, 博士, 副研究员, 研究方向为功能植物次生代谢。(Tel) 025-84390290; (E-mail) gushuqiao@126.com

通讯作者: 倪万潮, (Tel) 025-84390618; (E-mail) nwchao2002@aliyun.com

法尼基焦磷酸 (FPP) 氧化成紫穗槐-4, 11-二烯 (Amorpha-4, 11-diene), 并进一步氧化成青蒿素合成的前体青蒿酸 (Artemisinic acid)^[5]。同时青蒿中还含有大量其他萜类化合物, 青蒿植物精油主要成分包括单萜和倍半萜^[6-7]。因此, 了解青蒿 P450, 对于进一步了解 P450 在植物初生代谢和次生代谢中的作用具有一定的指导意义。

1 材料与方法

1.1 材料

本研究所采用的青蒿 (*A. annua*) P450 氨基酸序列下载自 NCBI (<http://www.ncbi.nlm.nih.gov/protein/?term=Artemisia+P450>) 蛋白质数据库, 水稻 (*Oryza sativa*) 和拟南芥 (*Arabidopsis thaliana*) P450 氨基酸序列下载自 P450 蛋白质数据库 (<http://drnelson.uthsc.edu/CytochromeP450.html>)。

1.2 方法

1.2.1 菊科植物青蒿 P450 基因的鉴定 从 NCBI 蛋白质数据库中下载菊科植物青蒿中所有的 P450 全长序列, 对其逐一进行氨基酸的一致性鉴定。对还没有进行家族分类的 P450 进行 NCBI blast (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), 结合下载的拟南芥和水稻序列寻找与其一致性较高序列。根据国际命名标准对其进行分类: 氨基酸序列一致性 (Identity) 在 40% 至 55% 之间的序列为同一家族 (Family), 一致性为 55% ~ 97% 的序列归入同一亚家族 (Subfamily), 氨基酸序列的一致性大于 97% 的序列是由同一基因的不同位点即等位基因编码^[8]。

1.2.2 菊科植物青蒿 P450 系统进化树的构建 根据青蒿 P450 的氨基酸序列, 通过 MUSCLE^[9] 对所鉴定的 P450 氨基酸序列进行比对, 采用 MEGA 6.0 软件 Neighbor joining 法构建进化树, 重复次数为 1 000 次^[10]。

1.2.3 青蒿和拟南芥、水稻 P450 系统进化树对比 在水稻和拟南芥 P450 中选择在青蒿中存在的 11 个家族所有 P450 全长氨基酸序列, 采用 MEGA 6.0 软件中 Neighbor joining 法对已鉴定的青蒿中氨基酸序列一起构建系统发育树, 重复次数为 1 000 次^[10]。

1.2.4 青蒿 P450 蛋白质结构分析

1.2.4.1 青蒿 P450 蛋白质一级结构理化特性分析 利用在线软件 Protparam (<http://web.expasy.org/protparam>) 对氨基酸数目、相对分子质量、理论等电

点、分子式、脂肪族氨基酸、蛋白质疏水性、不稳定系数等蛋白质一级结构理化特性进行分析。

1.2.4.2 青蒿 P450 蛋白质亚细胞定位和信号肽预测 使用 TargetP 在线程序 (<http://www.cbs.dtu.dk/services/TargetP/>) 对青蒿 P450 蛋白质进行亚细胞定位^[11]。信号肽预测使用在线软件 SignalP4.1 (<http://www.cbs.dtu.dk/services/SignalP/>) 进行^[12]。

1.2.4.3 青蒿 P450 蛋白质的二级结构分析 利用 ExPaSy 在线程序 SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html) 预测分析 α -螺旋、 β -折叠、转角等; 蛋白质晶体结构信息来源于 RCSB PDB (<http://www.rcsb.org>), 共 4 个细菌 P450 [P450cam (CYP101)、P450BM3 (CYP102)、P450eryF (CYP107A1) 和 P450terP (CYP108)], 1 个哺乳动物 P450 (CYP2C5), 1 个拟南芥丙二烯氧化合酶 (CYP74A), 1 个灰白银胶菊 (*Parthenium argentatum*) 丙二烯氧化合酶 AOS (CYP74A2)^[13]。将 6 个氨基酸序列与青蒿的 P450 用 ClustalW 软件进行同源比对, 选择一致性较高的 P450 (CYP2C5) 为模板^[14], 应用在线 ESPript 程序 (<http://esprict.ibcp.fr/ESPrict/ESPrict/>) 进行二级结构及基质结合位点的预测^[15]。

2 结果与分析

2.1 菊科植物青蒿 P450 基因的鉴定结果

从 NCBI 蛋白质数据库中共下载 76 条青蒿 P450 蛋白质氨基酸序列。其中具有全长的氨基酸序列 57 条, 包含 8 条 P450 还原酶蛋白质氨基酸序列。将剩余的 49 条序列进行氨基酸的一致性鉴定, 最后获得 41 个编码细胞色素 P450 单加氧酶的基因 (表 1), 其中有 14 个基因目前没有经过正式的家庭分类。对没有分类的 P450 进行 NCBI blast (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), 并结合下载的拟南芥和水稻 P450 氨基酸序列找到与其一致性较高的序列, 对其进行进一步确认, 鉴定结果见表 2。

