

刘 潮, 韩利红, 宋培兵, 等. 桑树潜在致敏原鉴定与生物信息学分析[J]. 江苏农业科学, 2018, 46(11): 24–30.  
doi:10.15889/j.issn.1002-1302.2018.11.006

# 桑树潜在致敏原鉴定与生物信息学分析

刘 潮, 韩利红, 宋培兵, 张亚萍, 唐利洲

(曲靖师范学院云南高原生物资源保护与利用研究中心/云南省高校云贵高原动植物多样性及生态适应性进化重点实验室, 云南曲靖 655011)

**摘要:**致敏原是能引发 I 型过敏反应的抗原性物质, 桑树的花粉和果实均能引起人的过敏反应。为了解桑树的潜在致敏原信息, 以报道的致敏原为探针, 搜索桑树蛋白数据库, 获得桑树潜在致敏原信息, 并对其家族分布、结构和功能进行分析。结果显示, 共获得 143 条桑树的潜在致敏原蛋白, 它们分属于 71 个不同的基因本体 (gene ontology, 简称 GO) 注释子条目和 49 个 pfam 家族, 且其结构和功能各不相同。其中多个成员属于常见的致敏家族, 如 Bet v 1 家族、Profilin 家族、Thaumatococcus 家族、RlpA-like\_DPBB 家族, 与已鉴定的致敏原具有较高的序列一致性和相似的三级结构。研究结果有助于过敏症的临床诊断和交叉反应的预警。

**关键词:**致敏原; 桑树; 过敏反应; 生物信息学

**中图分类号:** S888.2 **文献标志码:** A **文章编号:** 1002-1302(2018)11-0024-07

过敏属于世界性的健康问题, 植物来源的过敏原能引起人类的 I 型过敏反应, 人们既可以从上呼吸道吸入, 也可以通过存在于种类繁多的植物食品中的过敏原引起食物过敏反应。近年来, 食物过敏的患病率有所上升, 在西方国家中, 多达 8% 的幼儿和 2% 的成人受到影响<sup>[1]</sup>。桑树是一种常见的落叶树种, 其叶片多汁, 是桑蚕的主要饲料, 曾有报道指出, 桑树花粉能引起呼吸过敏反应和接触性荨麻疹<sup>[2]</sup>。桑树的果实桑葚能直接食用或用作酿酒原料, 也是常用的中草药成分, 过敏者食用桑葚能引发全身性荨麻疹, 导致呼吸困难、胸闷和腹绞痛等<sup>[3]</sup>。

致敏原属于多个蛋白家族, 结构复杂、功能多样。编码序列和致敏原的三维 (three dimensional, 简称 3D) 结构对引起致敏性至关重要。大多数过敏原都比较小、稳定、结构良好, 这些蛋白质家族多为种子贮藏蛋白、结构蛋白和防御相关的蛋白质——病程相关蛋白。序列或结构相似的蛋白容易引发免疫球蛋白 E (IgE) 交叉反应, 可以通过电脑模拟的方法预测其致敏性。计算工具能相当精确和快速地预测变应原分子的表位位置, 并将其用于确定分子结构, 从而确定 B 细胞表位<sup>[4]</sup>。在本研究中, 筛选了 63 个桑树潜在致敏原, 分为 13 个家族。这些潜在的抗原表位的 IgE 结合残基的功能研究将为后续研究提供借鉴。

## 1 材料与方法

### 1.1 致敏原蛋白序列的获取

收稿日期: 2017-09-19

基金项目: 云南省地方本科高校基础研究联合专项 (编号: 2017FH001-034); 地方高校国家级大学生创新创业训练计划 (编号: 201610684004, 201710684012); 云南省高校科技创新团队 (编号: 云教科[2014]14 号)。

作者简介: 刘 潮 (1980—), 男, 河北景县人, 博士, 讲师, 主要从事植物分子生物学研究。E-mail: liuchao@mail.qjnu.edu.cn。

通信作者: 唐利洲, 博士, 教授, 硕士生导师, 主要从事分子谱系地理学研究。E-mail: tanglizhou@163.com。

2016 年 10 月, 笔者通过食物过敏研究和资源计划 (FARRP) 的致敏原蛋白质数据库 ([www.allergenonline.org](http://www.allergenonline.org)) 和致敏原命名委员会官方网站 (<http://www.allergen.org/index.php>), 下载了截至 2016 年 8 月 4 日公布的 705 条植物和真菌致敏原蛋白序列数据。

