

18种观赏植物 *F3'5'H* 基因生物信息学分析曾慧兰<sup>1,2</sup>, 高阳<sup>1</sup>, 卢毅<sup>1</sup>

(1. 宜春学院生命科学与资源环境学院, 江西宜春 336000; 2. 江西省作物生长发育与调控重点实验室, 江西宜春 336000)

**摘要** 采用生物信息学方法对 18 种观赏植物类黄酮-3'-5'-羟化酶基因 (flavonoid-3'-5'-hydroxylase, *F3'5'H*) 的 mRNA 和氨基酸序列的理化性质、跨膜结构域、保守结构域、亚细胞定位、二级结构、三级结构和同源性进行预测与分析。结果表明, 绝大多数观赏植物的 *F3'5'H* 为亲水性稳定蛋白质, 以  $\alpha$  螺旋为主、无信号肽的跨膜蛋白质; 大多数定位于内质网膜上; 其三级结构模型为 5ylw. 1. A 铁锈醇合成酶, 为单链蛋白, 属于细胞色素 P450 基因家族; 同源保守氨基酸序列为“LPPGP”“AGTDTs”和“PFGAGRRICAG”。

**关键词** 生物信息学; 观赏植物; 氨基酸序列; 类黄酮-3'-5'-羟化酶

中图分类号 S68 文献标识码 A

文章编号 0517-6611(2023)07-0103-07

doi: 10.3969/j.issn.0517-6611.2023.07.026

开放科学(资源服务)标识码(OSID):



### Bioinformatical Analysis of Flavonoid-3'5'-Hydroxylase Genes Originated from 18 Ornamental Plants

ZENG Hui-lan<sup>1,2</sup>, GAO Yang<sup>1</sup>, LU Yi<sup>1</sup> (1. College of Life Science and Resources and Environment, Yichun University, Yichun, Jiangxi 336000; 2. Key Laboratory of Crop Growth, Development and Regulation, Yichun University, Yichun, Jiangxi 336000)

**Abstract** Predictive analyses of the mRNA sequence or amino acid sequences of the flavonoid 3'5'-hydroxylase gene (*F3'5'H*) originated from 18 ornamental plants and their physical and chemical parameters, transmembrane domains, conserved domains, subcellular localizations, secondary structure, tertiary structure and evolutionary relationships were explored using bioinformatics method. The results suggested that most *F3'5'H* of the 18 ornamental plants are hydrophilic stable proteins and transmembrane proteins, which are mainly made of alpha-helix and no signal peptides; they are mostly located on endoplasmic reticulum membrane and assigned to 5ylw. 1. A, which belongs to cytochrome P450 superfamily and is a monomer named ferruginol synthase in tertiary structure analysis; the Homologous conserved amino acid sequences are “LPPGP” “AGTDTs” and “PFGAGRRICAG”.

**Key words** Bioinformatical; Ornamental plants; Amino acid sequence; Flavonoid-3'5'-hydroxylase

花色是观赏植物的重要性状, 植物的花色主要是由黄酮类化合物、类胡萝卜素和生物碱 3 类物质含量决定的<sup>[1]</sup>, 而花色苷是黄酮类化合物的主要组成之一, 能控制花的橙红到蓝紫等颜色<sup>[2]</sup>。类黄酮-3'-5'-羟化酶 (flavonoid-3'-5'-hydroxylase, *F3'5'H*) 是花色苷代谢途径中的关键酶之一<sup>[3]</sup>。*F3'5'H* 催化花色苷母体 B 环 3', 5' 位形成羟基基团, 最终形成使花变蓝的飞燕草色素, 被称为“蓝色基因”<sup>[4]</sup>, 所以它对蓝色花的育种具有非常重要的作用<sup>[5]</sup>。

目前, 已从鹤望兰 (*Strelitzia reginae*)、矮牵牛 (*Petunia hybrida*)、非洲堇 (*Saintpaulia* spp.)、细茎石斛 (*Dendrobium moniliforme*) 等多种植物中分离出 *F3'5'H* 基因并对其功能进行了深入研究<sup>[6-9]</sup>。黄敏玲等<sup>[8]</sup>采用 RT-PCR 和 RACE 方法从鹤望兰 (*Strelitzia reginae*) 黄色花萼中克隆到类黄酮生物合成途径关键基因 *SrF3'5'H*, 并分析了其在鹤望兰不同花期和花瓣中的表达含量。Qi 等<sup>[10]</sup>利用转基因技术和 qRT-PCR 验证了蝴蝶兰属 (*Phalaenopsis*) *PhF3'5'H* 在花色中的功能。严黎等<sup>[11]</sup>通过利用 PCR 方法扩增了丹参 (*Salvia miltiorrhiza* Bunge) 中的 *F3'5'H* 基因全长, 并对该基因编码蛋白质的理化定性、亚细胞定位、结构域等进行了分析。方颖等<sup>[12]</sup>以华丽龙胆 (*Gentiana sino-ornata*) 为材料利用 RT-PCR 技术克隆了 *GsF3'5'H* 全长, 并进行了生物信息学分析。也有不少研究利用转录组测序技术关注不同花色植物中

*F3'5'H* 的表达与功能, 并进行生物信息学分析<sup>[2,13-14]</sup>。虽然当前植物中 *F3'5'H* 基因功能和表达情况的研究不少, 但目前对各植物中 *F3'5'H* 进行系统比较生物信息学相关的详细分析鲜有报道。

该研究利用生物信息学原理和软件, 对 NCBI GenBank 数据库中已正式发表的 18 种观赏植物的 *F3'5'H* 基因全长 mRNA 和氨基酸序列进行理化性质、亚细胞定位、跨膜结构、保守区域、二级结构、三级结构和同源性进行分析, 旨在为蓝色花卉的分子育种提供理论参考依据。

### 1 材料与方法

**1.1 材料** 从美国国家生物技术信息中心 (National Center for Biotechnology Information, NCBI) GenBank 数据库中检索已正式注册、物种来源明确的 18 种观赏植物的 *F3'5'H* 基因全长 mRNA 核酸序列及其编码的氨基酸序列 (表 1)。在这 18 种观赏植物中, 蓝紫色系花卉有 10 种, 红色系有 6 种, 其他色系 2 种。

