

李娜,樊裕璇,韩红艳,等.不同种植物中 *HMGR* 基因的生物信息学分析[J].山西农业科学,2025,53(3):27-43.

LI N, FAN Y X, HAN H Y, et al. Bioinformatics analysis of *HMGR* genes in different plants[J]. Journal of Shanxi Agricultural Sciences, 2025, 53(3):27-43.

doi:10.3969/j.issn.1002-2481.2025.03.04

## 不同种植物中 *HMGR* 基因的生物信息学分析

李娜<sup>1</sup>,樊裕璇<sup>1</sup>,韩红艳<sup>1</sup>,宋宇琴<sup>2</sup>,李六林<sup>2</sup>,赵红梅<sup>1</sup>

(1.晋中学院生物科学与技术系,山西晋中 030609;2.山西农业大学园艺学院,山西晋中 030801)

**摘要:**为鉴定不同种植物中的 *HMGR*(3-Hydroxy-3-methylglutaryl coenzyme A reductase)基因,采用生物信息学方法分析拟南芥、番茄、马铃薯等18种植物中 *HMGR* 蛋白的理化性质,并对其跨膜结构、亚细胞定位、修饰位点等特性进行分析,预测18种植物 *HMGR* 基因编码蛋白的高级结构和保守结构域。结果表明,甜瓜 *HMGR* 蛋白序列最长,由1244个氨基酸组成,草莓 *HMGR* 蛋白序列最短,由131个氨基酸组成;玉米 *HMGR* 蛋白的等电点最高,为10.12,草莓 *HMGR* 蛋白的等电点最低,为4.72;除芒果、玉米、小麦、杨树和桃 *HMGR* 蛋白带正电荷之外,其余植物均带负电荷。除玉米、杨树、草莓之外,其余植物的脂肪族指数均在90以上,推断 *HMGR* 大多为疏水蛋白;*HMGR* 蛋白在芒果中最稳定,在梨和杏中的最强疏水性最高,在杨树中的最强亲水性最高;除草莓外,其余植物中的 *HMGR* 都具有跨膜结构,其中甜瓜跨膜结构位置最多。对蛋白质修饰位点的分析结果显示,草莓中的 *HMGR* 蛋白的糖基化位点数量最少,杏的糖基化修饰位点最多,拟南芥的磷酸化位点数量最多,草莓的磷酸化位点数量最少,二级结构主要由 $\alpha$ -螺旋和无规卷曲构成。*HMGR* 基因基本被分为3类,且除玉米外,17种植物 *HMGR* 蛋白均包含 *HMG-CoA reductase Class I* 结构域。

**关键词:** *HMGR* 基因;生物信息学;理化性质;基因编码蛋白;保守结构域

中图分类号:Q943.2 文献标识码:A 文章编号:1002-2481(2025)03-0027-17

### Bioinformatics Analysis of *HMGR* Genes in Different Plants

LI Na<sup>1</sup>, FAN Yuxuan<sup>1</sup>, HAN Hongyan<sup>1</sup>, SONG Yuqin<sup>2</sup>, LI Liulin<sup>2</sup>, ZHAO Hongmei<sup>1</sup>

(1.Department of Biological Science and Technology, Jinzhong University, Jinzhong 030609, China;

2.College of Horticulture, Shanxi Agricultural University, Jinzhong 030801, China)

**Abstract:** To identify *HMGR*(3-Hydroxy-3-methylglutaryl coenzyme A reductase) genes in different plants, in this study, bioinformatics methods were used to analyze the physicochemical properties of *HMGR* protein in 18 plants including *Arabidopsis*, tomato, and potato. The characteristics of their transmembrane structure, subcellular localization, and modification sites were analyzed. The higher-order and conserved domains of the proteins encoded by *HMGR* genes in 18 plants were predicted. The results indicate that the *HMGR* protein sequence of melon was the longest, consisting of 1244 amino acids, while the *HMGR* protein sequence of strawberry was the shortest, consisting of 131 amino acids. The isoelectric point of the *HMGR* protein in corn was the highest, at 10.12, and that of strawberry was the lowest, at 4.72. Except for mango, corn, wheat, poplar, and peach, the *HMGR* proteins of other plants carried a negative charge. Except for corn, poplar, and strawberry, the aliphatic index of the *HMGR* proteins of other plants were above 90, suggesting that most *HMGR* proteins were hydrophobic. The *HMGR* protein was the most stable in mango. The strongest hydrophobicity was the highest in pear and apricot. The strongest hydrophilicity was the highest in poplar. Except for strawberry, *HMG*Rs in other plants had transmembrane structures, with the highest number of transmembrane structure locations in melon. The analysis of protein modification sites showed that the *HMGR* protein in strawberry had the fewest glycosylation sites. Apricot had the most glycosylation modification sites. *Arabidopsis* had the highest number of phosphorylation sites, while strawberry had the lowest number of phosphorylation sites. The secondary

收稿日期:2024-06-06

基金项目:山西省高等学校科技创新项目(2023L313);山西省基础研究计划(自由探索类)(202303021222261)

作者简介:李娜,讲师,博士,主要从事园艺植物功能物质开发研究,E-mail:lnjzyswx@163.com

通信作者:李六林,教授,博士,主要从事园艺植物栽培生理与品质提升研究,E-mail:tgliulin@163.com

structure was mainly composed of  $\alpha$ -helices and random coils. The *HMGR* gene was divided into 3 classes and the HMGR protein of 17 plants, except for corn, contained the HMG-CoA reductase Class I domain.

**Keywords:** *HMGR* gene; bioinformatics analysis; physicochemical properties; gene-encoded proteins; conserved domains

萜类化合物是由不同数量异戊二烯单元组成的一种天然化合物,广泛存在于自然界<sup>[1]</sup>。目前,科学家从生物中提取、鉴定出的萜类化合物超过 10 万种,包括单萜、倍半萜、二萜、二倍半萜、三萜和多萜等<sup>[1-2]</sup>。萜类化合物具有经济价值和药用价值,被广泛地应用于化妆品、食品和医药行业。从红豆杉树皮中首次提取的紫杉醇,是一类四环二萜类化合物,目前被应用于多种癌症的治疗<sup>[3]</sup>。

萜类化合物合成是通过类异戊二烯途径合成的,以五碳焦磷酸异戊烯基(Isopentenyl pyrophosphate, IPP)及其异构体二磷酸二甲基烯丙基(Dimethyl allyl diphosphate, DMAPP)为基本的合成单元<sup>[4-5]</sup>。IPP 和 DMAPP 的生物合成则通过 2 个途径来完成,即甲基赤藓糖醇 4-磷酸/脱氧木酮糖 5-磷酸途径(Methylerythritol 4-phosphate/deoxyxylulose 5-phosphate, MEP/DOXP)和甲羟戊酸酯(mevalonate, MEV 或 MVA)途径。MEP 通路仅在内质体运行,而 MVA 通路分布在细胞质、内质网和过氧化物酶体中<sup>[5]</sup>。3-羟基-3-甲基戊二酸单酰辅酶 A 还原酶(3-hydroxy-3-methylglutaryl co-enzyme A reductase, HMGR)是植物萜类代谢途径中的关键限速酶,可催化生成 MVA,在植物的整个生长发育中都起到重要作用<sup>[6-7]</sup>。*HMGR* 基因在多种植物中均有表达,在植物生长发育<sup>[8]</sup>、果实发育期间细胞的分裂<sup>[8]</sup>、果实成熟<sup>[8-9]</sup>、抗逆境<sup>[10-11]</sup>、抗病虫害<sup>[12]</sup>等方面发挥着关键作用。

*HMGR* 基因已从多个物种中被克隆出来,如

酿酒酵母<sup>[13]</sup>、土沉香<sup>[14]</sup>等。目前,已经对多个植物生物合成的关键酶 HMGR 的生物学功能进行了初步研究。郭思远等<sup>[15]</sup>研究表明, MVA 可以调节细胞生长,而 *HMGR* 基因介导的 MVA 生物合成是萜类生物合成的重要途径,并可启动植物内质网的形成。该基因的表达量对植物萜类化合物的生物合成有直接的影响。HMGR 可通过调控萜类物质的生物合成进而影响植物的香气,通过影响 MEP 途径和花色苷的合成调控色泽等<sup>[16]</sup>。而光照、病原体和机械损伤等也会反过来影响 *HMGR* 基因的表达<sup>[17]</sup>。

目前,通过生物信息学手段对基因进行研究已成为国内外的热点。首先,利用生物信息学方法,在多种生物学数据库中获得 *HMGR* 的序列信息、结构信息和功能注释;其次,利用生物信息学技术,结合生物信息学分析、序列分析、结构域预测、蛋白二级结构预测等技术,鉴定并预测 HMGR 的保守区域、功能位点、亚细胞定位、二级结构、结构域等关键特性。为深入理解 *HMGR* 基因的功能和作用机制提供了有力的支持。

## 1 材料和方法

### 1.1 试验材料

本试验选择的不同植物中 *HMGR* 基因序列和编号来自 NCBI(<https://www.ncbi.nlm.nih.gov/>)、Uniprot(<https://www.uniprot.org/>)2 个网站(表 1)。

表 1 材料信息

Tab.1 Material information

物种	Species	基因编号	Gene ID	染色体定位	Chromosome localization
拟南芥	Arabidopsis( <i>Arabidopsis thaliana</i> (L.) Heynh)	LOC843982			NC_003070.9
番茄	Tomato( <i>Solanum lycopersicum</i> L.)	LOC543702			NC_015439.3
马铃薯	Potato( <i>Solanum tuberosum</i> L.)	LOC102577654			NW_006239312.1
苹果	Apple( <i>Malus pumila</i> Mill.)	LOC103421562			NC_041803.1
芒果	Mango( <i>Mangifera indica</i> L.)	LOC123213049			NC_058140.1
甜瓜	Melon( <i>Cucumis melo</i> L.)	LOC103502601			NC_066859.1
玉米	Corn( <i>Zea mays</i> L.)	LOC100280335			NC_050099.1
小麦	Wheat( <i>Triticum aestivum</i> L.)	LOC123151295			NC_057812.1