### 1.2 桑树潜在致敏原的筛选

以致敏蛋白全序列为探针, 通过 Blast 桑树 (*Morus notabilis*) 蛋白数据库, 根据序列间的相似性筛选潜在的桑树致敏蛋白。桑树潜在致敏蛋白选择阈值为序列一致性 (identity)  $\geq 60\%$ <sup>[5]</sup> 和 E 值  $\leq 10^{-20}$  ([www.allergenonline.org](http://www.allergenonline.org))。

### 1.3 桑树潜在致敏原的结构和功能分析

利用 UniProt 数据库 (<http://www.uniprot.org/>) 对筛选的潜在致敏原蛋白进行结构域和基因本体 (gene ontology, 简称 GO) 功能分类分析。利用 WEGO 在线软件 (Web Gene Ontology Annotation Plotting, <http://wego.genomics.org.cn/>) 对蛋白 GO 富集度进行计算。

### 1.4 桑树潜在致敏原的三级结构预测

蛋白同源三级结构模型用 Swiss-Model<sup>[6]</sup> 构建完成。3D 结构的可视化使用 VMD 1.8.5 (<http://www.ks.uiuc.edu/Research/vmd/vmd-1.8.5/>) 完成。

## 2 结果与分析

### 2.1 桑树潜在致敏原的筛选

以从 FARRP 和 Allergome 数据库上下载的致敏原蛋白为探针, BLAST 桑树蛋白数据库, 以序列一致性 (identity)  $\geq 60\%$  和 E 值  $\leq 10^{-20}$  为筛选条件, 共获得 143 条桑树潜在致敏原 (其中 4 条注释为 Allergen 的序列低于筛选条件), 详见表 1。

### 2.2 桑树潜在致敏原的功能分类

GO 注释将基因产物分为细胞组分、分子功能和生物过程 3 类, 可以对基因及其蛋白的功能进行分类以了解其功能。在 143 条桑树潜在致敏原中, 有 113 条具有 GO 功能注释, 分析发现 71 条不同的 GO 注释子条目, 7 类 GO 子分类基因作

为细胞结构组分起作用,其中参与细胞和细胞部分组成的均占 46.4%,参与细胞器结构成分的占 32.1%。在分子功能分类中,具有结合功能的蛋白占 67.9%,具有催化活性的蛋白占 25.0%,具有结构分子活性的蛋白占 14.3%。13 类 GO 子

分类基因参与了生物学过程,其中参与细胞过程的占 39.3%,参与代谢过程的占 28.6%,参与应激反应的占 23.2%(图 1)。

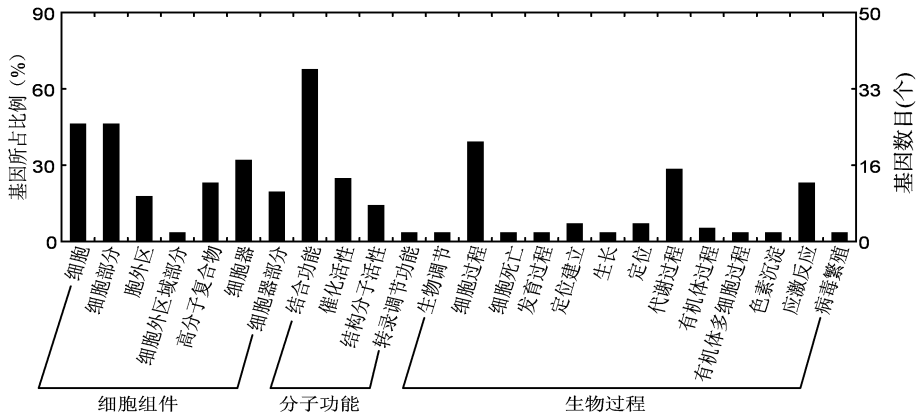


图1 桑树潜在致敏原 GO 功能分类

### 2.3 桑树潜在致敏原的家族分布

在 143 条桑树潜在致敏原中共查询到 172 个 pfam 保守功能域,分属于 49 个 pfam 家族(图 2),根据其所属 pfam 家族共分为 22 类(表 1)。其中热休克蛋白家族含有 HSP70 (PF00012)、HSP20 (PF00011) 和 HSP90 (PF00183) 3 类结构域,共有 25 条序列,该类蛋白是蛋白质折叠复合体的重要成分,可保护植物免受胁迫的损伤<sup>[7]</sup>,关于该家族成员致敏原活性的报道较少。含有 Profilin (PF00235) 功能域的有 3 条序列,且与已知致敏原序列一致性较高,该类蛋白在调节肌动蛋白聚合中起作用,在多种植物的花粉、果实或种子中具有致敏