**1.2 分析方法** 运用 ProtParam<sup>[15]</sup> (<https://web.expasy.org/protparam/>) 对氨基酸的基本理化性质即氨基酸数、分子式、相对分子量、理论等电点、稳定性系数、带负电残基总数 (Asp+Glu)、带正电残基总数 (Arg+Lys)、亲水性和脂肪族氨基酸指数进行在线分析; 通过 SignalP 4.1<sup>[16]</sup> (<https://services.healthtech.dtu.dk/service.php?SignalP-4.1>) 和 SignalP 5.0<sup>[17]</sup> (<https://services.healthtech.dtu.dk/service.php?SignalP-5.0>) 对各序列的信号肽进行预测; 利用 TMHMM 2.0<sup>[18]</sup> (<https://services.healthtech.dtu.dk/service.php?TMHMM-2.0>) 分析氨基酸序列的跨膜结构域, 并通过

**基金项目** 江西省作物生长发育调控实验室开放课题项目 (KFJJ-201704); 宜春学院博士科研启动项目 (2103360117013)。

**作者简介** 曾慧兰 (1986—), 女, 江西宜春人, 讲师, 博士, 从事花色机理和园艺植物病害研究。

**收稿日期** 2022-04-12

SMART<sup>[19]</sup>([http://smart.embl.de/smart/change\\_mode.pl](http://smart.embl.de/smart/change_mode.pl)) 验证跨膜区域;运用 PSORT<sup>[20]</sup>(<http://psort1.hgc.jp/form.html>) 在线软件对氨基酸序列进行亚细胞定位分析;利用 WebLogo 3<sup>[21]</sup>(<http://weblogo.threeplusone.com/create.cgi>) 进行保守区域分析;利用 SOPMA<sup>[22]</sup>([https://npsa-prabi.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=npsa\\_sopma.html](https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html)) 进行氨基酸

序列的二级结构分析;利用 SWISS-MODEL<sup>[23]</sup>(<https://swiss-model.expasy.org/>) 进行氨基酸序列的三级结构分析;利用 MEGA 11<sup>[24]</sup>(<https://www.megasoftware.net/>) 进行上述序列的多重比对,寻找保守区域,并采用邻接法(Neighbor-Joining,NJ)迭代数为 999 次建立系统发育树,分析其在遗传学上的亲缘关系。

表 1 18 种观赏植物及其类黄酮-3'5'-羟化酶基因的 GenBank 登录号

Table 1 18 ornamental plants and their GenBank accessions of flavonoid-3'5'-hydroxylase

序号 Serial number	物种 Species	拉丁名 Scientific name	科属 Family & Genus	花色 Flower color	登录号 GenBank No.	蛋白质 Protein ID
1	矮牵牛	<i>Petunia hybrida</i>	茄科矮牵牛属	红色	D14588.1	BAA03438.1
2	飞燕草	<i>Delphinium grandiflorum</i>	毛茛科飞燕草属	蓝色	AB818394.1	BAO66642.1
3	非洲堇	<i>Saintpaulia hybrid</i>	苦苣苔科非洲堇属	蓝紫色	AB596830.1	BAJ23912.1
4	洋桔梗	<i>Eustoma grandiflorum</i>	龙胆科洋桔梗属	蓝紫色	U72654.2	AAB17562.1
5	一品红	<i>Euphorbia pulcherrima</i>	大戟科大戟属	红色	KY273442.1	AUV64095.1
6	鸳鸯茉莉	<i>Brunfelsia brasiliensis</i>	茄科鸳鸯茉莉属	白蓝色	JQ678765.1	AFF59221.1
7	虎头兰×黄蝉兰	<i>Cymbidium tracyanum × Cymbidium iridioides</i>	兰科兰属	红色	MN879325.1	QKD76841.1
8	烟台翠雀花	<i>Delphinium chefoense</i>	毛茛科翠雀属	蓝色	KX825847.1	AQL59239.1
9	新墨西哥钩钟柳	<i>Penstemon neomexicanus</i>	车前科钩钟柳属	蓝色	KM388828.1	AIY51702.1
10	大花钩钟柳	<i>Penstemon grandiflorus</i>	车前科钩钟柳属	红色	KM576693.1	AIX10134.1
11	长筒蓝曼陀罗	<i>Ichroma cyaneum</i>	茄科紫铃花属	蓝色	KJ094354.1	AIY22752.1
12	荷兰鸢尾	<i>Iris hollandica</i>	鸢尾科鸢尾属	蓝色	KJ396194.1	AHX02955.1
13	细茎石斛	<i>Dendrobium moniliforme</i>	兰科石斛属	白色	HQ412560.1	AEB96145.1
14	金鱼草	<i>Antirrhinum kelloggii</i>	玄参科金鱼草属	浅紫色	AB547159.1	BAJ16328.1
15	蓝眼菊	<i>Osteospermum hybrid</i>	菊科蓝眼菊属	—	DQ250711.1	ABB29899.1
16	瓜叶菊	<i>Pericallis cruenta</i>	菊科瓜叶菊属	蓝色	DQ257626.1	ABB43030.1
17	非洲菊	<i>Gerbera hybrida</i>	菊科大丁草属	红色	DQ218417.1	ABA64468.1
18	菊花	<i>Chrysanthemum morifolium</i>	菊科菊属	红色	KF313549.1	AGX01487.1

注:“—”表示未检索到相关信息。

Note: “—” indicates that there is no accurate information on the color of the plant.

2 结果与分析

2.1 F3'5'H 理化性质分析 通过 ProtParam 在线软件对 18 种观赏植物 F3'5'H 基因所编码蛋白质的理化性质进行分析(表 2),结果表明,18 种观赏植物 F3'5'H 基因编码的氨基酸数为 502~535,平均为 510.1;相对分子量为 55 827.39~58 080.35,平均为 56 737.91;理论等电点为 6.52~9.31,平均为 8.39;不稳定性系数为 28.58~47.99,平均为 38.365;带负电残基总数(Asp+Glu)为 49~60,平均为 54.5;带正电残基

总数(Arg+Lys)为 52~66,平均为 58.9;亲水性平均系数为 -0.146~0.085,平均为 -0.041;脂肪族氨基酸指数为 90.22~106.54,平均为 96.42。从表 2 可知,有 11 种观赏植物的 F3'5'H 为稳定的亲水性蛋白质,占比 61.11%;有 4 种观赏植物的 F3'5'H 为不稳定的亲水性蛋白质,占比 22.22%;有 2 种观赏植物的 F3'5'H 为稳定的疏水性蛋白质,占比 11.11%;有 1 种观赏植物的 F3'5'H 为不稳定的疏水性蛋白质,占比 5.56%。

表 2 18 种观赏植物类黄酮-3'5'-羟化酶理化性质

Table 2 Physicochemical properties of flavonoid-3'5'-hydroxylase originated from 18 ornamental plants