续表 1 材料信息  
Tab.1(Continued) Material information

物种	Species	基因编号	Gene ID	染色体定位	Chromosome localization
辣椒	Chilli( <i>Capsicum annuum</i> L.)	LOC107866760		NC_061114.1	
大豆	Soybean( <i>Glycine max</i> (L.) Merr.)	LOC100815480		NC_038252.2	
烟草	Tobacco( <i>Nicotiana tabacum</i> L.)	LOC107832729		NW_015864719.1	
杨树	Poplar( <i>Populus trichocarpa</i> L.)	LOC7489052		NC_037295.2	
石榴	Pomegranate( <i>Punica granatum</i> L.)	LOC116199997		NC_045129.1	
桃	Peach( <i>Prunus persica</i> (L.) Batsch)	LOC18770980		NC_034015.1	
枣	Jujube( <i>Ziziphus jujuba</i> Mill.)	LOC107425621		NC_083387.1	
杏	Apricot( <i>Prunus dulcis</i> L.)	LOC117636586		NC_047657.1	
梨	Pear( <i>Pyrus bretschneideri</i> )	LOC103949512		NW_025999391.1	
草莓	Strawberry( <i>Fragaria vesca</i> Duch.)	LOC105349649		NC_020492.1	

## 1.2 蛋白质理化性质及亲水性分析

在 ProtParam3.0 在线平台 (<https://web.expasy.org/protparam/>) 分析蛋白质的理化性质;在 ProtScale 3.0 在线平台 (<https://web.expasy.org/protscale/>) 分析蛋白质的亲水性和疏水性。

## 1.3 蛋白质跨膜结构域及亚细胞定位分析

在德泰生物科技有限公司平台 (<http://www detaibio.com/tools/transmembrane.html>) 上分析蛋白质的跨膜结构域,再用 PSORT Prediction (<https://www.genscript.com/psort.html>) 分析蛋白质的亚细胞定位。

## 1.4 蛋白质修饰位点预测

用 NetOGlyc4.0 (<https://www.genscript.com/psort.html>) 分析预测 HMGR 蛋白的 O-糖基化位点<sup>[18]</sup>;用 NetNGlyc1.0 (<https://services.healthtech.dtu.dk/services/NetNGlyc-1.0/>) 预测 HMGR 蛋白的 N-糖基化位点;用 NetPhos3.1 (<https://services.healthtech.dtu.dk/services/NetPhos-3.1/>) 预测 HMGR 蛋白的磷酸化位点。

## 1.5 蛋白质结构预测分析

用 Prabi (<https://doua.prabi.fr/software/cap3>) 分析预测蛋白质的二级结构;用 PDB (<https://www.rcsb.org/>)、RaptorX (<http://raptorx.uchicago.edu/StructPredV2/predict/>) 或 Swiss-Model (<https://swissmodel.expasy.org/interactive>) 预测蛋白质的三级结构。

## 1.6 进化树的构建以及保守结构域的预测分析

首先,用 ClustalX 对蛋白质序列进行比对<sup>[19]</sup>;其次,用 DNAMAN 进行比较,绘制比对图谱;用

MEGA11 作系统发育进化树;用 MEME 在线网站 (<http://meme-suite.org/tools/meme>) 分析预测蛋白质的保守结构域。

## 2 结果与分析

### 2.1 蛋白质理化性质及亲水性分析

HMGR 基因编码蛋白的理化性质结果如表 2 所示,甜瓜的序列最长,由 1 244 个氨基酸组成;草莓的序列最短,由 131 个氨基酸组成;其余植物蛋白序列为 343~608 个氨基酸不等。甜瓜的分子质量最大,为 135 546.89 u;草莓的分子质量最小,为 13 958.05 u;其余植物的分子质量在 37 510.64~64 978.66 u。玉米蛋白质的等电点最高,为 10.12;草莓最低,为 4.72。甜瓜最大氨基酸组成是苏氨酸,占比 34.9%;玉米最大氨基酸组成是精氨酸,占比 13.9%;辣椒、烟草和石榴最大氨基酸组成是丝氨酸,分别占比 10.2%、9.9% 和 9.8%;大豆、杨树和桃最大氨基酸组成为亮氨酸,分别占比 10.1%、9.9% 和 10.5%;草莓最大氨基酸组成为谷氨酸,占比 11.5%;其余植物最大氨基酸组成均为丙氨酸,占比在 8.2%~11.7%。除芒果、玉米、小麦、杨树和桃蛋白质带正电荷之外,其余植物均带负电荷。苹果、芒果、桃、枣、杏和梨的不稳定系数低于 40,较稳定;拟南芥、番茄、马铃薯、甜瓜、玉米、小麦、辣椒、大豆、烟草、杨树、石榴和草莓的不稳定系数都高于 40,处于不稳定状态。除玉米、杨树、草莓之外,其余植物的脂肪族指数均在 90 以上,推断 HMGR 大多为疏水蛋白。

表 2 蛋白质理化性质

Tab.2 Physicochemical properties of proteins

蛋白来源 Protein source	氨基酸序 列长度/aa Amino acid sequence length	分子质量/u Molecular weight	蛋白质 等电点 Protein iso- electric point	最大氨基酸组 成/占比/% Maximum amino acid composi- tion/proportion	Asp、Glu总数/个 Total number of Asp, Glu	Arg、Lys总数/个 Total number of Arg, Lys	不稳定 系数 Instabil- ity factor	脂肪族 指数 Aliphatic index
拟南芥 Arabidopsis	592	63 597.73	6.23	Ala/9.5	59	56	43.19	90.29
番茄 Tomato	343	37 510.64	5.74	Ala/8.2	37	35	56.60	96.09
马铃薯 Potato	596	63 916.95	6.31	Ala/9.1	59	57	44.76	95.22
苹果 Apple	608	64 936.95	6.24	Ala/10.2	62	59	35.99	92.22
芒果 Mango	556	59 182.33	7.83	Ala/9.2	51	53	30.64	93.13
甜瓜 Melon	1 244	135 546.89	6.68	Thr/34.9	126	123	44.47	90.88
玉米 Corn	583	63 478.83	10.12	Arg/13.9	68	87	60.24	75.08
小麦 Wheat	557	59 033.00	7.90	Ala/11.7	53	55	49.92	97.56
辣椒 Chilli	570	61 437.43	5.99	Ser/10.2	55	49	51.07	91.07
大豆 Soybean	587	62 640.87	6.83	Leu/10.1	55	54	42.51	95.04
烟草 Tobacco	604	64 732.58	6.64	Ser/9.9	59	58	46.24	93.31
杨树 Poplar	588	62 704.15	7.52	Leu/9.9	54	55	43.94	89.76
石榴 Pomegranate	584	62 270.47	6.48	Ser/9.8	59	57	43.52	90.80
桃 Peach	562	60 608.79	7.53	Leu/10.5	54	55	36.10	96.87
枣 Jujube	603	64 677.38	6.00	Ala/10.1	62	58	38.25	91.71
杏 Apricot	605	64 687.63	6.63	Ala/10.2	62	61	34.37	93.34
梨 Pear	608	64 978.66	6.24	Ala/10.4	62	59	36.70	92.38
草莓 Strawberry	131	13 958.05	4.72	Glu/11.5	19	13	56.96	86.41

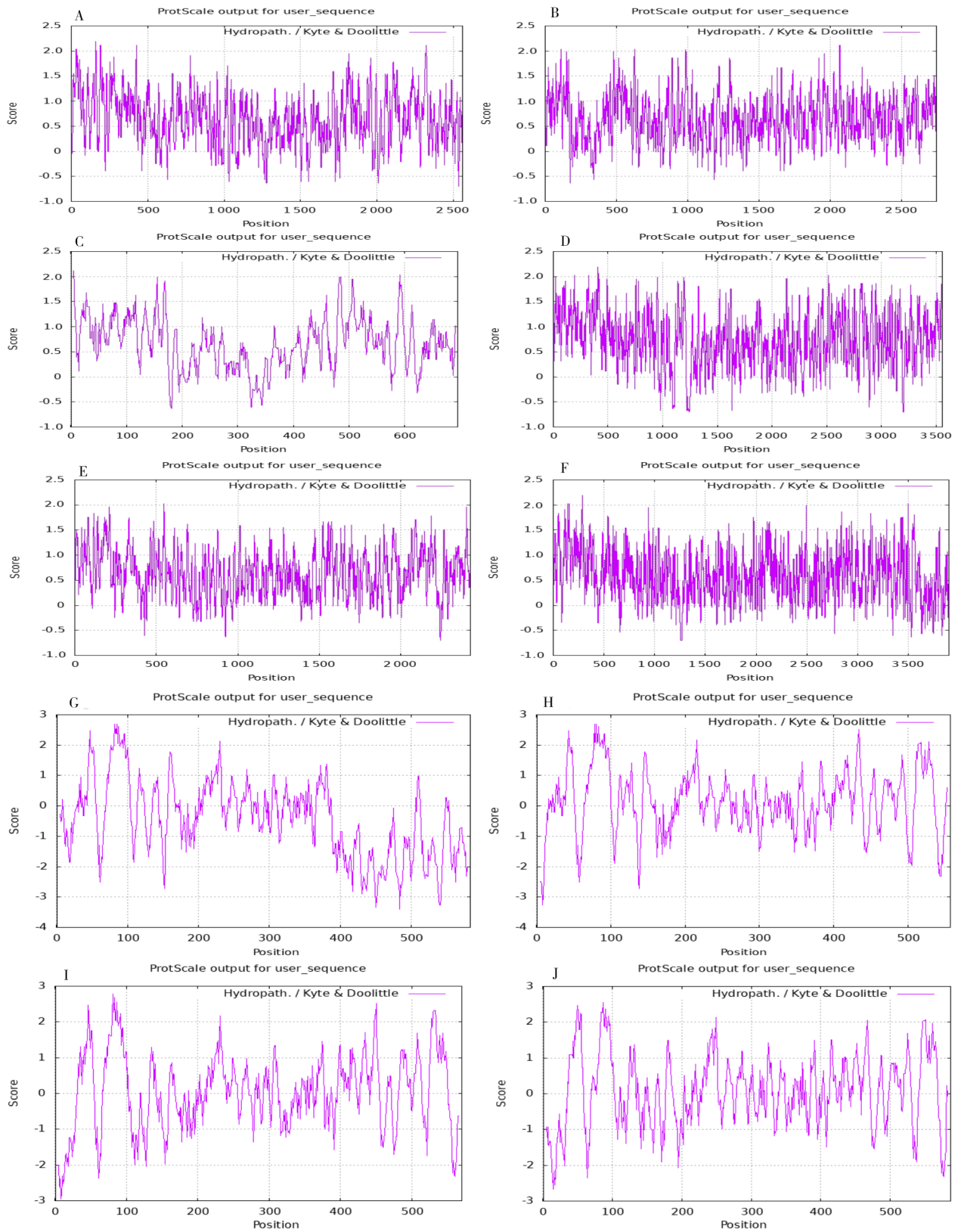
对不同种植物中 HMGR 的亲水性分析结果发现(表 3、图 1、2),玉米和草莓的总平均亲水性为负数;桃的总平均亲水性最低,为 0.038;苹果的总平均亲水性最高,为 0.734;其余植物的总平均亲