2.2 青蒿 P450 基因系统进化

从青蒿 P450 系统进化树 (图 1) 可以看出, 青蒿的 41 个基因分布相对集中, 共分布在 3 个基因簇的 11 个家族中。其中 Clan72 中的 CYP72 有 2 个, Clan85 中的 CYP90 有 1 个, CYP716 有 4 个。其余

表 1 预测的青蒿细胞色素 P450 基因分布

Table 1 Distribution of the P450 genes from *Artemisia annua*

GenBank 登录号	基因簇	家族	亚家族	蛋白质编号	编码产物	参考文献
DQ268763.1	71	71	71A	ABB82944.1 *	ADOS	[16]
DQ315671.1	71	71	71A	ABC41927.1 *	ADOS	[17]
DQ363131	71	706	---	ABC94480.1	F3H	[6]
DQ363132	85	90	90A	ABC94481.1	S23H	[6]
DQ363133	85	716	716D	ABC94482.1	T13H	[6]
DQ363134	85	716	716A	ABC94483.1	T5H	[6]
DQ370065	71	706	---	ABD15097.1	P450	[6]
DQ453967	71	71	71A	ABE57266.1 *	ADOS	[18]
DQ667171	71	71	71A	ABG49366.1 *	ADOS	[6]
ABC91755.1	71	71	71A	ABC91755.1 *	ADOS	[6]
EF197889	71	71	71A	ABM88788.1 *	ADOS	[6]
EU684540	71	71	71A	ABI31728.1 *	ADOS	[6]
DQ872632	71	71	71A	ACF74516.1 *	ADOS	[6]
GU318226	71	73	73A	ADO16181.1	CaOS	[6]
GU318227	71	83	---	ADO16182.1	P450	[6]
GU318228.1	71	83	---	ADO16183.1	P450	[6]
GU318229	72	72	CYP72A	ADO16184.1	P450	[6]
GU318230	71	92	92A	ADO16185.1	P450	[6]
HQ315834.1	71	71	71A	ADU25498.1 *	ADOS	[6]
JN594504	71	71	71A	AFO64615.1 *	ADOS	[6]
JN594505	72	72	72A	AFO64616.1	P450	[6]
JN594506	71	81	81B	AFO64617.1	Fa10, 12, 14H;	[6]
JQ254992.1	71	71	71A	AFP19100.1 *	ADOS	[19]
KC594703.1	71	98	98A	AGN54071.1	P450	[20]
KC594704.1	71	82	82A	AGN54072.1	P450	[20]
KF309251.1	85	716	716A	AHF22083.1 *	P450	-
KF309250.1	85	716	716D	AHF22082.1 *	P450	-
AB706288	71	71	71A	BAM68808.1 *	ADOS	[6-7]
AB706289	71	71	71A	BAM68809.1 *	ADOS	[7]
AB706290	71	71	71A	BAM68810.1 *	ADOS	[7]
AB706291.1	71	71	71A	BAM68811.1 *	ADOS	[7]
AB706292.1	71	71	71A	BAM68812.1 *	ADOS	[7]
AB706293.1	71	71	71A	BAM68813.1 *	ADOS	[7]
AB750729.1	71	71	71A	BAM68814.1 *	ADOS	[7]
AB750730.1	71	71	71A	BAM68815.1 *	ADOS	[7]
AB750731.1	71	71	71A	BAM68816.1 *	ADOS	[7]
AB750732.1	71	71	71A	BAM68817.1 *	ADOS	[7]
AB750733.1	71	71	71A	BAM68818.1 *	ADOS	[7]
AB750734.1	71	71	71A	BAM68819.1 *	ADOS	[7]
AB750735.1	71	71	71A	BAM68820.1 *	ADOS	[7]
AB750736.1	71	71	71A	BAM68821.1 *	ADOS	[7]

* 表示蛋白质的分布由前人报道;---表示不确定;ADOS:紫槐二烯氧化酶;F3H:类黄酮 3 羟化酶;S23H:类固醇 23 羟化酶;T13H:紫杉醇 13 羟化酶;Fa10, 12, 14H:脂肪酸氧化酶;CaOS:肉桂酸氧化酶。