序号 Serial number	物种 Species	氨基酸数 Number of amino acids	蛋白质分子式 Formula	相对分子量 Molecular weight	理论等电点 Theoretical pI	不稳定性系数 Instability index (II)	分类 Classification	带负电残基总数 (Asp+Glu) Total number of negatively charged residues	带正电残基总数 (Arg+Lys) Total number of positively charged residues	亲水性平均系数 Grand average of hydrophobicity (GRAVY)	脂肪族氨基酸指数 Aliphatic index
1	矮牵牛	506	C <sub>2536</sub> H <sub>4050</sub> N <sub>688</sub> O <sub>707</sub> S <sub>33</sub>	56 548.22	9.05	31.79	稳定	52	60	-0.058	93.72
2	飞燕草	506	C <sub>2544</sub> H <sub>4060</sub> N <sub>688</sub> O <sub>706</sub> S <sub>24</sub>	56 649.97	9.31	38.33	稳定	55	65	-0.095	94.64
3	非洲堇	508	C <sub>2571</sub> H <sub>4088</sub> N <sub>688</sub> O <sub>706</sub> S <sub>29</sub>	57 290.78	9.19	45.70	不稳定	56	66	-0.146	90.22
4	洋桔梗	510	C <sub>2566</sub> H <sub>4052</sub> N <sub>704</sub> O <sub>704</sub> S <sub>27</sub>	56 894.31	9.08	45.17	不稳定	49	57	-0.064	95.27
5	一品红	516	C <sub>2592</sub> H <sub>4126</sub> N <sub>690</sub> O <sub>729</sub> S <sub>31</sub>	57 613.32	7.64	36.88	稳定	56	57	-0.051	96.98
6	鸳鸯茉莉	506	C <sub>2541</sub> H <sub>4055</sub> N <sub>681</sub> O <sub>714</sub> S <sub>28</sub>	56 466.96	8.78	34.26	稳定	54	59	-0.039	96.96

接下表

续表 2

序号 Serial number	物种 Species	氨基酸数 Number of amino acids	蛋白质分子式 Formula	相对分子量 Molecular weight	理论等电点 Theoretical pI	不稳定系数 Instability index (II)	分类 Classification	带负电残基总数 (Asp+Glu) Total number of negatively charged residues	带正电残基总数 (Arg+Lys) Total number of positively charged residues	亲水性平均系数 Grand average of hydropathicity (GRAVY)	脂肪族氨基酸指数 Aliphatic index
7	虎头兰×黄蝉兰	503	C <sub>2517</sub> H <sub>4018</sub> N <sub>686</sub> O <sub>692</sub> S <sub>27</sub>	55 827.39	8.75	47.99	不稳定	57	62	-0.018	98.53
8	烟台翠雀花	502	C <sub>2525</sub> H <sub>4031</sub> N <sub>701</sub> O <sub>704</sub> S <sub>24</sub>	56 242.50	9.31	38.17	稳定	54	64	-0.101	94.80
9	新墨西哥钓钟柳	514	C <sub>2571</sub> H <sub>4070</sub> N <sub>696</sub> O <sub>724</sub> S <sub>34</sub>	57 404.87	8.36	36.44	稳定	55	58	-0.067	91.48
10	大花钓钟柳	514	C <sub>2581</sub> H <sub>4079</sub> N <sub>699</sub> O <sub>721</sub> S <sub>32</sub>	57 463.95	8.69	39.51	稳定	54	59	-0.083	90.91
11	长筒蓝曼陀罗	509	C <sub>2571</sub> H <sub>4083</sub> N <sub>685</sub> O <sub>709</sub> S <sub>31</sub>	56 927.72	8.84	28.58	稳定	54	60	-0.008	96.44
12	荷兰鸢尾	535	C <sub>2589</sub> H <sub>4131</sub> N <sub>729</sub> O <sub>734</sub> S <sub>27</sub>	58 080.35	8.42	44.94	不稳定	60	63	-0.043	90.99
13	细茎石斛	504	C <sub>2516</sub> H <sub>4015</sub> N <sub>683</sub> O <sub>699</sub> S <sub>28</sub>	55 914.39	6.52	46.23	不稳定	60	57	0.007	101.01
14	金鱼草	510	C <sub>2544</sub> H <sub>4048</sub> N <sub>694</sub> O <sub>709</sub> S <sub>28</sub>	56 598.03	9.07	36.00	稳定	50	58	-0.037	95.65
15	蓝眼菊	514	C <sub>2549</sub> H <sub>4049</sub> N <sub>695</sub> O <sub>724</sub> S <sub>17</sub>	56 560.43	6.53	39.72	稳定	57	53	-0.025	101.50
16	瓜叶菊	504	C <sub>2536</sub> H <sub>4034</sub> N <sub>696</sub> O <sub>706</sub> S <sub>26</sub>	56 403.73	8.67	37.35	稳定	51	55	-0.011	98.83
17	非洲菊	512	C <sub>2541</sub> H <sub>4051</sub> N <sub>689</sub> O <sub>716</sub> S <sub>19</sub>	56 318.44	8.12	28.99	稳定	53	55	0.010	101.15
18	菊花	508	C <sub>2545</sub> H <sub>4032</sub> N <sub>688</sub> O <sub>708</sub> S <sub>15</sub>	56 077.09	6.77	34.52	稳定	54	52	0.085	106.54

**2.2 F3'5'H 信号肽和跨膜区域分析** 信号肽的主要作用是促进蛋白分泌到胞外,通常位于氨基酸序列的 N 端,是引导新合成的蛋白质向分泌通路转移的短肽链<sup>[25]</sup>,一般长度为 5~30 个氨基酸。通过 SignalP 4.1 和 5.0 对 F3'5'H 编码的蛋白质进行信号肽分析(表 3),结果表明,所测的 16 种观赏植物的 F3'5'H 均无信号肽,只有虎头兰×黄蝉兰和细茎石斛具有信号肽,信号肽所在位置分别为 1~24 位和 1~20 位氨

基酸,且都与分泌(Sec/secretory)通路相关的能被一型信号肽酶(SPI)切除的信号肽。从表 3 可知,有 15 种植物的 F3'5'H 为跨膜蛋白,具有内膜区域、外膜区域和跨膜区域;有 3 种植物的 F3'5'H 无跨膜区域,只有外膜区域。经 SMART 分析验证,TMHMM 2.0 所测的跨膜区域与 SMART 分析的结构域一致。

表 3 18 种观赏植物类黄酮-3'5'-羟化酶跨膜区域与信号肽分析

Table 3 Transmembrane region and signal peptide analysis of flavonoid-3'5'-hydroxylase originated from 18 ornamental plants