水性均在 0.042~0.664 范围内。最强疏水性均位于 Ile(异亮氨酸)上,其中,杏和梨的最强疏水性最高,均为 3.267;最强亲水性均位于 Arg(精氨酸)上,其中杨树的最强亲水性最高,为-3.589。

表 3 蛋白质亲水性

Tab.3 Protein hydrophilicity

蛋白来源 Protein source	总平均亲水性 Total average hydrophilicity	疏水性最强 氨基酸 The most hydro- phobic amino acid	疏水性最强 氨基酸位置 The position of the most hydro- phobic amino acids	最强疏水性 Strongest hy- drophobicity	亲水性最强 氨基酸 The most hy- drophilic amino acid	亲水性最强 氨基酸位置 The position of the most hydro- philic amino acids	最强亲水性 Strongest hy- drophilicity
拟南芥 Arabidopsis	0.641	Ile	161,162	2.189	Arg	2 534	-0.700
番茄 Tomato	0.653	Ile	2 067	2.111	Arg	178	-0.633
马铃薯 Potato	0.664	Ile	5	2.111	Arg	181	-0.633
苹果 Apple	0.734	Ile	410	2.189	Arg	1 243,1 244	-0.700
芒果 Mango	0.642	Ile	548	2.022	Arg	2 244,2 245	-0.700
甜瓜 Melon	0.060	Ile	96,721	2.611	Arg	76,701	-2.667
玉米 Corn	-0.465	Ile	82,85	2.689	Arg	484	-3.400
小麦 Wheat	0.142	Ile	78,81	2.689	Arg	8	-3.256
辣椒 Chilli	0.057	Ile	81	2.778	Arg	8	-2.944
大豆 Soybean	0.073	Ile	87	2.544	Arg	15	-2.667
烟草 Tobacco	0.083	Ile	90	2.800	Arg	71	-2.400
杨树 Poplar	0.046	Ile	88	2.622	Arg	17,18	-3.589
石榴 Pomegranate	0.042	Ile	55	2.789	Arg	6	-2.889
桃 Peach	0.038	Ile	53	2.556	Arg	19	-3.278
枣 Jujube	0.079	Ile	95	3.033	Arg	26	-2.889
杏 Apricot	0.110	Ile	91	3.267	Arg	128	-2.767
梨 Pear	0.090	Ile	91	3.267	Arg	128	-3.578
草莓 Strawberry	-0.085	Ile	101	1.956	Arg	23	-2.556

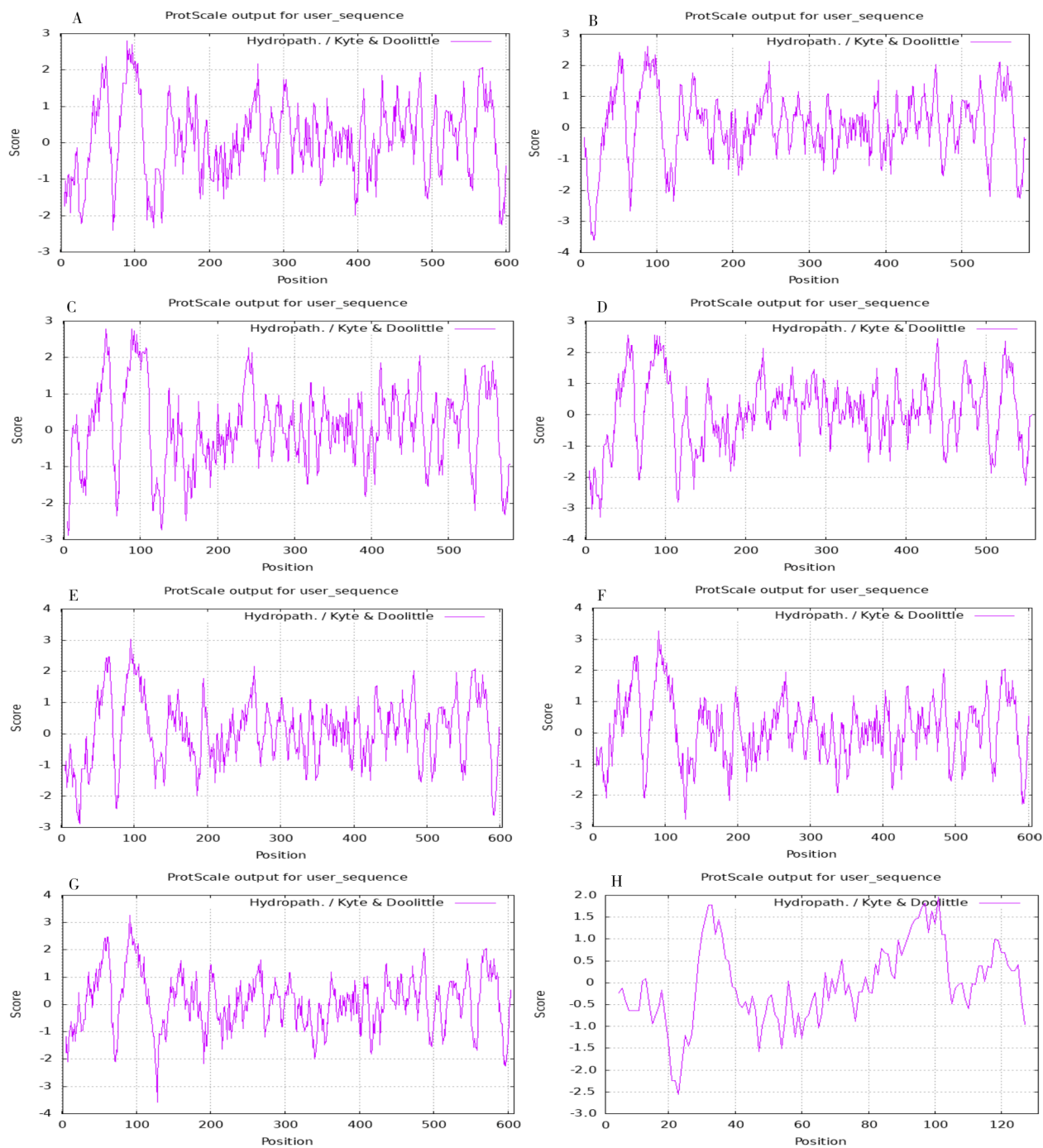


A~J分别为拟南芥、番茄、马铃薯、苹果、芒果、甜瓜、玉米、小麦、辣椒、大豆

A~J were arabidopsis, tomato, potato, apple, mango, melon, corn, wheat, chilli, soybean, respectively

图1 10种植物HMGR基因的亲水性分析

Fig.1 Hydrophilic analysis of HMGR gene of 10 plants



A~H 分别为烟草、杨树、石榴、桃、枣、杏、梨、草莓

A~H were tobacco, poplar, pomegranate, peach, jujube, apricot, pear, strawberry, respectively

图 2 8 种植物 *HMGR* 基因的亲水性分析

Fig.2 Hydrophilic analysis of *HMGR* gene of 8 plants

## 2.2 蛋白质跨膜结构及亚细胞定位分析

通过对拟南芥、番茄等 18 种植物的 *HMGR* 蛋白进行亚细胞定位, 结果发现(表 4), 18 个品种中除草莓外, 其余植物中的 *HMGR* 都具有跨膜结构。其中, 番茄、马铃薯、苹果、芒果、甜瓜、烟草、

杨树和杏的 *HMGR* 定位在内质网上; 小麦、石榴、桃和枣的亚细胞定位在叶绿体上; 拟南芥亚细胞定位在叶绿体和线粒体上; 玉米的亚细胞定位在过氧化物酶体上; 辣椒定位在线粒体上; 大豆定位在细胞质中; 梨定位在细胞质和内质网上。