原活性。肽基脯氨酰顺反异构酶 (Pro\_isomerase, 编号 PF00160),也称为亲环素 (cyclophilin),在哺乳动物中是免疫抑制剂环孢素 A (CsA) 的主要靶细胞组分,也具有分子伴侣和细胞信号转导的功能,桑树潜在致敏原中有 12 个成员含有 Pro\_isomerase 功能域。类甜蛋白 (thauMATIN-like protein, 简称 TLP) 是一类能响应胁迫的蛋白家族,多种植物水果或花粉的 TLPs 具有致敏原活性<sup>[8]</sup>,桑树有 4 个潜在的 TLP 致敏原成员。捕光色素 A-B 结合蛋白在植物光合作用中发挥重要作用,在芹菜中的 Api g 3 具有致敏原活性<sup>[9]</sup>,桑树潜在致敏原中有 5 个成员具有 Chloroa\_b-bind (PF00504) 功能域。

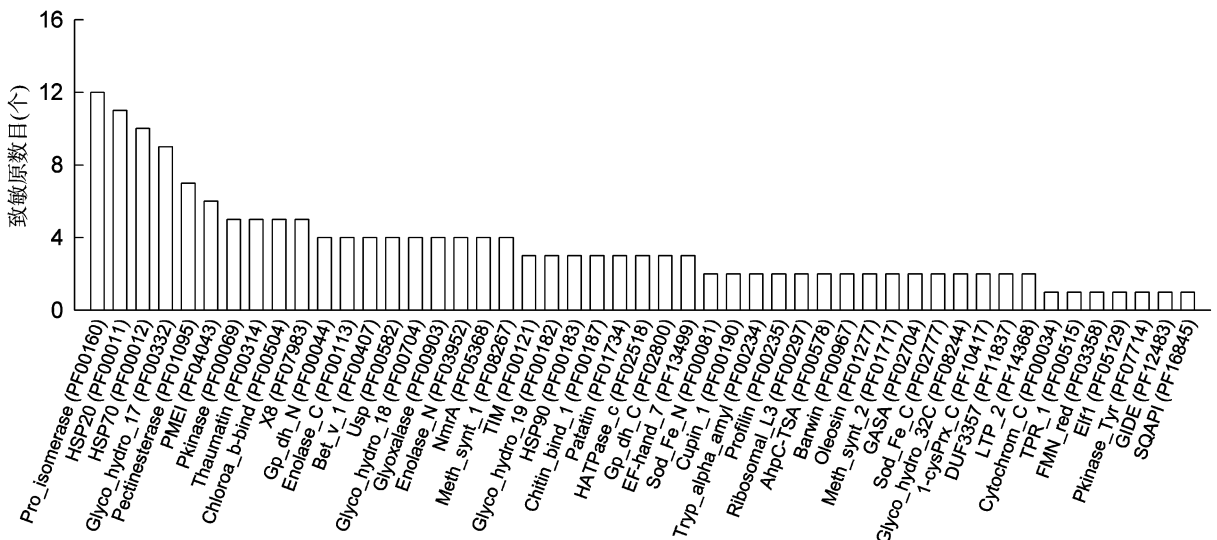


图2 桑树潜在致敏原保守结构域分布

有部分序列同时含有多个功能域,如 I484\_024225 同时含有 HSP90 (PF00183) 和 HATPase\_c (PF02518) 2 个功能域, I484\_020224 同时含有 Glyco\_hydro\_18 (PF00704) 和 Thaumatin (PF00314) 2 个功能域, I484\_006986 则同时含有 Glyco\_hydro\_17 (PF00332)、Pkinase (PF00069) 和 Profilin