序号 Serial number	物种 Species	内膜区域 Inside	外膜区域 Outside	跨膜结构 Transmembrane helix	是否有信号肽 Signal peptide	有信号肽的概率 Probability %	信号肽所在氨基酸位置 Acid position of signal peptide	信号肽类型 Type of signal peptide
1	矮牵牛	1~6,59~169	27~35,190~506	7~26,36~58,170~189	否	6.527	—	—
2	飞燕草	1~4	25~506	5~24	否	2.487	—	—
3	非洲堇	1~4	23~508	5~22	否	6.169	—	—
4	洋桔梗	—	1~510	—	否	1.263	—	—
5	一品红	1~6	30~516	7~29	否	2.430	—	—
6	鸳鸯茉莉	1~6,59~169	30~38,190~506	7~29,39~58,170~189	否	6.921	—	—
7	虎头兰×黄蝉兰	1	20~503	2~19	有	69.334	1~24	SP(Sec/SPI)
8	烟台翠雀花	1~4	23~502	5~22	否	33.727	—	—
9	新墨西哥钓钟柳	1~6	30~514	7~29	否	0.604	—	—
10	大花钓钟柳	1~6	30~514	7~29	否	0.464	—	—
11	长筒蓝曼陀罗	1,61~208	24~37,232~509	2~23,38~60,209~231	否	2.152	—	—
12	荷兰鸢尾	24~232	1~3,256~535	4~23,233~255	否	4.641	—	—
13	细茎石斛	—	1~504	—	有	33.383	1~20	SP(Sec/SPI)
14	金鱼草	—	1~510	—	否	1.194	—	—
15	蓝眼菊	22~65	1~3,89~514	4~21,66~88	否	3.187	—	—
16	瓜叶菊	1	25~504	2~24	否	11.753	—	—
17	非洲菊	22~65	1~3,89~512	4~21,66~88	否	16.254	—	—
18	菊花	1	22~508	2~21	否	11.522	—	—

注:“—”表示空值,Sec 表示“分泌通路”,SPI 表示“一型信号肽酶”。

Note: “—” indicates none, Sec represents secretory channel, SPI represents SPase I.

**2.3 F3'5'H 亚细胞定位分析** 通过 PSORT 在线分析了 8 种观赏植物 F3'5'H 的亚细胞定位(表 4),根据定位频率和可靠性可知,该类蛋白位于内质网膜的可能性最大,为

0.640~0.820,平均可靠性为 0.786;其次为质膜、内质网腔、细胞核、过氧化物酶体、细胞外、高尔基体,最后为线粒体基质空间和叶绿体类囊体膜上。结合细胞生物学知识,可预测

得 F3'5'H 可能有以下几类运输方式来行使催化功能:一类是 F3'5'H 蛋白在游离核糖体中合成后,通过信号肽锚定于内质网膜上,通过跨膜运输和膜泡方式分选到过氧化物酶体等细胞器行使催化功能,如虎头兰×黄蝉兰;第二类是在游离核糖体中合成后,通过信号肽锚定于内质网膜上,通过分泌方式分选到过氧化物酶体等细胞器行使催化功能,如细茎石斛;第三类是在游离核糖体中合成后转移至内质网膜上,以跨膜运输方式经内质网腔再直接运输到胞外行使催化功能,

如飞燕草、一品红、烟台翠雀花、新墨西哥钓钟柳、大花钓钟柳、荷兰鸢尾;第四类是在游离核糖体中合成后转移至内质网膜上,通过跨膜运输方式分选到线粒体、叶绿体、高尔基体、过氧化物酶体等细胞器后行使催化功能,如非洲堇、长筒蓝曼陀罗、蓝眼菊、瓜叶菊、菊花;第五类是直接由细胞核内合成,不经跨膜方式,可能直接在细胞质中的线粒体、叶绿体等细胞器的特定部位中行使催化功能,如洋桔梗。

表 4 18 种观赏植物类黄酮-3'5'-羟化酶亚细胞定位可靠性

Table 4 Subcellular localization certainties of flavonoid-3'5'-hydroxylase originated from 18 ornamental plants

序号 Serial number	物种 Species	内质网膜 Endoplasmic reticulum membrane	质膜 Plasma membrane	细胞核 Nucleus	线粒体 基质空间 Mitochondrial matrix space	内质网腔 Endoplasmic reticulum (lumen)	高尔基体 Golgi body	叶绿体 类囊体膜 Chloroplast thylakoid membrane	过氧化物酶体 Peroxisome	胞外 Outside
1	矮牵牛	0.820	0.190	0.108	—	0.100	—	—	—	—
2	飞燕草	0.820	0.190	—	—	0.100	—	—	—	0.100
3	非洲堇	0.685	0.640	0.108	—	—	0.460	—	—	—
4	洋桔梗	—	—	0.880	0.440	—	—	0.280	0.430	—
5	一品红	0.820	0.190	—	—	0.100	—	—	—	0.100
6	鸳鸯茉莉	0.820	0.190	0.108	—	0.100	—	—	—	—
7	虎头兰×黄蝉兰	0.820	0.190	—	—	0.100	—	—	0.144	—
8	烟台翠雀花	0.820	0.190	—	—	0.100	—	—	—	0.100
9	新墨西哥钓钟柳	0.820	0.190	—	—	0.100	—	—	—	0.100
10	大花钓钟柳	0.820	0.190	—	—	0.100	—	—	—	0.100
11	长筒蓝曼陀罗	0.820	0.190	0.108	—	—	—	—	0.132	—
12	荷兰鸢尾	0.820	0.190	—	—	0.100	—	—	—	0.100
13	细茎石斛	0.820	0.190	—	—	0.100	—	—	0.109	—
14	金鱼草	0.640	0.685	0.180	—	—	0.370	—	—	—
15	蓝眼菊	0.685	0.640	—	—	0.100	0.460	—	—	—
16	瓜叶菊	0.820	0.190	—	—	0.100	—	—	0.223	—
17	非洲菊	0.685	0.640	—	—	0.100	0.460	—	—	—
18	菊花	0.820	0.190	—	—	0.100	—	—	0.184	—

注:“—”表示空值。

Note: “—” indicates none.