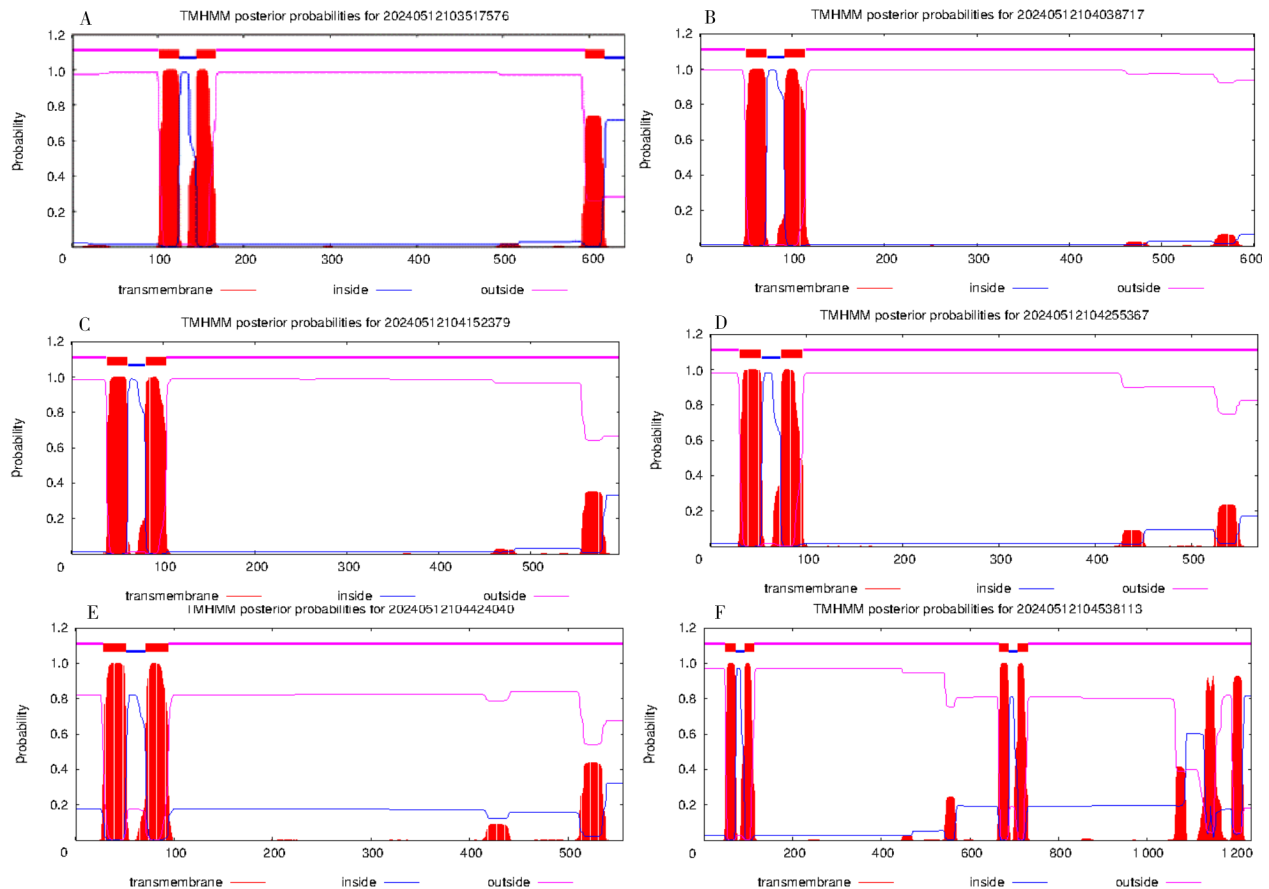
表4 蛋白质跨膜结构及亚细胞定位

Tab.4 Transmembrane structure and subcellular localization of proteins

蛋白来源 Protein source	跨膜结构位置 Transmembrane structure location	亚细胞定位 Subcellular localization	蛋白来源 Protein source	跨膜结构位置 Transmembrane structure location	亚细胞定位 Subcellular localization
拟南芥 Arabidopsis	102-124, 145-167, 596-618	叶绿体、线粒体	大豆 Soybean	37-59, 72-94	细胞质
番茄 Tomato	50-72, 92-114	内质网	烟草 Tobacco	44-66, 86-108	内质网
马铃薯 Potato	39-61, 81-103	内质网	杨树 Poplar	38-60, 81-103	内质网
苹果 Apple	31-53, 74-96	内质网	石榴 Pomegranate	42-64, 85-107	叶绿体
芒果 Mango	29-51, 72-94	内质网	桃 Peach	40-62, 83-105	叶绿体
甜瓜 Melon	49-71, 92-114, 665-687, 708-730	内质网	枣 Jujube	48-70, 91-113	叶绿体
玉米 Corn	38-57, 78-100	过氧化物酶体	杏 Apricot	44-66, 87-109	内质网
小麦 Wheat	74-96	叶绿体	梨 Pear	44-66, 87-109	细胞质、内质网
辣椒 Chilli	34-56, 77-99	线粒体	草莓 Strawberry		

跨膜区预测结果中有 Inside、Outside、Transmembrane 等 3 个值。从图 3 和图 4 可以看出,有 3 种颜色区域,其中,Inside(蓝色)代表胞内区,值越大,说明该氨基酸更可能定位在胞内区;Outside(紫色)代表胞外区,值越大,说明该氨基酸更有可

能定位在胞外区;Transmembrane(红色)代表跨膜区,值越大,说明该氨基酸更有可能存在于跨膜区。这 18 种植物中,除草莓外,红色面积均较大,代表大部分存在跨膜结构,其中甜瓜跨膜结构位置最多,而草莓几乎不存在跨膜结构。

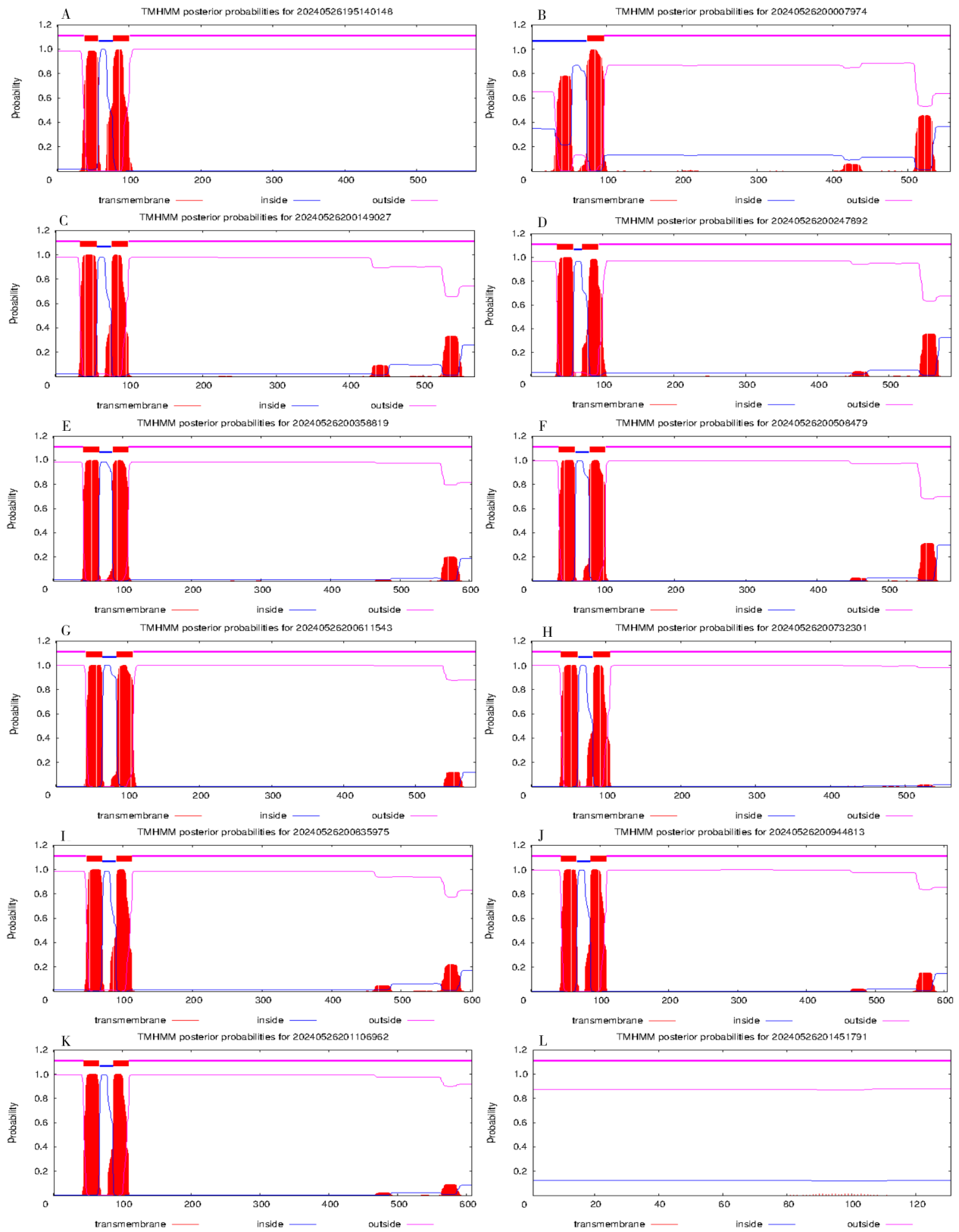


A~F 分别为拟南芥、番茄、马铃薯、苹果、芒果、甜瓜

A-F were arabidopsis, tomato, potato, apple, mango, melon, respectively

图3 6种植物HMGR基因的跨膜结构分析

Fig.3 Transmembrane structure analysis of HMGR gene of 6 plants



A~L 分别为玉米、小麦、辣椒、大豆、烟草、杨树、石榴、桃、枣、杏、梨、草莓

A~L were corn, wheat, chilli, soybean, tobacco, poplar, pomegranate, peach, jujube, apricot, pear, strawberry, respectively

图 4 12 种植物 *HMGR* 基因的跨膜结构分析

Fig.4 Transmembrane structure analysis of *HMGR* gene of 12 plants

### 2.3 O、N-糖基化位点、磷酸化位点预测

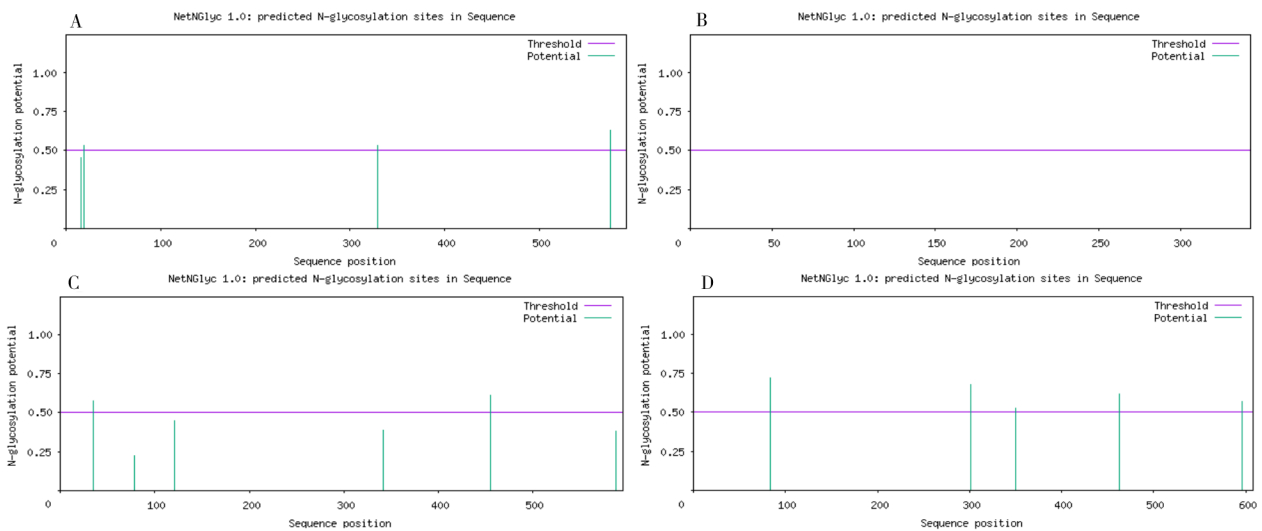
蛋白质的糖基化能修饰蛋白质,调节蛋白质功能。N-糖基化起始于内质网,完成于高尔基体;然而,O-糖基化修饰仅发生于高尔基体区。对蛋白质修饰位点的分析结果显示(表5、图5~8),草莓中的HMGR蛋白的糖基化位点数量最少,N-糖基

化0个,O-糖基化1个;杏的糖基化修饰位点最多,N-糖基化4个,O-糖基化4个。拟南芥的磷酸化位点数量最多,为67个;草莓的磷酸化位点数量最少,为17个。甜瓜的丝氨酸(S)数量最多,为43个;拟南芥的苏氨酸(T)数量最多,为26个;马铃薯、甜瓜和烟草的酪氨酸(Y)数量最多,均为7个。