(PF00235) 3 个功能域。

### 2.4 桑树潜在致敏原的结构和功能分析

蛋白初级结构是高级结构的基础,具有序列相似性的蛋白,其三维结构可能有明显不同,蛋白序列或结构,尤其是高级结构在其发挥功能过程中起决定作用,蛋白三维结构表层

表 1 桑树潜在致敏原蛋白信息

基因编号	<i>E</i> 值	相似度 (%)	致敏原	物种	蛋白注释
Bet_v_1 (PF00407):病程相关家族 10/ Bet v 1 家族					
L484_001697	$7.00 \times 10^{-20}$	62	Que a 1	白橡	Bet v 1 家族
L484_001696	$7.00 \times 10^{-56}$	49	Pru av 1	甜樱桃	核糖核酸酶
L484_007480	$9.00 \times 10^{-54}$	49	Pyr c 1	西洋梨	核糖核酸酶
L484_001939	$4.00 \times 10^{-5}$	31	Pyr c 1	西洋梨	核糖核酸酶
Tryp_alpha_amyl (PF00234)/ LTP_2 (PF14368):植物脂质转运蛋白					
L484_006084	$7.00 \times 10^{-35}$	60	Mor n 3	黑桑	非特异性脂质转移蛋白
L484_002202	$2.00 \times 10^{-49}$	86	Mor n 3	黑桑	非特异性脂质转移蛋白
L484_013260	$7.00 \times 10^{-24}$	62	Api g 6	白芹	非特异性脂质转移蛋白
L484_027077	$3.00 \times 10^{-34}$	74	Sola 16	番茄	非特异性脂质转移蛋白
Cupin_1 (PF00190):Cupin 超家族					
L484_009809	0	62	Car i 4	薄壳山核桃	豆球蛋白 A
L484_009810	0	63	Car i 4	薄壳山核桃	豆球蛋白 A
Profilin (PF00235):抑制蛋白家族					
L484_005874	$1.00 \times 10^{-84}$	85	Mal d 4	苹果	抑制蛋白
L484_026290	$7.00 \times 10^{-84}$	88	Hev b 8	橡胶树	抑制蛋白
L484_006986	$8.00 \times 10^{-81}$	90	Pru p 4	巴旦木	抑制蛋白
Thaumatococin (PF00314):类甜蛋白					
L484_022589	$3.00 \times 10^{-117}$	66	Pru p 2	巴旦木	类甜蛋白
L484_022590	$4.00 \times 10^{-111}$	65	Pru p 2	巴旦木	类甜蛋白
L484_022591	$2.00 \times 10^{-79}$	62	Pru p 2	巴旦木	类甜蛋白
L484_022592	$2.00 \times 10^{-112}$	67	Pru p 2	巴旦木	类甜蛋白
DPBB_1 (PF03330):含 DPBB 的稀有脂蛋白 A,扩展蛋白					
L484_007509	$6.00 \times 10^{-105}$	79	Act d 5	美味猕猴桃	Kiwellin
L484_007512	$3.00 \times 10^{-101}$	76	Act d 5	美味猕猴桃	Kiwellin
L484_007513	$1.00 \times 10^{-74}$	75	Act c 5	中华猕猴桃	Kiwellin
L484_007514	$7.00 \times 10^{-99}$	75	Act d 5	美味猕猴桃	Kiwellin
L484_007511	$2.00 \times 10^{-90}$	73	Act c 5	中华猕猴桃	Kiwellin
Chitin_bind_1 (PF00187/ PF00967):几丁质识别蛋白,类橡胶蛋白					
L484_014360	$5.00 \times 10^{-158}$	70	Mus a 2	小果野蕉	内切几丁质酶 1
L484_014362	$3.00 \times 10^{-150}$	70	Mus a 2	小果野蕉	内切几丁质酶 1
L484_024568	$1.00 \times 10^{-67}$	77	Hev b 6	橡胶树	橡胶蛋白前体
L484_024570	$3.00 \times 10^{-92}$	70	Hev b 6	橡胶树	橡胶蛋白前体
烯醇化酶 (PF00113/ PF03952):烯醇化酶					
L484_022084	0	60	Hev b 9	橡胶树	烯醇化酶
L484_004312	0	67	Hev b 9	橡胶树	烯醇化酶
L484_001983	0	85	Hev b 9	橡胶树	烯醇化酶
L484_003561	0	91	Hev b 9	橡胶树	烯醇化酶
HSP70 (PF00012):热激蛋白 P70					
L484_013955	0	73	Pen c 19	橘青霉	热激蛋白 P70
L484_018081	0	76	Pen c 19	橘青霉	热激蛋白 P70
L484_026318	0	76	Pen c 19	橘青霉	热激蛋白 P70
L484_026327	0	76	Pen c 19	橘青霉	热激蛋白 P70
L484_018080	0	75	Pen c 19	橘青霉	热激蛋白 P70
L484_025201	0	76	Pen c 19	橘青霉	热激蛋白 P70
L484_002848	$1.00 \times 10^{-119}$	81	Cor a 10	欧洲榛	鲁米诺结合蛋白
L484_000362	$3.00 \times 10^{-131}$	77	Cor a 10	欧洲榛	鲁米诺结合蛋白
L484_007819	0	92	Cor a 10	欧洲榛	鲁米诺结合蛋白
L484_002884	0	79	Cor a 10	欧洲榛	鲁米诺结合蛋白
HSP20 (PF00011):热激蛋白 P20					
L484_000338	$7.00 \times 10^{-73}$	82	Cas s 9	欧洲栗	胞质 I 型小热激蛋白
L484_000756	$2.00 \times 10^{-77}$	75	Cas s 9	欧洲栗	胞质 I 型小热激蛋白
L484_005861	$3.00 \times 10^{-77}$	77	Cas s 9	欧洲栗	胞质 I 型小热激蛋白