表 2 预测的青蒿细胞色素 P450 与其同源序列的一致性

Table 2 Similarity of P450 amino acid sequence from *A. annua* and other species

蛋白质编号	参考序列			一致性 (%)	结果
	参比序列 ID	分类	物种		
ABC94480.1	AJD25237.1	CYP706C35	<i>Salvia miltiorrhiza</i>	48	CYP706
	AJD25238.1	CYP706G11	<i>S.miltiorrhiza</i>	48	CYP706
ABC94481.1	XP_004240946.1	CYP90A1	<i>S.lycopersicum</i>	71	CYP90A
ABC94482.1	AHF22082.1	CYP716D22	<i>Artemisia annua</i>	66	CYP716D
	AJD25250.1	CYP716D25	<i>S.miltiorrhiza</i>	56	CYP716D
ABC94483.1	AHF22083.1	CYP716A14v2	<i>A.annua</i>	99	CYP716A
	AJD25248.1	CYP716A89	<i>S.miltiorrhiza</i>	66	CYP716A
ABD15097.1	AJD25237.1	CYP706C35	<i>S.miltiorrhiza</i>	45	CYP706
	AJD25238.1	CYP706G11	<i>S.miltiorrhiza</i>	43	CYP706
ADO16181.1	ABC69413.1	CYP73A47v2	<i>Nicotianatabacum</i>	89	CYP73A
ADO16182.1	XP_002301993.2	CYP83A1	<i>Populstrichocarpa</i>	53	CYP83
	KHN02071.1	CYP83B1	<i>P.trichocarpa</i>	50	CYP83
ADO16183.1	XP_002301993.2	CYP83A1	<i>P.trichocarpa</i>	52	CYP83
	KHN02071.1	CYP83B1	<i>P.trichocarpa</i>	50	CYP83
ADO16184.1	XP_002263609.	CYP72A219	<i>Vitisvinifera</i>	55	CYP72A
ADO16185.1	ABC69381.1	CYP92A2v2	<i>N.tabacum</i>	75	CYP92A
AFO64616.1	H2DH21.1	CYP72A219	<i>Panax ginseng</i>	64	CYP72A
AFO64617.1	ABC69377.1	CYP81B2v2	<i>N.tabacum</i>	61	CYP81B
AGN54071.1	XP_002308860.1	CYP98A3	<i>P.trichocarp</i>	74	CYP98A
AGN54072.1	XP_007017416.1	CYP 82A3	<i>Theobroma cacao</i>	56	CYP 82A

全部分布于 Clan 71 中,且大多集中在 CYP71 家族中。来源于青蒿中的 16 个基因,除 *AaABG49366.1* 外,都在一个分支中,一致性在 97%~99%,在分类上被视为同一基因的等位基因^[8]。来源于青蒿属的其他物种的 P450 一致性为 82%~86%。

2.3 青蒿、拟南芥和水稻 P450 系统进化树的对比

从青蒿和拟南芥和水稻的 P450 系统进化树(图 2)中可以看出,双子叶植物拟南芥中所特有的^[21],水稻中不存在的 CYP82、CYP83 和 CYP716 3 个家族在青蒿中都出现了。有趣的是在拟南芥中不存在,而在单子叶植物中特有的 CYP92 家族居然在

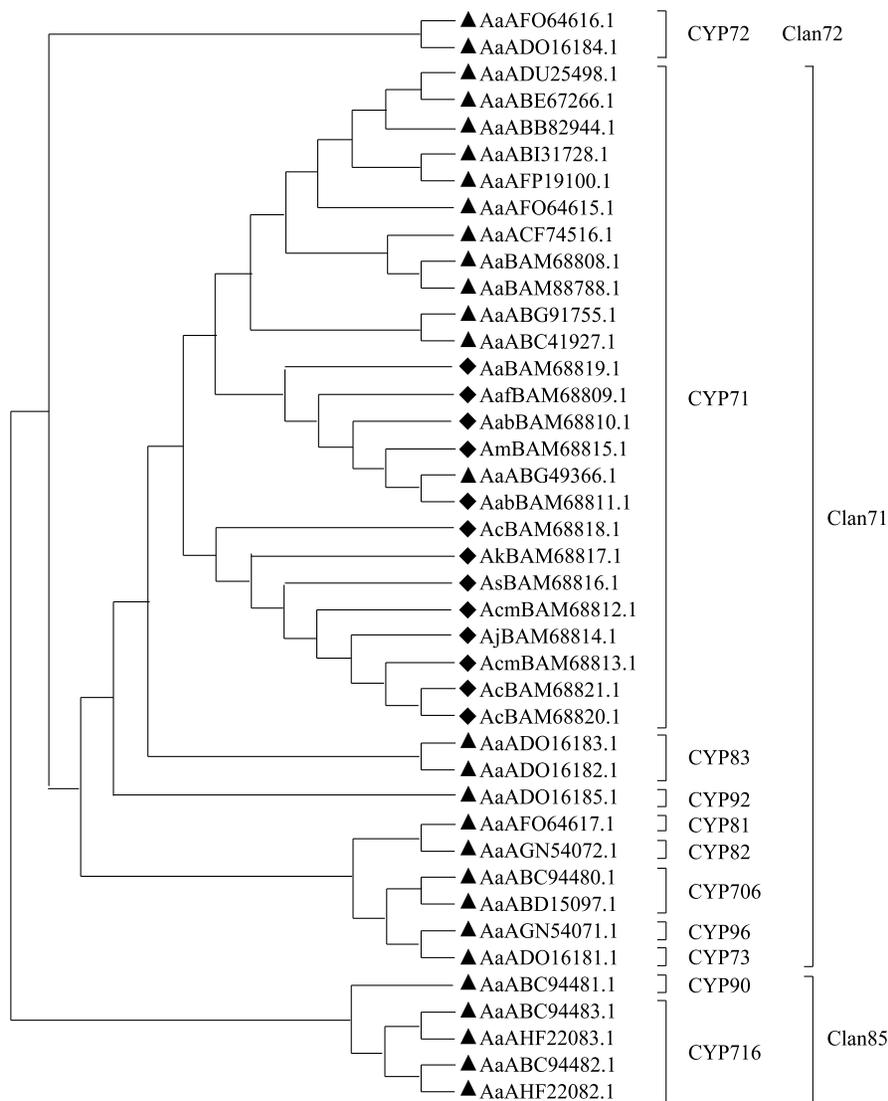
青蒿中出现。而 Nelson 等^[21]提到的菊科中出现的 8 个家族 CYP73、CYP75、CYP76、CYP81、CYP93、CYP701、CYP74、CYP87,在青蒿中只有 CYP73 和 CYP81 2 个家族出现。