表5 蛋白质修饰位点

Tab.5 Protein modification sites

蛋白来源 Protein source	糖基化位点数量/个 Number of glycosylation sites		磷酸化位点数量/个 Number of phosphorylation sites			
	N-糖基化 N-glycosylation	O-糖基化 O-glycosylation	丝氨酸 Serine	苏氨酸 Threonine	酪氨酸 Tyrosine	总数 Total
拟南芥 Arabidopsis	3	3	35	26	6	67
番茄 Tomato	0	2	24	10	4	38
马铃薯 Potato	2	1	35	16	7	58
苹果 Apple	5	1	28	19	4	51
芒果 Mango	2	3	31	14	2	47
甜瓜 Melon	3	2	43	12	7	62
玉米 Corn	1	3	27	8	1	36
小麦 Wheat	2	4	31	13	2	46
辣椒 Chilli	3	2	38	16	6	60
大豆 Soybean	3	1	40	17	2	59
烟草 Tobacco	3	4	42	17	7	66
杨树 Poplar	3	3	35	25	5	65
石榴 Pomegranate	3	1	36	14	6	56
桃 Peach	3	2	24	17	5	46
枣 Jujube	3	2	29	19	4	52
杏 Apricot	4	4	27	18	4	49
梨 Pear	4	3	30	19	4	53
草莓 Strawberry	0	1	10	5	2	17

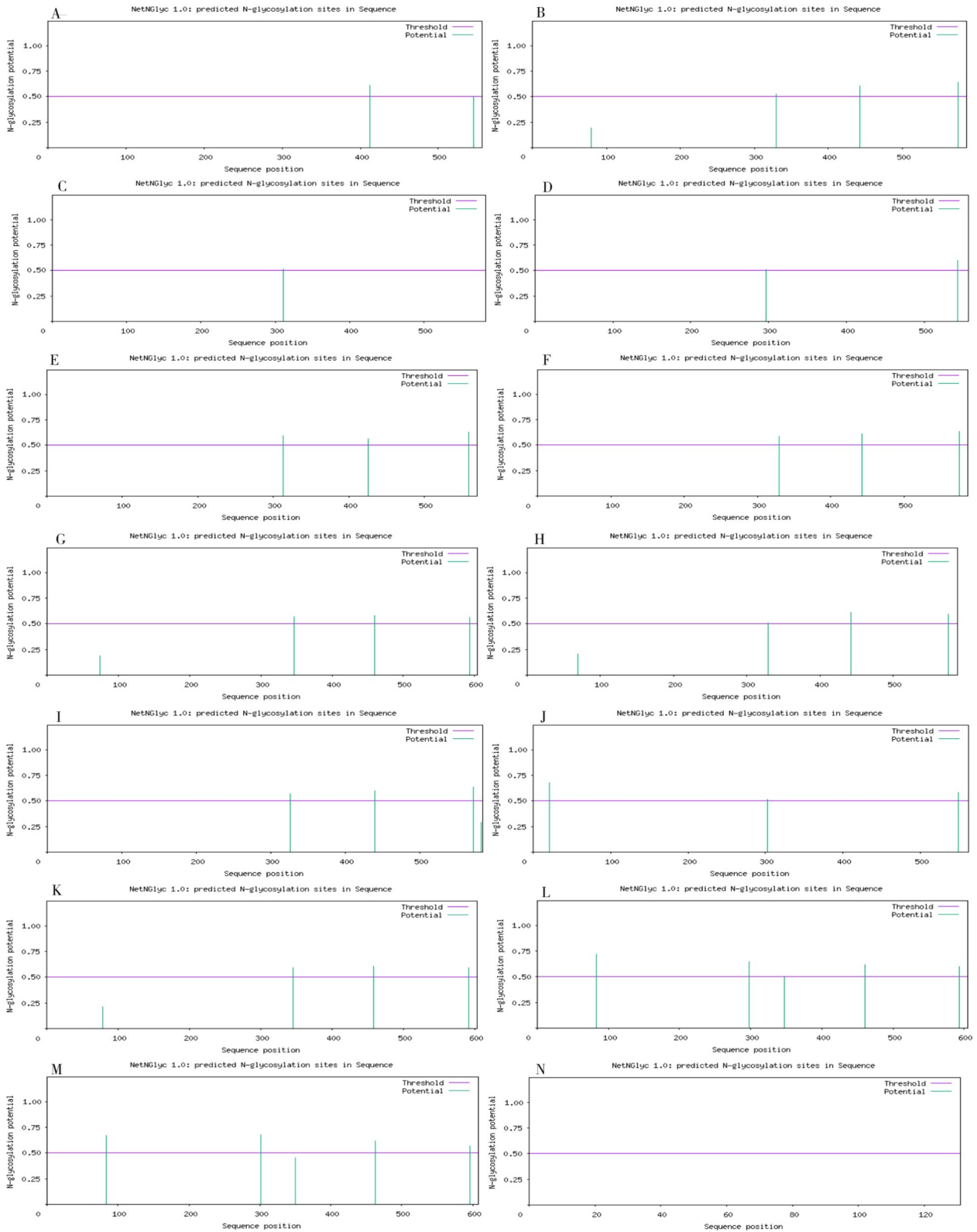


A~D 分别为拟南芥、番茄、马铃薯、苹果

A~D were arabidopsis, tomato, potato, apple, respectively

图5 4种植物HMGR基因N-糖基化位点预测

Fig.5 Prediction of N-glycosylation site of HMGR gene of 4 plants

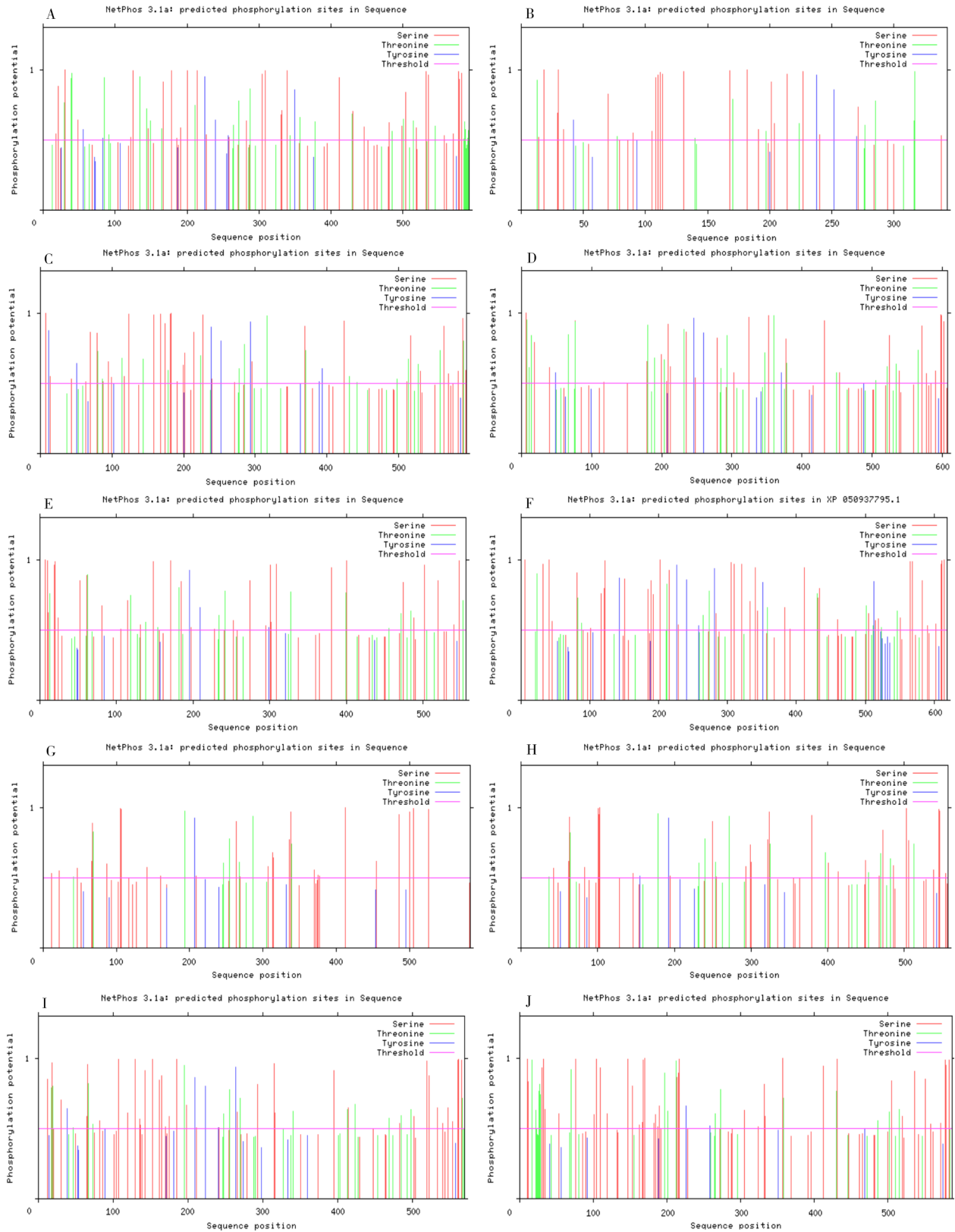


A~N 分别为芒果、甜瓜、玉米、小麦、辣椒、大豆、烟草、杨树、石榴、桃、枣、杏、梨、草莓

A~N were mango, melon, corn, wheat, chilli, soybean, tobacco, poplar, pomegranate, peach, jujube, apricot, pear, strawberry, respectively