续表 1

基因编号	<i>E</i> 值	相似度 (%)	致敏原	物种	蛋白注释
L484_006023	$1.00 \times 10^{-64}$	70	Cas s 9	欧洲栗	胞质 I 型小热激蛋白
L484_006026	$3.00 \times 10^{-67}$	73	Cas s 9	欧洲栗	胞质 I 型小热激蛋白
L484_006031	$6.00 \times 10^{-63}$	69	Cas s 9	欧洲栗	胞质 I 型小热激蛋白
L484_006032	$4.00 \times 10^{-58}$	68	Cas s 9	欧洲栗	胞质 I 型小热激蛋白
L484_006033	$2.00 \times 10^{-61}$	62	Cas s 9	欧洲栗	胞质 I 型小热激蛋白
L484_006034	$5.00 \times 10^{-73}$	68	Cas s 9	欧洲栗	胞质 I 型小热激蛋白
L484_025213	$7.00 \times 10^{-73}$	82	Cas s 9	欧洲栗	胞质 I 型小热激蛋白
L484_025214	$7.00 \times 10^{-81}$	80	Cas s 9	欧洲栗	胞质 I 型小热激蛋白
L484_025215	$1.00 \times 10^{-73}$	76	Cas s 9	欧洲栗	胞质 I 型小热激蛋白
乙二醛酶 (PF00903):乙二醛酶,博来霉素抗性蛋白,双加氧酶超家族					
L484_018049	$1.00 \times 10^{-163}$	77	Ory s GLO I	水稻	乙二醛酶
L484_007241	$2.00 \times 10^{-152}$	71	Ory s GLO I	水稻	乙二醛酶
L484_003661	$4.00 \times 10^{-161}$	76	Ory s GLO I	水稻	乙二醛酶
L484_000361	$1.00 \times 10^{-45}$	73	Ory s GLO I	水稻	乙二醛酶
Chloroa_b - bind (PF00504):叶绿素 a/b 结合蛋白					
L484_018739	$7.00 \times 10^{-118}$	67	Api g 3	白芹	叶绿素 a/b 结合蛋白
L484_020016	$2.00 \times 10^{-150}$	77	Api g 3	白芹	叶绿素 a/b 结合蛋白
L484_013866	$4.00 \times 10^{-162}$	84	Api g 3	白芹	叶绿素 a/b 结合蛋白
L484_002855	$7.00 \times 10^{-174}$	88	Api g 3	白芹	叶绿素 a/b 结合蛋白
L484_002856	$8.00 \times 10^{-173}$	87	Api g 3	白芹	叶绿素 a/b 结合蛋白
HATPase_c (PF02518)/HSP90 (PF00183):组氨酸激酶,DNA 螺旋酶 B 和类热激蛋白 P90 ATP 酶					
L484_020213	0	65	Asp f 12	烟曲霉	热激蛋白 P90
L484_024225	0	67	Asp f 12	烟曲霉	热激蛋白 P90
L484_024226	0	67	Asp f 12	烟曲霉	热激蛋白 P90
EF - hand_7 (PF13499):EF 手型结构域(EF - hand domain pair)					
L484_027230	$1.00 \times 10^{-42}$	77	Ole e 3	油橄榄	钙结合花粉致敏原(Polcalcin)
L484_006237	$3.00 \times 10^{-30}$	70	Phl p 7	梯牧草	Polcalcin
L484_027232	$2.00 \times 10^{-20}$	63	Phl p 7	梯牧草	Polcalcin
Ribosomal_L12 (IPR027534)/Ribosomal_L3 (PF00297):核糖体蛋白 L12 和 L3					
L484_003256	$8.00 \times 10^{-27}$	60	Alt a 12	链格孢	60S 酸性核糖体蛋白