2.4 青蒿 P450 蛋白质的结构特性

选择代表不同家族和亚家族的 19 个青蒿 P450 分析其特征特性。代表性 P450 氨基酸数为 473~528,平均为 500;相对分子量在 53 977.9 至 61 127.3 之间,平均为 56 882.1(表 3)。根据理论等电点(pI)将青蒿 P450 蛋白质分为 2 类:一类呈酸性,pI 值介于 6.00 和 6.69 之间,包括 CYP706、CYP83 和

CYP82 共 3 个家族中的 5 个蛋白质;另一类呈碱性, pI 值介于 8.03 和 9.25 之间,包括 CYP90、CYP716、CYP73、CYP72、CYP92、CYP71、CYP81、CYP98 共 8 个家族中的 14 个蛋白质。不稳定指数介于 32.19 和 42.58 之间,平均值为 39.64,总体上青蒿 P450 稳定性不太好。脂溶性指数在 89.75 和 102.11 之间,平均值也高达 96.37,而其亲水性指数皆为负

值,故青蒿 P450 都是亲脂性蛋白质。从表 3 可以看出,青蒿 P450 中除 ADO16185.1 (CYP92A)、ABC94480.1 (CYP706)、ADO16181.1 (CYP73A) 外,其余都是分泌途径中的蛋白质。信号肽分析结果表明,除 ADO16181.1 (CYP73A) 外,其余都含有跨膜区,为跨膜蛋白质。



▲代表序列来源于青蒿属中的青蒿,◆代表序列来源于青蒿属中的其他物种。

图 1 青蒿 P450 基因系统进化树

Fig.1 The phylogenetic tree of P450 from *A. annua*

利用在线程序 SOPMA 进行二级结构预测。结果(表 3)表明,本研究所涉及的 P450 家族蛋白质二级结构以 α -螺旋为主要构成元件,平均所占比例达 39.64%,其次是无规则卷曲、 β -折叠和转角,平均所

占比例分别为 28.48%、16.57%和 8.28%。以编号为 BAM68808.1 的蛋白质为例,绘制蛋白质二级结构(图 3)。从图 3 中可见, α -螺旋为主要构成元件,无规则卷曲、 β -折叠和转角分散在整个蛋白质中。

表 3 青蒿 P450 家族蛋白质的生物信息

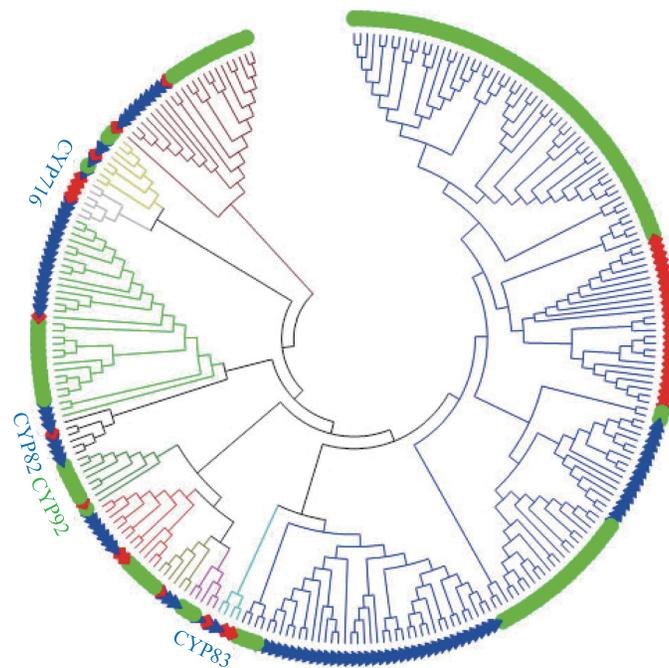
Table 3 The biological information of P450 protein from *A. annua*