图 6 14 种植物 HMGR 基因 N-糖基化位点预测

Fig.6 Prediction of N-glycosylation site of HMGR gene of 14 plants

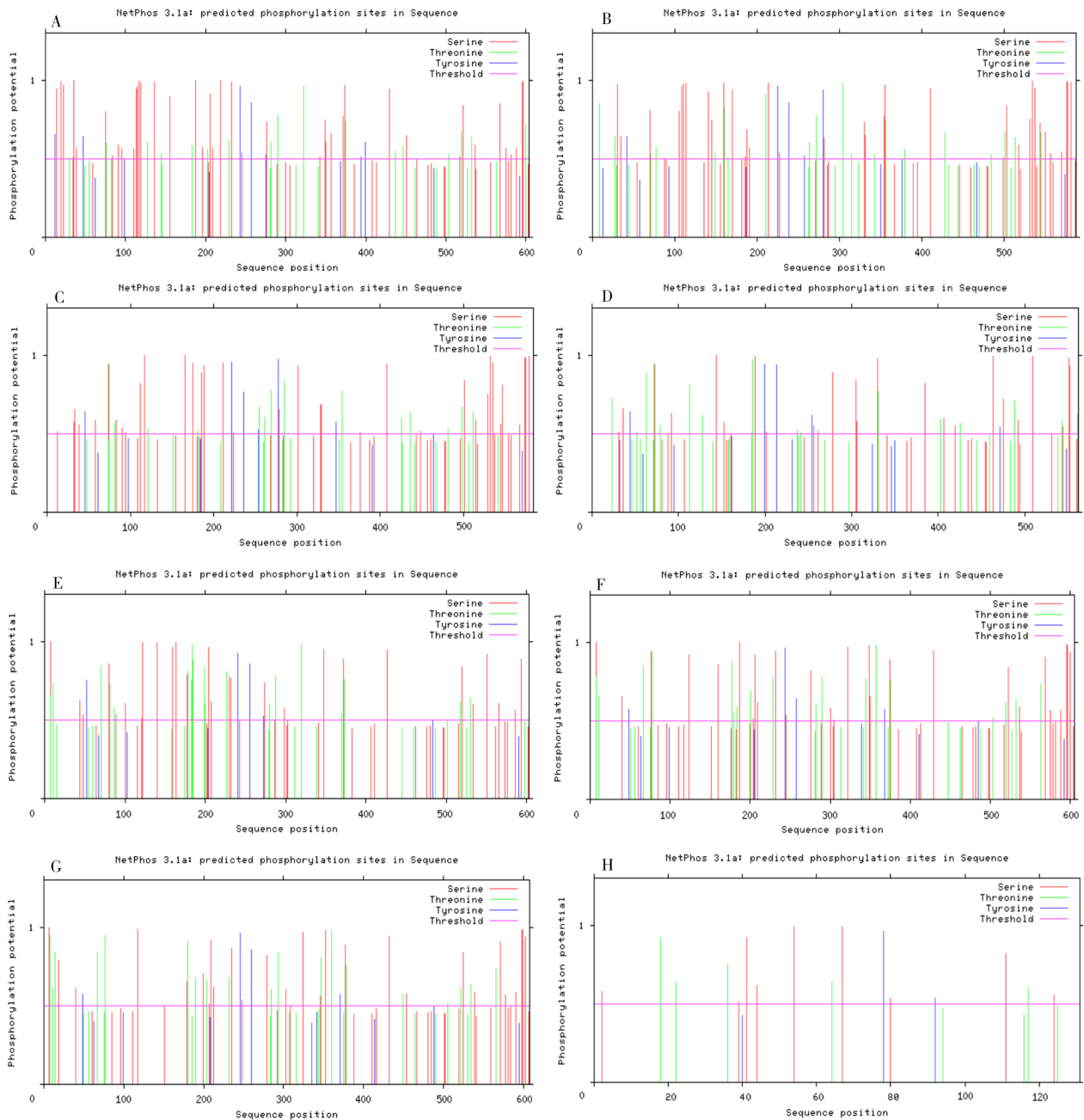


A~J 分别为拟南芥、番茄、马铃薯、苹果、芒果、甜瓜、玉米、小麦、辣椒、大豆

A~J were arabidopsis, tomato, potato, apple, mango, melon, corn, wheat, chilli, soybean, respectively

图 7 10 种植物 *HMGR* 基因的磷酸化位点预测

Fig.7 Phosphorylation site prediction of *HMGR* gene of 10 plants



A~H分别为烟草、杨树、石榴、桃、枣、杏、梨、草莓

A-H were tobacco, poplar, pomegranate, peach, jujube, apricot, pear, strawberry, respectively

图8 8种植物HMGR基因的磷酸化位点预测

Fig.8 Phosphorylation site prediction of HMGR gene of 8 plants

#### 2.4 蛋白质二级及三级结构预测

蛋白质二级结构预测结果显示(表6、图9),小麦的 $\alpha$ -螺旋在所有二级结构中占比最高,为47.40%;玉米最低,为34.99%;其余植物介于37.40%~46.22%。玉米中 $\beta$ -转角占比较高,为7.89%;烟草中 $\beta$ -转角占比最低,为4.14%;其余16种植物占比介于4.36%~6.97%,总体 $\beta$ -转角占

比差异不大。无规卷曲占比在31.42%~45.97%;扩展链占二级结构的比例较低,为11.15%~15.44%。总的来说,在 $\alpha$ -螺旋、 $\beta$ -转角、无规卷曲和扩展链这4种二级结构中,占比最高的是 $\alpha$ -螺旋,其次是无规卷曲,二者共占比约80%,说明在这18种植物中,二级结构主要由 $\alpha$ -螺旋和无规卷曲构成,还含有少部分的 $\beta$ -转角和扩展链。

表6 蛋白质二级结构

Tab.6 Protein secondary structure

%

蛋白来源 Protein source	$\alpha$ -螺旋 $\alpha$ -Spiral	$\beta$ -转角 $\beta$ -Corners	无规卷曲 Random curl	扩展链 Extended chains	蛋白来源 Protein source	$\alpha$ -螺旋 $\alpha$ -Spiral	$\beta$ -转角 $\beta$ -Corners	无规卷曲 Random curl	扩展链 Extended chains
拟南芥 Arabidopsis	41.59	4.83	38.63	14.95	大豆 Soybean	43.78	5.79	35.95	14.48
番茄 Tomato	41.43	4.99	38.77	14.81	烟草 Tobacco	40.07	4.14	43.05	12.75
马铃薯 Potato	43.12	4.36	40.77	11.74	杨树 Poplar	44.56	6.97	34.52	13.95
苹果 Apple	46.22	5.45	34.97	13.36	石榴 Pomegranate	42.47	5.31	38.01	14.21
芒果 Mango	42.99	5.22	37.77	14.03	桃 Peach	42.88	5.16	37.19	14.77
甜瓜 Melon	44.12	5.28	35.78	14.82	枣 Jujube	41.96	4.98	38.47	14.59
玉米 Corn	34.99	7.89	45.97	11.15	杏 Apricot	43.64	4.46	36.86	15.04
小麦 Wheat	47.40	5.75	31.42	15.44	梨 Pear	42.93	4.77	37.34	14.97
辣椒 Chilli	42.98	4.39	37.19	15.44	草莓 Strawberry	37.40	4.58	43.51	14.50

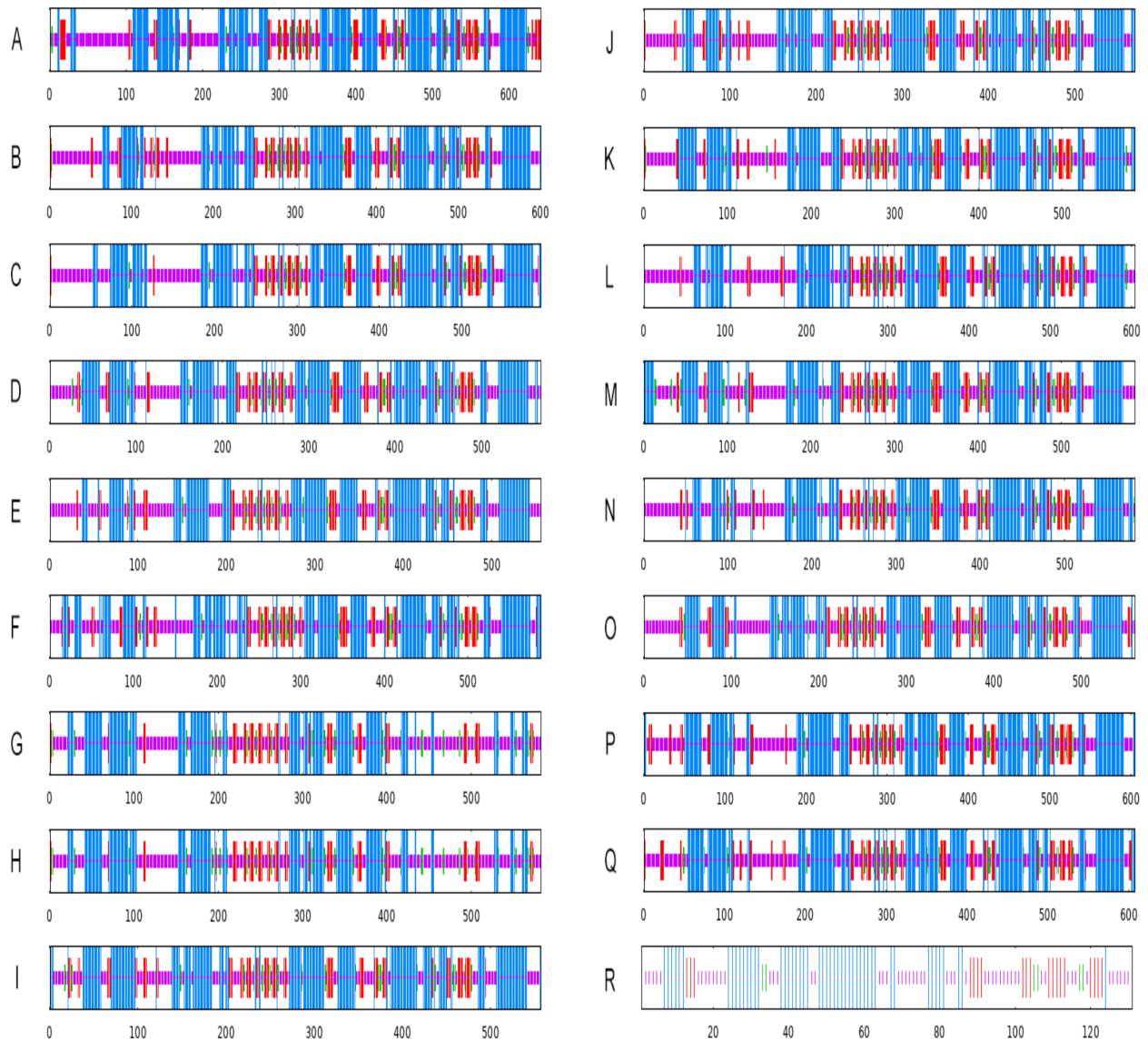


图9 蛋白质二级结构

Fig.9 Protein secondary structure

蛋白质的三级结构是在其二级或超二级结构上,受疏水作用、氢键、范德华力、静电作用等力的影响而发生盘曲折叠而成的一种稳定的空间结构。将从 Uniprot 中获取的不同植物 HMGR 基因的蛋白质序列分别通过 Swiss Model 进行三级结

构预测,结果表明(图 10),HMGR 主要是由  $\alpha$ -螺旋通过无规卷曲连接在一起,除番茄和草莓外,其余 18 种植物 HMGR 均有  $\alpha$ -螺旋、 $\beta$ -转角、无规则卷曲和延伸链组成的复杂结构区。