**2.4 F3'5'H 二级结构和三级结构分析** 通过 SOPMA 软件在线预测了 18 种观赏植物 F3'5'H 的二级结构(表 5),表明 F3'5'H 主二级结构主要以  $\alpha$  螺旋为主,其次为无规则卷曲、延伸主链,最少的二级结构是  $\beta$  转角。通过 SWISS-MODEL 同源建模预测结果中可知(表 6),所测的所有蛋白质序列一致度均大于 30%,适合用同源建模法进行三级结构预测,即符合 SWISS-MODEL 建模质量评估条件,同时质量评估系数 GMQE 和 QMEAND 值均较大(取值范围为 0~1),建模可靠。从表 6 可知,所测的 18 种观赏植物 F3'5'H 的三级结构均建模至同一蛋白模型 5ylw.1.A 铁锈醇合成酶(图 1),为单链蛋白,属于细胞色素 P450 基因家族。

**2.5 F3'5'H 同源比对和系统发育分析** 经 MEGA 11 中 ClustalW 比对分析,18 种 F3'5'H 编码氨基酸的保守位点总计 121 个,明显的保守序列有起始于 48 位的“LPPGP”序列、起始于 334 位的“AGTDTS”序列和起始于 469 位的“PF-GAGRRICAG”序列。从 WebLogo 3 结果可知(图 2),保守序列和 ClustalW 分析序列基本一致。

表 5 18 种观赏植物类黄酮-3'5'-羟化酶二级结构分析

Table 5 Secondary structure analysis of flavonoid-3'5'-hydroxylase originated from 18 ornamental plants

序号 Serial number	物种 Species	$\alpha$ 螺旋 Alpha helix//%	延伸主链 Extended strand//%	$\beta$ 转角 Beta turn %	无规则卷曲 Random Coil//%
1	矮牵牛	50.00	9.49	4.15	36.36
2	飞燕草	48.22	12.06	3.75	35.97
3	非洲堇	48.03	11.22	4.13	36.61
4	洋桔梗	48.43	10.98	4.51	36.08
5	一品红	45.93	12.79	4.84	36.43
6	鸳鸯茉莉	49.21	11.07	3.75	35.97
7	虎头兰×黄蝉兰	48.91	10.34	4.97	35.79
8	烟台翠雀花	48.41	12.15	4.98	34.46
9	新墨西哥钓钟柳	48.83	11.87	4.47	34.82
10	大花钓钟柳	48.25	12.06	5.64	34.05
11	长筒蓝曼陀罗	48.33	11.39	5.30	34.97
12	荷兰鸢尾	44.49	11.03	4.86	39.63
13	细茎石斛	48.81	10.71	5.16	35.32
14	金鱼草	49.41	10.78	4.71	35.10
15	蓝眼菊	49.42	11.48	4.47	34.63
16	瓜叶菊	48.41	10.91	4.56	36.11
17	非洲菊	48.24	10.55	5.08	36.13
18	菊花	50.20	9.45	5.51	34.84

表 6 18 种观赏植物类黄酮-3'5'-羟化酶三级结构分析

Table 6 Tertiary structure analysis of flavonoid-3'5'-hydroxylase originated from 18 ornamental plants

序号 Serial number	物种 Species	序列一致性 Sequence identity//%	序列相似性 Sequence similarity	覆盖率 Coverage	全局模型 质量评估 Global Model Quality Estimate (GMQE)	QMEAND Co Global	预测模型模板 ID PDB Model ID
1	矮牵牛	31.62	0.36	0.92	0.68	0.70±0.05	5ylw. 1. A
2	飞燕草	33.19	0.37	0.92	0.70	0.72±0.05	5ylw. 1. A
3	非洲堇	35.20	0.38	0.88	0.69	0.74±0.05	5ylw. 1. A
4	洋桔梗	34.66	0.38	0.93	0.67	0.70±0.05	5ylw. 1. A
5	一品红	31.18	0.36	0.90	0.67	0.71±0.05	5ylw. 1. A
6	鸳鸯茉莉	31.69	0.37	0.92	0.69	0.71±0.05	5ylw. 1. A
7	虎头兰×黄蝉兰	33.33	0.37	0.93	0.71	0.73±0.05	5ylw. 1. A
8	烟台翠雀花	33.48	0.37	0.93	0.70	0.72±0.05	5ylw. 1. A
9	新墨西哥钓钟柳	32.61	0.37	0.90	0.68	0.72±0.05	5ylw. 1. A
10	大花钓钟柳	34.95	0.38	0.89	0.69	0.74±0.05	5ylw. 1. A
11	长筒蓝曼陀罗	33.11	0.37	0.87	0.67	0.71±0.05	5ylw. 1. A
12	荷兰鸢尾	33.70	0.38	0.86	0.66	0.71±0.05	5ylw. 1. A
13	细茎石斛	33.12	0.37	0.92	0.68	0.70±0.05	5ylw. 1. A
14	金鱼草	33.19	0.37	0.92	0.70	0.72±0.05	5ylw. 1. A
15	蓝眼菊	32.61	0.38	0.90	0.69	0.71±0.05	5ylw. 1. A
16	瓜叶菊	31.75	0.37	0.92	0.68	0.69±0.05	5ylw. 1. A
17	非洲菊	33.90	0.38	0.92	0.71	0.73±0.05	5ylw. 1. A
18	菊花	32.04	0.38	0.92	0.69	0.70±0.05	5ylw. 1. A

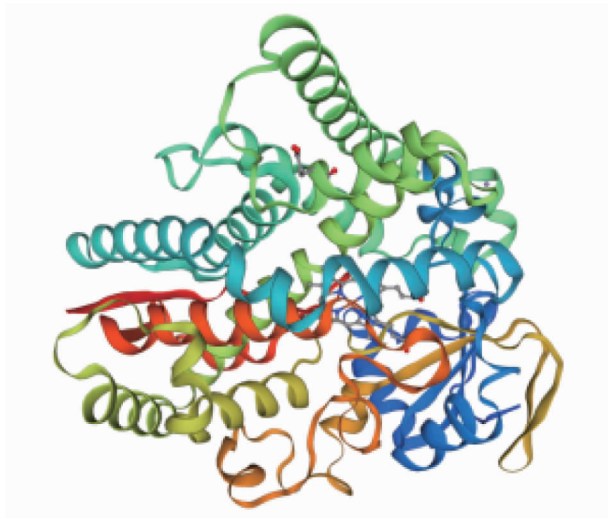


图 1 18 种观赏植物 F3'5'H 三维结构模型

Fig. 1 3D model of flavonoid-3'5'-hydroxylase originated from 18 ornamental plants

从构建的系统发育树(图 3)可知,属于同科植物的 F3'5'H 其亲缘关系较近,分属茄科、菊科、兰科、毛茛科、车前科的 13 种观赏植物在科内 F3'5'H 亲缘关系近。在不同科属间,茄科与龙胆科的洋桔梗和大戟科的一品红 F3'5'H 亲缘关系近;车前科与玄参科的金鱼草和苦苣苔科的非洲堇 F3'5'H 亲缘关系较近;鸢尾科的荷兰鸢尾与菊科 F3'5'H 亲缘关系较近。