蛋白质编号	家族	氨基酸数	理化性质						次级结构占比(%)				亚细胞定位	跨膜区预测
			相对分子量	理论等电点(pI)	分子式	不稳定指数	脂溶指数	亲水性指数	α -螺旋	β -折叠	转角	无规则卷曲		
ABC94480.1	CYP706	528	60 906.5	6.22	C ₂₇₇₅ H ₄₃₄₀ N ₇₂₂₀ 778S ₂₀	33.83	100.91	-0.159	47.16	18.37	8.90	25.57	-	TM
ABC94481.1	CYP90A	473	54 170.6	8.49	----	39.36	92.54	-0.129	42.07	21.14	9.30	27.48	S	TM
ABC94482.1	CYP716D	480	54 813.9	9.14	C ₂₅₁₈ H ₃₉₁₄ N ₆₅₈₀ 673S ₂₀	34.59	92.81	-0.070	43.13	17.92	10.00	28.96	S	TM
ABC94483.1	CYP716A	474	53 977.9	8.06	C ₂₄₉₁ H ₃₈₄₁ N ₆₂₉₀ 671S ₂₀	41.53	90.72	-0.095	45.15	14.56	10.34	29.96	S	TM
ABD15097.1	CYP706	534	61 127.3	6.00	C ₂₇₅₅ H ₄₃₁₁ N ₇₂₅₀ 798S ₂₄	32.19	93.78	-0.216	41.81	17.03	10.56	30.60	S	TM
ADO16181.1	CYP73A	505	57 799.2	9.10	C ₂₆₃₄ H ₄₁₇₂ N ₇₁₂₀ 721S ₁₄	42.58	100.36	-0.202	42.07	21.14	9.30	27.48	S	TM
ADO16182.1	CYP83	515	58 165.2	6.69	C ₂₆₃₆ H ₄₁₅₈ N ₆₈₆₀ 756S ₁₉	41.29	99.38	-0.136	39.61	17.28	12.04	31.07	-	No-TM
ADO16183.1	CYP83	491	55 390.0	6.65	C ₂₅₀₄ H ₃₉₄₆ N ₆₅₄₀ 721S ₂₀	42.09	97.29	-0.151	45.15	14.56	10.34	29.96	S	TM
ADO16184.1	CYP72A	516	58 703.6	9.25	C ₂₆₅₇ H ₄₂₂₉ N ₇₁₅₀ 732S ₂₅	40.23	95.79	-0.166	41.81	17.03	10.56	30.60	S	TM
ADO16185.1	CYP92A	513	58 762.2	8.06	C ₂₆₆₄ H ₄₂₀₆ N ₇₀₀₀ 753S ₂₁	40.73	93.47	-0.258	40.35	18.32	10.14	31.19	M	TM
ADU25498.1	CYP71A	488	54 892.7	9.00	C ₂₄₇₄ H ₃₉₇₅ N ₆₆₅ O ₇₁₁ S ₁₅	43.74	100.33	-0.170	45.49	16.80	8.40	29.30	S	TM
AF064616.1	CYP72A	515	58 836.1	9.06	C ₂₆₇₂ H ₄₁₈₀ N ₇₁₄ O ₇₄₃ S ₂₀	39.39	89.75	-0.237	45.24	18.06	10.68	26.02	S	TM
AF064617.1	CYP81B	502	56 923.4	8.64	C ₂₅₇₉ H ₄₁₀₇ N ₆₈₃ O ₇₂₁ S ₂₂	42.28	102.11	-0.062	41.63	19.12	9.96	29.28	S	TM
AFP19100.1	CYP71A	488	54 922.7	9.00	C ₂₄₇₉ H ₃₉₇₇ N ₆₆₅ O ₇₀₉ S ₁₅	43.6	100.53	-0.158	46.11	16.80	8.40	28.69	S	TM
AGN54071.1	CYP98A	510	57 889.1	8.99	C ₂₆₁₇ H ₄₁₃₄ N ₇₁₄ O ₇₂₈ S ₂₀	33.76	92.35	-0.254	38.24	20.39	11.96	29.41	S	TM
AGN54072.1	CYP82A	523	59 215.6	6.30	C ₂₆₈₀ H ₄₂₂₉ N ₇₀₉ O ₇₅₈ S ₂₂	36.02	101.01	-0.019	44.36	22.56	7.07	26.00	S	TM
AHF22082.1	CYP716D	479	54 547.4	8.03	C ₂₅₀₈ H ₃₈₈₄ N ₆₄₄ O ₆₈₀ S ₁₉	40.31	95.70	-0.055	49.27	15.87	9.81	25.05	S	TM
AHF22083.1	CYP716A	474	54 005.9	8.05	C ₂₄₉₃ H ₃₈₄₅ N ₆₂₉ O ₆₇₁ S ₂₀	42.17	91.54	-0.093	45.78	14.56	10.13	29.54	S	TM
BAM68808.1	CYP71A	495	55 710.7	9.13	C ₂₅₁₄ H ₄₀₄₂ N ₆₇₄ O ₇₁₈ S ₁₆	43.48	100.69	-0.154	46.67	16.57	8.28	28.48	S	TM
平均值	---	500	56 882.1	8.10	---	39.64	96.37	-0.147	43.74	17.79	9.80	28.67		

亚细胞定位栏中,-表示没有定位,S表示分泌路径,M表示线粒体;跨膜区预测栏中,TM表示转膜区,No-TM表示没有转膜区。

利用 ESPript 在线程序对青蒿 P450 进行同源比对和二级结构预测,发现青蒿 P450 一级结构同源性不高,但它们却有着非常类似的二级结构(图 3)。将这些序列与 CYP2C5(1DT6)的 X-射线晶体结构图比对,发现有 6 个基质识别位点(SRSs),且所有基质结合位点都分布在不同结构元件中,识别特异的底物。所有青蒿 P450 都含有典型的保守结