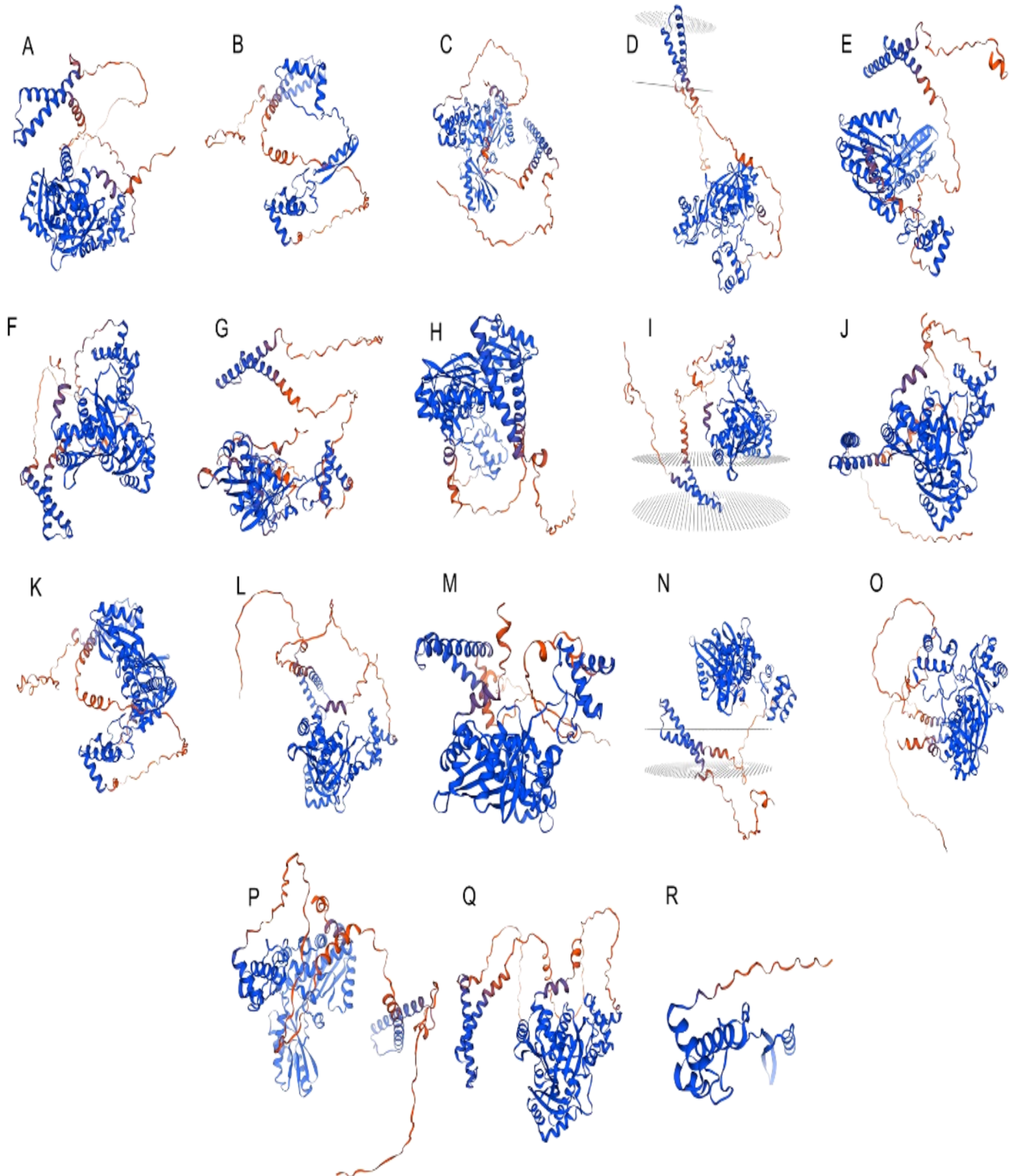


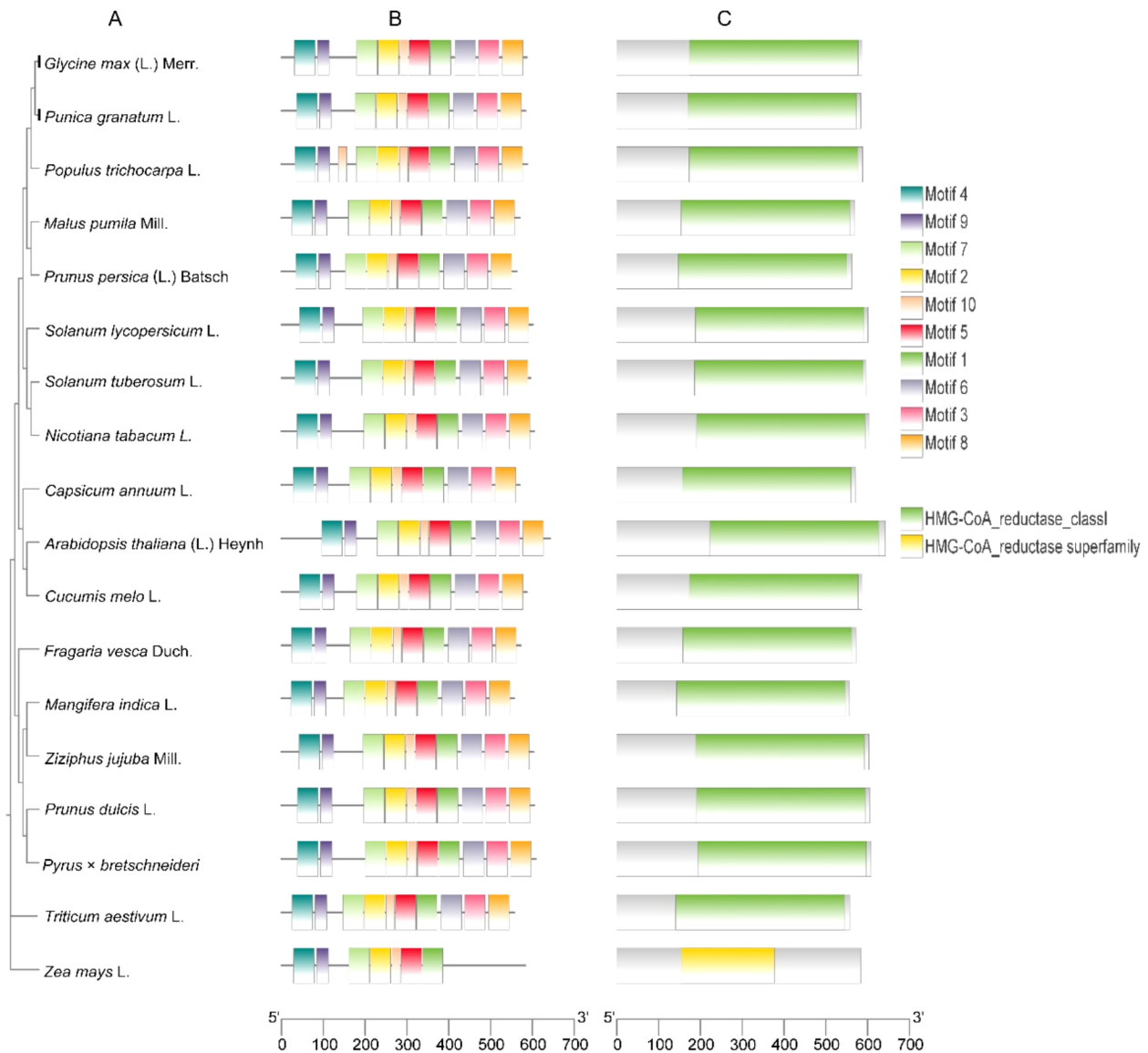
图 10 HMGR 蛋白三级结构

Fig.10 Tertiary structure of HMGR protein

## 2.5 进化树的构建及保守结构域预测分析

对18种植物保守结构域及进化关系进行分析发现, HMGR蛋白基本被分成三类, 小麦和玉米HMGR蛋白各被分成一类, 其他植物HMGR蛋白被分成一类。16种植物HMGR蛋白所在的类群中又被分成两类, 其中, 草莓、芒果、枣、杏和梨这5种植物HMGR蛋白被分成一类, 其余11种植物

HMGR蛋白被分成一类。除玉米HMGR蛋白只具有7个motif外, 其余17种植物HMGR蛋白均有10个motif, 且杨树具有2个motif10, 猜测玉米HMGR蛋白在进化过程中丢失或者发生插入。除玉米外, 17种植物HMGR蛋白均包含HMG-CoA reductase Class I结构域(图11)。



A. 系统发育进化树; B. 蛋白基序; C. 保守结构域

A. Phylogenetic tree; B. Protein motif; C. Conserved domains

图11 HMGR蛋白系统发育进化树及保守结构域预测

Fig.11 Phylogenetic tree and conserved domain prediction of HMGR protein

## 3 结论与讨论

HMGR是一个庞大的基因家族, 在该家族中,

HMGR基因的表达水平很有可能影响其他萜类合成终产物的比例。姜鸣<sup>[20]</sup>研究发现, HMGR基因的表达水平有助于提高鞘翅(Coleoptera)昆虫体