### 3 结论与讨论

观赏植物的类黄酮-3'5'-羟化酶(F3'5'H)是花青苷代谢途径中重要的关键酶,对蓝紫色花色的形成有重要的作用。目前越来越多的研究关注 F3'5'H 的基因功能,并对其功能进行验证<sup>[3,10,14]</sup>,但国内外极少数研究从生物信息学角度关注各种观赏植物 F3'5'H 基因和蛋白质序列的差异。该研究从 GenBank 中搜索已发表的 F3'5'H 全长 mRNA 序列及

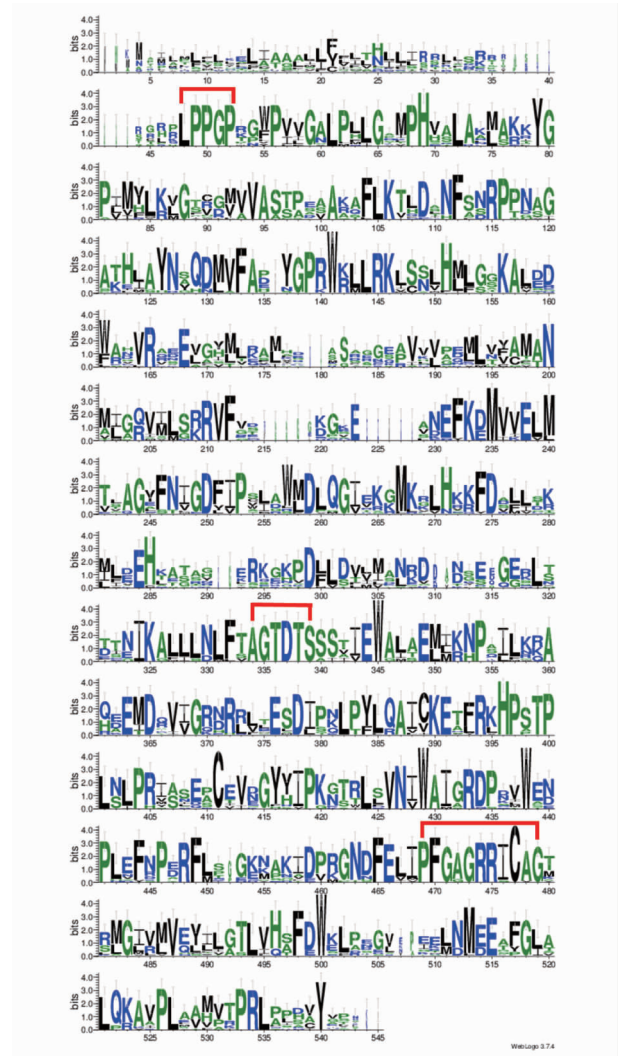


图 2 18 种观赏植物 F3'5'H 保守序列

Fig. 2 Conserved regions of flavonoid-3'5'-hydroxylase originated from 18 ornamental plants

其编码的蛋白质,共筛选出 18 种观赏植物的核酸和蛋白质序列,从理化性质、亚细胞定位、保守序列、二级结构、三级结

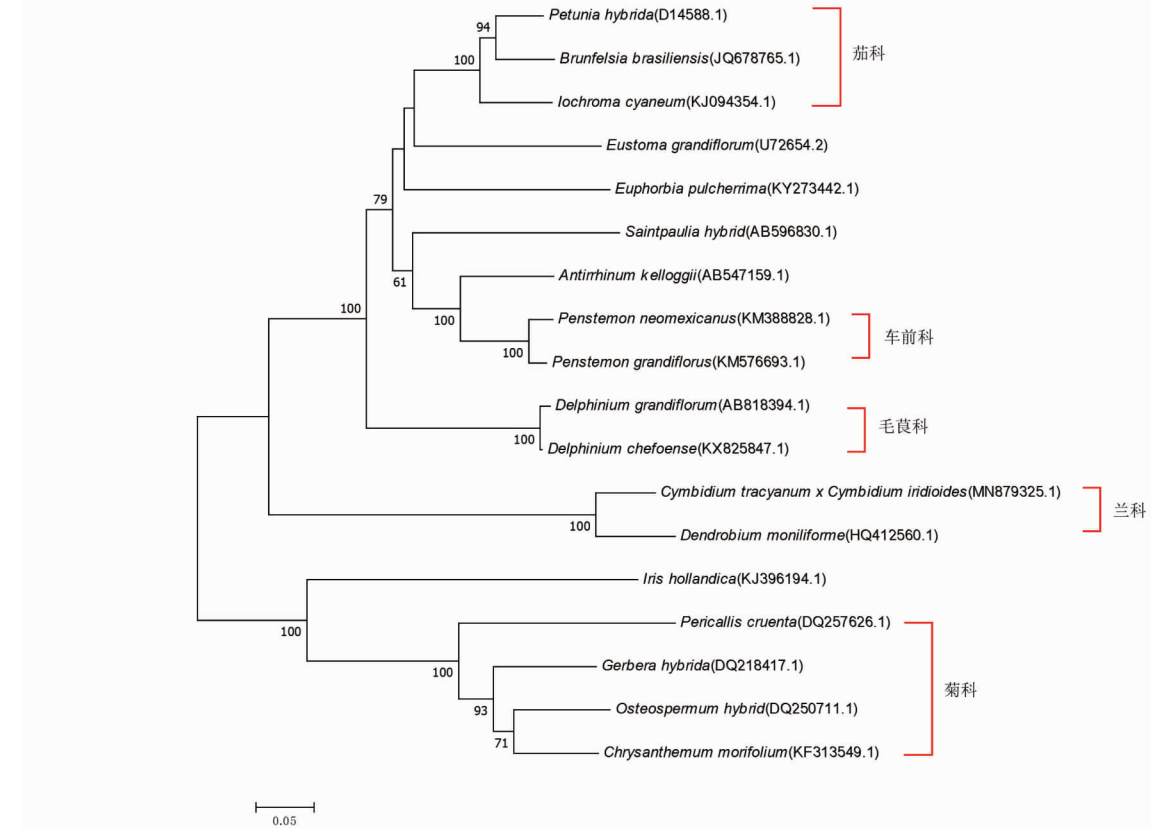


图 3 18 种观赏植物基于 F3'5'H 编码氨基酸序列的系统发育树

Fig. 3 Phylogenetic tree based on amino acid sequences of flavonoid-3'5'-hydroxylase originated from 18 ornamental plants