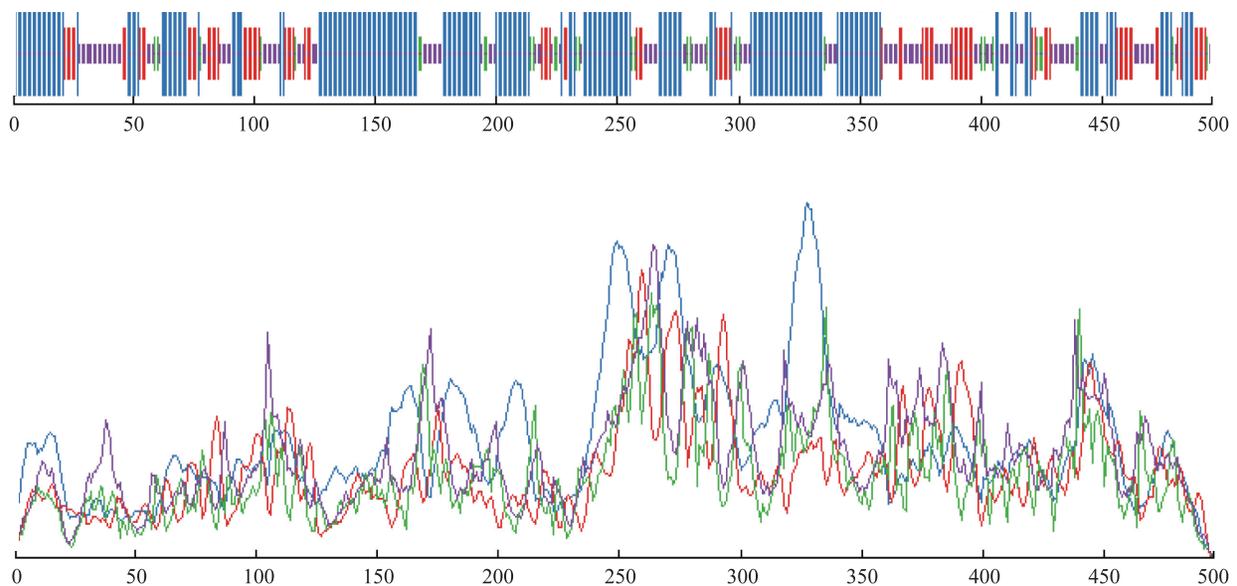
构域:PERF 结构域(PxRx)、K 螺旋(ExxR)、C 末端的血红素结合域(FxxGxxxCxG)以及 N 端的跨膜区(图 4)。血红素结合域是 P450 蛋白的特征结构,该结构域中存在与铁元素形成硫醇盐离子键的绝对保守的半胱氨酸残基。二级结构预测结果表明,青蒿中不同 P450 氨基酸序列间存在严格的保守基序,这些保守的基序通常分散于 P450 氨基酸序列



11个基因家族以不同的颜色表示,顺时针方向依次为CYP71(蓝色)、CYP83(水绿色)、CYP98(橄榄色)、CYP706(红色)、CYP92(绿色)、CYP73(紫红色)、CYP82(黑色)、CYP81(酸橙色)、CYP716(银白色)、CYP90(黄色)、CYP72(褐紫红色)。外圈的▲蓝色三角表示氨基酸序列来源于拟南芥,●酸橙色圆代表氨基酸序列来源于水稻,◆红色菱形表示氨基酸序列来源于青蒿。