内单萜类物质的合成。*HMGR* 基因的表达水平也可以通过影响萜烯、萜醇、萜醛、萜酮和萜酯等萜类化合物合成,从而影响植物的香气<sup>[2,21]</sup>。笔者以拟南芥、番茄、马铃薯、苹果、芒果和甜瓜等 18 种植物为例,用生物信息学方法研究其 *HMGR* 蛋白的理化性质、亲水性和疏水性、亚细胞定位以及二、三级结构预测等,结果表明,这 18 种植物中的 *HMGR* 蛋白大多为疏水蛋白,总平均亲水性介于 -0.465~0.734。总平均亲水性是将蛋白中所有氨基酸的亲水性值之和除以序列中的残基数<sup>[22]</sup>,蛋白的疏水性强有助于建立本身的三级结构。不稳定性指数是指蛋白质在体外测试中稳定性的参考值,当预测值小于 40 时,提示蛋白质的稳定性好。这个预测值是基于 GURUPRASAD 等<sup>[23]</sup>的研究结果,该研究对 12 种不稳定蛋白和 32 种稳定蛋白进行统计分析,发现不稳定蛋白与稳定蛋白中存在一定的二肽。根据这些二肽对不稳定蛋白质的影响,为每个二肽分配一个不稳定的权重值。对于给定的蛋白质,这些权重值的总和归一化到其序列的长度,有助于区分不稳定和稳定的蛋白质<sup>[23]</sup>。本研究中除苹果、芒果、桃、枣、杏和梨之外,其余植物的 *HMGR* 蛋白均处于不稳定状态。亚细胞定位大多位于内质网和叶绿体上,除草莓外,其余 17 种植物均具有跨膜结构,这与赵荣荣<sup>[24]</sup>、盖江涛等<sup>[25]</sup>研究结果一致。CAMPOS 等<sup>[26]</sup>研究表明,植物 *HMGR* 主要针对内质网,在膜中具有相同的拓扑结构。定位于内质网的 *HMGR* 亚基可能起到信号肽的作用,指导内质网完成蛋白结构修饰<sup>[27]</sup>。本研究预测到拟南芥、小麦、石榴、桃和枣 5 种植物的 *HMGR* 蛋白定位于叶绿体,可能与 *HMGR* 影响植物成熟和果实色泽形成的功能有关。18 种植物的糖基化位点最大值和最小值分别为 8 个和 1 个,磷酸化位点最小值和最大值也有较大差别,分别为 17 个和 67 个;中华豆芫菁 *HMGR* 蛋白预测有 2 个 O-糖基化位点,4 个 N-糖基化位点和 2 个 O-GlcNAc 糖基化位点,与植物中 *HMGR* 蛋白有相似之处<sup>[20]</sup>。在二级结构中,除了玉米、烟草和草莓的无规卷曲占比大于  $\alpha$ -螺旋之外,其余 15 种植物中 *HMGR* 蛋白的二级结构均是  $\alpha$ -螺旋占比最高,无规卷曲次之, $\beta$ -转角占比最少。通过分析发现,不同品种植物中 *HMGR* 基因进化相对保守,因此,其对不同植物的影响机理也比较类似。

## 参考文献:

- [1] 刘晶晶,崔光周,段旺军,等.生态因素影响烟草香味特征的研究进展[J].河南农业科学,2023,52(2):1-11.  
LIU J J, CUI G Z, DUAN W J, et al. Research progress on the effects of ecological factors on tobacco aroma characteristics[J]. Journal of Henan Agricultural Sciences, 2023, 52(2): 1-11.
- [2] 郑婷,魏灵珠,程建徽,等.植物 3-羟基-3-甲基戊二酰辅酶 A 还原酶 (*HMGR*) 研究进展[J].植物生理学报,2022,58(6):1037-1044.  
ZHENG T, WEI L Z, CHENG J H, et al. Research progress of 3-hydroxy-3-methylglutaryl-CoA reductase in plants[J]. Plant Physiology Journal, 2022, 58(6): 1037-1044.
- [3] 孙凡淑,党舒扬,郑东然,等.紫杉醇生物合成与转录调控机制研究进展[J].生物工程学报,2024,40(5):1380-1405.  
SUN F S, DANG S Y, ZHENG D R, et al. Advances in paclitaxel biosynthesis and transcriptional regulation mechanisms[J]. Chinese Journal of Biotechnology, 2024, 40(5): 1380-1405.
- [4] LU J, YAN S, XUE Z. Biosynthesis and functions of triterpenoids in cereals[J]. Journal of Advanced Research, 2025, 71: 155-171.
- [5] RODRÍGUEZ-CONCEPCIÓN M. Early steps in isoprenoid biosynthesis: multilevel regulation of the supply of common precursors in plant cells[J]. Phytochemistry Reviews, 2006, 5(1): 1-15.
- [6] ZHANG M, LIU H, WANG Q, et al. The 3-hydroxy-3-methylglutaryl-coenzyme A reductase 5 gene from *Malus domestica* enhances oxidative stress tolerance in *Arabidopsis thaliana*[J]. Plant Physiology and Biochemistry, 2020, 146: 269-277.
- [7] DONG N, PONCIANO G, MCMAHAN C M, et al. Overexpression of 3-hydroxy-3-methylglutaryl coenzyme A reductase in *Parthenium argentatum* (guayule) [J]. Industrial Crops and Products, 2013, 46: 15-24.
- [8] 刘展. 苹果 *HMGR1* 基因克隆及在番茄中的功能分析[D].泰安:山东农业大学,2016.  
LIU Z. Gene cloning of *MdHMGR1* from apple and its functional analysis in tomato[D]. Tai'an: Shandong Agricultural University, 2016.
- [9] 陈俊宇,姚玲窈,许张婷,等.掌叶覆盆子 *HMGR* 基因家族的鉴定及表达分析[J].核农学报,2024,38(6):1024-1034.  
CHEN J Y, YAO L T, XU Z T, et al. Identification and expression analysis of *HMGR* gene family in *Rubus chingii* Hu[J]. Journal of Nuclear Agricultural Sciences, 2024, 38(6): 1024-1034.
- [10] GHASEMI S, KUMLEH H H, KORDROSTAMI M, et al. Drought stress-mediated alterations in secondary metabolites and biosynthetic gene expression in cumin plants: Insights from gene-specific and metabolite-level analyses[J]. Plant Stress, 2023, 10: 100241.
- [11] 刁秀楠. 大豆 *HMGR* 基因家族的克隆、表达分析和功能探究[D].太原:山西农业大学,2019.  
DIAO X N. Cloning, expression analysis and functional exploration of the *HMGR* gene family in *Glycine max*[D]. Taigu: Shanxi Agricultural University, 2019.
- [12] 陈小保,李帅广,易善泽,等.阳春砂 *HMGR* 基因 cDNA 克隆及生物信息学分析[J].暨南大学学报(自然科学与医学版),

- 2016, 37(4):285-292.
- CHEN X B, LI S G, YI S Z, et al. Molecular cloning and bioinformatic analysis of *HMGR* gene in *Amomum villosum*[J]. Journal of Jinan University (Natural Science & Medicine Edition), 2016, 37(4):285-292.
- [13] 梁会超, 胡宗凤, 梁兰, 等. 过表达HMGR催化结构域以优化酵母工程菌人参二醇代谢途径的研究[J]. 药学研究, 2016, 35(8):444-448.
- LIANG H C, HU Z F, LIANG L, et al. Optimization of the protopanoxadiol metabolic pathway by overexpression of the catalytic domain of HMGR in *Saccharomyces cerevisiae*[J]. Journal of Pharmaceutical Research, 2016, 35(8):444-448.
- [14] XU Y H, YANG X, WEI J H, et al. Cloning and expression analysis on 3-hydroxy-3-methylglutaryl-coenzyme A reductase from *Aquilaria sinensis*[J]. Chinese Herbal Medicines, 2013, 5(3):182-188.
- [15] 郭思远, 闫琦, 李佳. 东北雷公藤 *DXR* *HMGR* 基因克隆及生物信息学分析[J]. 中国现代中药, 2019, 21(11):1482-1488.
- GUO S Y, YAN Q, LI J. Cloning and bioinformatic analysis of *DXR* and *HMGR* genes from *Tripterygium regelii*[J]. Modern Chinese Medicine, 2019, 21(11):1482-1488.
- [16] 武莹. 雷公藤 *HMGR* 基因克隆、表达及调控萜类次生代谢的初步探究[D]. 杨凌:西北农林科技大学, 2012.
- WU Y. Cloning of genes encoding *HMGR* from *Tripterygium wilfordii* and its metabolic regulation[D]. Yangling: Northwest A&F University, 2012.
- [17] HOLMBERG N, HARKER M, WALLACE A D, et al. Co-expression of N-terminal truncated 3-hydroxy-3-methylglutaryl CoA reductase and C24-sterol methyltransferase type 1 in transgenic tobacco enhances carbon flux towards end-product sterols[J]. The Plant Journal, 2003, 36(1):12-20.
- [18] HANSEN J E, LUND O, TOLSTRUP N, et al. NetOglyc: prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility[J]. Glycoconjugate Journal, 1998, 15(2):115-130.
- [19] KOHLI D K, BACHHAWAT A K. CLOURE: Clustal output reformatter, a program for reformatting ClustalX/ClustalW outputs for SNP analysis and molecular systematics[J]. Nucleic Acids Research, 2003, 31(13):3501-3502.
- [20] 姜鸣. 中华豆芫菁 *HMGR* 基因克隆及斑蝥素生物合成途径初探[D]. 杨凌:西北农林科技大学, 2012.
- JIANG M. Cloning of *HMGR* gene of *Epicauta Chinensis* and preliminary study of biosynthesis pathway of cantharidin[D]. Yangling: Northwest A&F University, 2012.
- [21] KIM Y J, LEE O R, OH J Y, et al. Functional analysis of 3-hydroxy-3-methylglutaryl coenzyme A reductase encoding genes in triterpene saponin-producing ginseng[J]. Plant Physiology, 2014, 165(1):373-387.
- [22] KYTE J, DOOLITTLE R F. A simple method for displaying the hydropathic character of a protein[J]. Journal of Molecular Biology, 1982, 157(1):105-132.
- [23] GURUPRASAD K, BHASKER REDDY B V, PANDIT M W. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting *in vivo* stability of a protein from its primary sequence[J]. Protein Engineering, Design and Selection, 1990, 4(2):155-161.
- [24] 赵荣荣. 梨 *TPS* 和 *HMGR* 基因家族鉴定及 *PcHMGR1* 基因功能研究[D]. 泰安:山东农业大学, 2018.
- ZHAO R R. The identification of *TPS* and *HMGR* gene family in pear and functional analysis of *PcHMGR1* Gene[D]. Tai'an: Shandong Agricultural University, 2018.
- [25] 盖江涛, 王鹏, 罗海燕, 等. 芒果 *HMGR* 家族基因的鉴定、系统进化及表达分析[J]. 分子植物育种, 2021, 19(16):5279-5285.
- GAI J T, WANG P, LUO H Y, et al. Identification, phylogenetic evolution and expression analysis of the mango (*Mangifera indica*) *HMGR* family genes[J]. Molecular Plant Breeding, 2021, 19(16):5279-5285.
- [26] CAMPOS N, BORONAT A. Targeting and topology in the membrane of plant 3-hydroxy-3-methylglutaryl coenzyme A reductase[J]. The Plant Cell, 1995, 7(12):2163-2174.
- [27] 鞠世杰, 阎秀峰. 高等植物的 3-羟基-3-甲基戊二酰辅酶 A 还原酶[J]. 植物生理学通讯, 2004, 40(1):104-110.
- JU S J, YAN X F. 3-hydroxy-3-methylglutaryl coenzyme A reductase in higher plants[J]. Plant Physiology Communications, 2004, 40(1):104-110.