构、同源性等方面进行分析。从理化性质和二级结构来看,虽然不同物种间的 F3'5'H 蛋白质理化性质略有差异,但大部分物种的不稳定性系数为小于 40,亲水性平均系数为负值,且二级结构均以  $\alpha$  螺旋为主,表明该类酶为以  $\alpha$  螺旋为主的稳定亲水性单链蛋白,这与杨晓娜等<sup>[26]</sup> 分析结果相一致。从三级结构来看,所测 18 种观赏植物均建模至蛋白模型 5ylw. 1. A 铁锈醇合成酶,由此可推断,在所有的分析指标中,三维结构可能是与其功能密切相关的一个指标,为有待进一步挖掘的功能指标。跨膜结构域、信号肽和亚细胞定位分析表明,15 种植物的 F3'5'H 为 P450 基因家族跨膜蛋白,具有内膜区域、外膜区域和跨膜区域,也进一步说明了该蛋白的形成是经过了跨膜运输加工后行使功能,且虎头兰×黄蝉兰和细茎石斛有信号肽指引蛋白质通过分泌通路进行转移,这与杨晓娜等<sup>[26]</sup> 分析结果基本一致。从序列同源性特征来看,F3'5'H 有 3 个较保守的序列,分别是起始于 48 位的“LPPGP”序列连接膜锚定位点和酶蛋白球体部分的结合域、起始于 334 位的“AGTDTS”序列氧分子形成结合域和起始于 469 位的“PFGAGRRICAG”序列血红素结合域,这与杨晓娜等<sup>[26]</sup> 分析结果一致。从系统发育树来看,F3'5'H 蛋白质序列可明显表示出不同植物间的亲缘关系,保守性较高,有利于区分不同物种间的亲缘关系,该特性可能使其用于物种间的鉴定,提高物种分辨率。

近期虽然已有研究进行了高等植物 F3'5'H 基因 cDNA

及其氨基酸序列的生物学分析<sup>[26]</sup>,但从该分析来看并未将所列植物的具体信息进行详细剖析,且多数植物的信息不够具体全面,甚至有些序列与该研究相同,但得出了不同的结果。原因可能为在线软件的版本问题,或者虽然杨晓娜等<sup>[26]</sup> 的研究列出 23 种高等植物的信息,但并未将该 23 种植物的全部信息详细列出,而是进行了概述与总结。

生物信息学是采用数学、统计学和计算机等方法分析生物学、生物化学等数据的一门综合性学科。当前基因组学、转录组学和蛋白质组学测序技术的发展,促进了生物信息学的发展,而生物信息学也从各方面更好地辅助了生物学和育种技术的发展。利用生物信息学分析了 18 种观赏植物 F3'5'H 蛋白质序列的基本信息并做了详细的剖析,可为后期蓝色花的育种技术提供理论依据。

#### 参考文献

- [1] TANAKA Y, BRUGLIERA F, CHANDLER S. Recent progress of flower colour modification by biotechnology[J]. International journal of molecular sciences, 2009, 10(12): 5350-5369.
- [2] GUO L P, WANG Y J, DA SILVA J A T, et al. Transcriptome and chemical analysis reveal putative genes involved in flower color change in *Paeonia* 'Coral Sunset' [J]. Plant physiology and biochemistry, 2019, 138: 130-139.
- [3] HUANG H, HU K, HAN K T, et al. Flower colour modification of chrysanthemum by suppression of F3'H and overexpression of the exogenous *Senecio cruentus* F3'5'H gene[J]. PLoS One, 2013, 8(11): 1-12.
- [4] TANAKA Y, BRUGLIERA F. Flower colour and cytochromes P450[J]. Philosophical transactions of the royal society B: Biological sciences, 2013,

- 368:1-14.
- [5] 张泰然,张和臣,武荣花. 蓝色花形成分子机理研究进展[J]. 植物学报,2020,55(2):216-227.
- [6] SATO M,KAWABE T,HOSOKAWA M,et al. Tissue culture-induced flower-color changes in *Saintpaulia* caused by excision of the transposon inserted in the *flavonoid 3',5'-hydroxylase (F3'5'H)* promoter[J]. Plant cell reports,2011,30(5):929-939.
- [7] 李莉,祁银燕,解燕,等. 矮牵牛编码 *F3'5'H* 的蓝色基因表达载体构建及转化[J]. 西北植物学报,2011,31(6):1090-1096.
- [8] 黄敏玲,樊荣辉. 鹤望兰类黄酮 3',5'-羟化酶基因 *SrF3'5'H* 的克隆及表达分析[J]. 中国细胞生物学学报,2012,34(11):1117-1122.
- [9] WHANG S S,UM W S,SONG I J,et al. Molecular analysis of anthocyanin biosynthetic genes and control of flower coloration by flavonoid 3',5'-hydroxylase (*F3'5'H*) in *Dendrobium moniliforme*[J]. Journal of plant biology,2011,54(3):209-218.
- [10] QI Y Y,LOU Q,QUAN Y H,et al. Flower-specific expression of the *Phalaenopsis* flavonoid 3',5'-hydroxylase modifies flower color pigmentation in *Petunia* and *Lilium*[J]. Plant cell, tissue and organ culture,2013,115(2):263-273.
- [11] 严黎,刘璇菁,杨成民,等. 丹参 *F3'5'H* 基因克隆及其序列分析[J]. 世界中医药,2020,15(5):689-695,701.
- [12] 方颖,黄启群,金雪花. 华丽龙胆 *GsF3'5'H* 和 *GsFNS* 基因的克隆及表达分析[J]. 西北植物学报. 2020,40(12):2023-2030.
- [13] LIANG C Y,RENGASAMY K P,HUANG L M,et al. Assessment of violet-blue color formation in *Phalaenopsis* orchids[J]. BMC plant biology,2020,20(1):1-16.
- [14] LU C F,LI Y J,CUI Y M,et al. Isolation and functional analysis of genes involved in polyacylated anthocyanin biosynthesis in blue *Senecio cruentus*[J]. Frontiers in plant science,2021,12:1-20.
- [15] GASTEIGER E,HOOGLAND C,GATTIKER A. Protein identification and analysis tools on the ExpASY server[M]//WALKER J M. The proteomics protocols handbook. Totowa,NJ:Humana Press,2005:571-607.
- [16] PETERSEN T N,BRUNAK S,VON HEIJNE G,et al. SignalP 4.0:Discriminating signal peptides from transmembrane regions[J]. Nature methods,2011,8(10):785-786.
- [17] ALMAGRO ARMENTEROS J J,TSIRIGOS K D,SØNDERBY C K,et al. SignalP 5.0 improves signal peptide predictions using deep neural networks[J]. Nature biotechnology,2019,37(4):420-423.
- [18] MÖLLER S,CRONING M D,APWEILER R. Evaluation of methods for the prediction of membrane spanning regions[J]. Bioinformatics,2001,17(7):646-653.
- [19] LETUNIC I,KHEDKAR S,BORK P. SMART:Recent updates,new developments and status in 2020[J]. Nucleic acids research,2021,49:D458-D460.
- [20] NAKAI K,KANEHISA M. Expert system for predicting protein localization sites in gram-negative bacteria[J]. Proteins-structure function and bioinformatics,1991,11(2):95-110.
- [21] CROOKS G E,HON G,CHANDONIA J M,et al. WebLogo: A sequence logo generator[J]. Genome research,2004,14(6):1188-1190.
- [22] GEOURJON C,DELEAGE G. SOPMA:Significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments[J]. Computer applications in the biosciences,1995,11(6):681-684.
- [23] WATERHOUSE A,BERTONI M,BIENERT S,et al. SWISS-MODEL:Homology modelling of protein structures and complexes[J]. Nucleic acids research,2018,46(W1):W296-W303.
- [24] TAMURA K,STECHE G,KUMAR S. MEGA11: Molecular evolutionary genetics analysis version 11[J]. Molecular biology and evolution,2021,38(7):3022-3027.
- [25] 郑斌,詹希美. 信号肽序列及其在蛋白质表达中的应用[J]. 生物技术通讯,2005,16(3):296-298.
- [26] 杨晓娜,陈自宏,陈宏艳,等. 高等植物类黄酮 3'5'-羟化酶基因 cDNA 及其氨基酸序列的生物信息学分析[J]. 保山学院学报,2021,40(2):24-32.