图2 青蒿、拟南芥和水稻的P450系统进化树

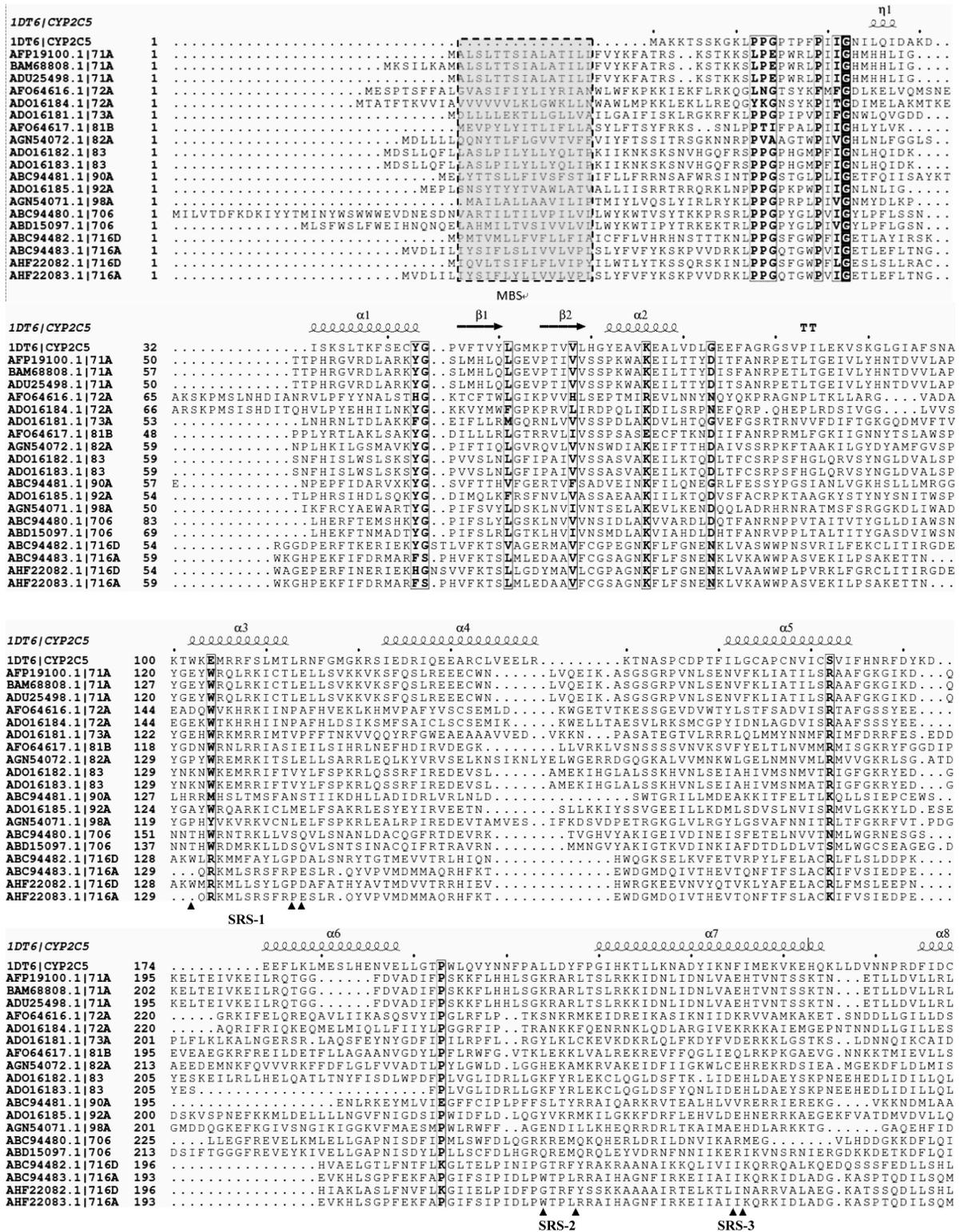
Fig.2 The phylogenetic tree of P450 from *A. annua*, *Arabidopsis* and rice



蓝色表示 α -螺旋,红色表示 β -折叠,绿色表示转角,紫色表示无规则卷曲。

图3 利用程序SOPMA进行青蒿P450蛋白质(BAM68808.1)二级结构预测

Fig.3 Prediction of secondary structure of P450(BAM68808.1) from *A. annua* by SOPMA



黑色表示序列一致, 灰色表示序列保守; 实线方框表示蛋白质保守域; 虚线半透明方框表示 P450 中典型的保守结构域; 序列下方的小三角表示与 CYP2C5 结合的基质位点(SRS)的氨基酸残基; 序列上方标有 α 和 η 的波形曲线分别表示 α-螺旋和 3₁₀-螺旋; 标有 β 的箭头表示 β-单链; 标有 TT 字母的表示 α-和 β-转角; 联配的 3DAM 序列每 10 个氨基酸残基上方用一个黑点标示。MBS: 可能的膜结合基序。

图 4 利用程序 ESPrnt 进行青蒿 P450 蛋白质二级结构预测
Fig.4 Prediction of secondary structure of P450 from *A. annua* using ESPrnt

中的5个区域,它们与血红素结合区($F_{XX}G_{XXX}C_X$ G),共同组成保守的P450结构中心^[14]。

3 讨论

P450是生物界无处不在的超级大家族,植物P450超家族被归为10个基因簇,127个家族,其中水稻中有48个家族,拟南芥中有45个家族^[21-22],有9个家族是双子叶植物拟南芥所特有的,10个家族是单子叶植物水稻所特有。本研究共鉴定了41个青蒿P450,分布在3个基因簇的11个家族中。有趣的是,作为双子叶植物的青蒿,目前鉴定的仅有的11个家族中,既包括了拟南芥所特有的3个家族CYP82、CYP83和CYP716,又包括了水稻特有的1个家族CYP92。而Nelson等提到的菊科中出现的8个家族中^[21],只有CYP73和CYP81 2个家族出现。造成这种现象的原因,主要是由于拟南芥和水稻的基因组测序已经完成,其P450的鉴定已趋于饱和,而菊科基因组的研究还远远落后于拟南芥和水稻,故还会有新的P450不断补充。另一方面,也提示我们菊科植物作为一个古老的物种,有着更为丰富的P450家族蛋白质有待开发、鉴定。

所有的P450蛋白质都呈现相似的二级结构,尽管它们的序列一致性不高(低于20%)^[23]。为了更好地理解P450结构和功能之间的关系,我们以CYP2C5(1DT6)的X-射线晶体结构图为模板^[14],对青蒿P450进行了氨基酸序列联配和次级结构元件的预测。发现所有的青蒿P450都含有典型的保守结构域,包括PERF结构域($PxRx$)、K螺旋($E_{xx}R$)和C末端的血红素结合域($F_{xx}G_{xxx}C_xG$),另外某些氨基酸残基也具有高度的保守性,这些保守结构域在蛋白质的折叠、组装中起着关键性作用^[14]。而可变区主要出现在基质结合区。P450是最大的酶家族之一,能催化多种底物,这就需要该蛋白质能结合多种底物,行使各自的功能。

参考文献:

- [1] OMURA T, SATO R. A new cytochrome in liver microsomes[J]. J Biol Chem, 1962, 237:1375-1376.
- [2] ESTABROOK R W, COOPER D Y, ROSENTHAL O. The light reversible carbon monoxide inhibition of the steroid C21-hydroxylase system of the adrenal cortex [J]. Biochem Z, 1963, 338:741-755.
- [3] WERCK-REICHHART D, FEYEREISEN R. Cytochromes P450: a success story[J]. Genome Biol, 2000, 1(6):30031-30039.
- [4] FU C, XIONG J, MIAO W. Genome-wide identification and characterization of cytochrome P450 monooxygenase genes in the ciliate *Tetrahymena thermophila* [J]. BMC Genomics, 2009, 10(208): 1-21.
- [5] BOSMAN A, MENDIS K N. A major transition in malaria treatment: the adoption and deployment of artemisinin-based combination therapies [J]. Am J Trop Med Hyg, 2007, 77: 193-197.
- [6] MISRA A, CHANOTIYA C S, GUPTA M M, et al. Characterization of cytochrome P450 monooxygenases isolated from trichome enriched fraction of *Artemisia annua* L. leaf [J]. Gene, 2012, 510(2):193-201.
- [7] KOMORI A, SUZUKI M, SEKI H, et al. Comparative functional analysis of CYP71AV1 natural variants reveals an important residue for the successive oxidation of amorpha-4,11-diene [J]. FEBS Lett, 2013, 587(3):278-284.
- [8] NELSON D R. Cytochrome P450 nomenclature, 2004 [J]. Methods Mol Biol, 2006, 320: 1-10.
- [9] EDGAR R C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity [J]. BMC Bioinf, 2004, 5: 113-131.
- [10] HALL B G. Building phylogenetic trees from molecular data with MEGA [J]. Mol Biol Evol, 2013, 30(5):1229-1235.
- [11] EMANUELSSON O, NIELSEN H, BRUNAK S, et al. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence [J]. J Mol Biol, 2000, 300: 1005-1016.
- [12] PETERSEN T N, BRUNAK S, VON HEIJNE G, et al. SignalP 4.0: discriminating signal peptides from transmembrane regions [J]. Nature Methods, 2011, 8:785-786.
- [13] LI L, CHANG Z, PAN Z, et al. Modes of heme binding and substrate access for cytochrome P450CYP74A revealed by crystal structures of allene oxide synthase [J]. PNAS, 2008, 105(37): 13883-13888.
- [14] WILLIAMS P A, COSME J, SRIDHAR V, et al. Mammalian microsomal cytochrome P450 monooxygenase: structural adaptations for membrane binding and functional diversity [J]. Mol Cell, 2000, 5(1):121-131.
- [15] GOUET P, ROBERT X, COURCELLE E. ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins [J]. Nucleic Acids Res, 2003, 31(13): 3320-3323.
- [16] RO D K, PARADISE E M, OUELLET M, et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast [J]. Nature, 2006, 440(7086):940-943.
- [17] TEOH K H, POLICHUK D R, REED D W, et al. *Artemisia annua* L. (Asteraceae) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin [J]. FEBS

- Lett, 2006, 580 (5):1411-1416.
- [18] OLSSON M E, OLOFSSON L M, LINDAHL A L, et al. Localization of enzymes of artemisinin biosynthesis to the apical cells of glandular secretory trichomes of *Artemisia annua* L. [J]. Phytochemistry, 2009, 70 (9):1123-1128.
- [19] TING H M, WANG B, RYDEN A M, et al. The metabolite chemotype of *Nicotianabes glauca* transiently expressing artemisinin biosynthetic pathway genes is a function of CYP71AV1 type and relative gene dosage[J]. New Phytol, 2013, 199 (2):352-366.
- [20] NAIR P, MISRA A, SINGH A, et al. Differentially expressed genes during contrasting growth stages of *Artemisia annua* for artemisinin content[J]. PLoS ONE, 2013, 8 (4): 6037510-6037512.
- [21] NELSON D R, SCHULER M A, PAQUETTE S M, et al. Comparative genomics of rice and *Arabidopsis*. Analysis of 727 cytochrome P450 genes and pseudogenes from a monocot and a dicot[J]. Plant Physiol, 2004, 135 (2): 756-772.
- [22] NELSON D R, WERCK-REICHHART D. A P450-centric view of plant evolution[J]. Plant J, 2011, 66 (1): 194-211.
- [23] GRAHAM S E, PETERSON J A. How similar are P450s and what can their differences teach us? [J]. Arch Biochem Biophys, 1999, 369(1):24-29.

(责任编辑:张震林)