## (上接第 98 页)

- [13] 孙云飞. 草鱼 (*Ctenopharyngodon idellus*) 混养系统氮磷收支和池塘水质与底质的比较研究[D]. 青岛:中国海洋大学,2013.
- [14] 刘梅,原居林,练青平,等. 池塘内循环水养殖模式对养殖塘上覆水-沉积物-间隙水磷时空分布特征及释放通量的影响[J]. 水生生物学报,2021,45(5):1045-1056.
- [15] 唐玥. 典型水生植物对水体氮磷的净化功能研究[D]. 上海:华东师范大学,2020.
- [16] 张家洋,陈丽丽,李慧. 水生植物对富营养化水体除磷去氮的研究概述[J]. 西北师范大学学报(自然科学版),2013,49(1):115-120.
- [17] 袁东海,高士祥,任全进,等. 几种挺水植物净化生活污水总氮和总磷效果的研究[J]. 水土保持学报,2004,18(4):77-80,92.
- [18] 崔虎,王莉霞,欧洋,等. 湿地生态系统磷迁移转化机制研究进展[J]. 水生生态学杂志,2020,41(2):105-112.
- [19] 李林锋,年跃刚,蒋高明. 植物吸收在人工湿地脱氮除磷中的贡献[J]. 环境科学研究,2009,22(3):337-342.
- [20] 阙天洋,王娟娟. 浅析浅水湖泊苦草-铁细菌-土壤除磷体系[J]. 安徽农学通报,2020,26(19):133-134,137.
- [21] 陈照方,陈凯,杨司嘉. 水生植物对淡水生态系统的修复效果[J]. 分子植物育种,2019,17(13):4501-4506.

## (上接第 102 页)

- [3] 中华人民共和国农业部. 鸡饲养标准:NY/T 33—2004[S]. 北京:中国农业出版社,2005:11-15.
- [4] 中华人民共和国农业部. 家禽生产性能名词术语和度量统计方法:NY/T 823—2004[S]. 北京:中国农业出版社,2005:1-7.
- [5] 黄玲玲,赵婉秋,李志梁,等. “鹤山鸡”体尺、屠宰性能、肉质测定及相关分析[J]. 中国家禽,2019,41(7):67-69.
- [6] 肖涛,李辉,唐诗雯,等. 黔东南小香鸡生长性能与肉质分析[J]. 畜牧与兽医,2021,53(10):8-13.
- [7] 谭玉文,文正亚,周敏,等. 宁都黄鸡生长、体尺与屠宰性状的典型相关分析[J]. 中国畜牧杂志,2021,57(12):102-105.
- [8] 李乃宾,杨芬霞,陈浩波,等. 热应激对 12 周龄麒麟鸡屠宰性能和肉质品质的影响[J]. 家禽科学,2013(8):11-15.
- [9] 肖燕丽,朱欣,陈光燕,等. 贵州 4 个地方鸡种体尺和屠宰性能的比较[J]. 贵州畜牧兽医,2015,39(3):19-21.
- [10] 陈宽维,李慧芳,张学余,等. 肉鸡肌纤维与肉质关系研究[J]. 中国畜牧杂志,2002,38(6):6-7.
- [11] 吴信生,陈国宏,陈宽维,等. 中国部分地方鸡种肌肉组织学特点及其肉品质的比较研究[J]. 江苏农学院学报,1998,19(4):52-58.
- [12] 姬改革,束婧婷,单艳菊,等. 鸡皮肤毛囊性状研究进展[J]. 中国家禽,2019,41(10):46-49.
- [13] 陈宽维. 后禽流感时代的黄羽肉鸡育种[J]. 中国家禽,2013,35(21):2-4.
- [14] 邹乐勤,邓继贤,徐文文,等. 广西地方鸡冰鲜屠体外观性状的测定和相关分析[J]. 黑龙江畜牧兽医,2020(20):36-38.
- [15] KALMENDAL R,TAUSON R. Effects of a xylanase and protease, individually or in combination, and an ionophore coccidiostat on performance, nutrient utilization, and intestinal morphology in broiler chickens fed a wheat-soybean meal-based diet [J]. Poultry science,2012,91(6):1387-1393.
- [16] PANDA A K,RAMA RAO S S,RAJU M V,et al. Effect of probiotic (*Lactobacillus sporogenes*) feeding on egg production and quality,yolk cholesterol and humoral immune response of White Leghorn layer breeders[J]. Journal of the science of food and agriculture,2008,88(1):43-47.
- [17] 巨晓军. IgY 在雏鸡生长早期器官中转移规律及其对生长性能和免疫机能的影响[D]. 扬州:扬州大学,2016.