L484_013777	$8.00 \times 10^{-33}$	75	Pru du 5	扁桃	60S 酸性核糖体蛋白
L484_009000	$2.00 \times 10^{-40}$	77	Pru du 5	扁桃	60S 酸性核糖体蛋白
L484_001372	0	68	Asp f 23	烟曲霉	60S 核糖体蛋白 L3
L484_004471	0	67	Asp f 23	烟曲霉	60S 核糖体蛋白 L3
Glyco_hydro_17 (PF00332):糖基水解酶家族 17/ 内切 -1,3 - $\beta$ - 葡萄糖苷酶					
L484_000600	$1.00 \times 10^{-33}$	63	Ole e 10	油橄榄	葡聚糖内酯 -1,3 - $\beta$ - 葡萄糖苷酶
L484_025235	$5.00 \times 10^{-34}$	60	Ole e 10	油橄榄	葡聚糖内酯 -1,3 - $\beta$ - 葡萄糖苷酶
L484_011859	$2.00 \times 10^{-32}$	60	Ole e 10	油橄榄	葡聚糖内酯 -1,3 - $\beta$ - 葡萄糖苷酶
L484_003338	0	66	Ole e 9	油橄榄	葡聚糖内酯 -1,3 - $\beta$ - 葡萄糖苷酶
L484_027431	0	65	Ole e 9	油橄榄	葡聚糖内酯 -1,3 - $\beta$ - 葡萄糖苷酶
L484_016565	$6.00 \times 10^{-135}$	63	Hev b 2	橡胶树	葡聚糖内酯 -1,3 - $\beta$ - 葡萄糖苷酶
L484_016574	$2.00 \times 10^{-151}$	60	Hev b 2	橡胶树	葡聚糖内酯 -1,3 - $\beta$ - 葡萄糖苷酶
L484_004210	0	69	Hev b 2	橡胶树	葡聚糖内酯 -1,3 - $\beta$ - 葡萄糖苷酶
Glyco_hydro_18 (PF00704):糖苷水解酶家族 18					
L484_022481	$2.00 \times 10^{-82}$	66	Hev b 14	橡胶树	酸性内切几丁质酶
L484_022482	$1.00 \times 10^{-87}$	70	Hev b 14	橡胶树	酸性内切几丁质酶
L484_020224	$1.00 \times 10^{-84}$	63	Hev b 14	橡胶树	酸性内切几丁质酶
L484_011486	$2.00 \times 10^{-102}$	80	Hev b 14	橡胶树	酸性内切几丁质酶
Pro_isomerase (PF00160):亲环素型肽基脯氨酰顺反异构酶/ CLD					
L484_016061	$1.00 \times 10^{-69}$	74	Sola l 5	番茄	亲环素
L484_024957	$1.00 \times 10^{-111}$	88	Cat r 1	长春花	亲环素
L484_016063	$2.00 \times 10^{-106}$	81	Cat r 1	长春花	亲环素
L484_016064	$3.00 \times 10^{-60}$	81	Cat r 1	长春花	亲环素

续表 1

基因编号	<i>E</i> 值	相似度 (%)	致敏原	物种	蛋白注释
L484_013897	$3.00 \times 10^{-36}$	62	Asp f 27	烟曲霉	亲环素
L484_024532	$5.00 \times 10^{-97}$	76	Sola l 5	番茄	亲环素
L484_008715	$6.00 \times 10^{-27}$	68	Cat r 1	长春花	亲环素
L484_022005	$1.00 \times 10^{-85}$	71	Sola l 5	番茄	亲环素
L484_002381	$2.00 \times 10^{-24}$	61	Asp f 27	烟曲霉	亲环素
L484_017475	$1.00 \times 10^{-72}$	66	Asp f 27	烟曲霉	亲环素
L484_023973	$4.00 \times 10^{-66}$	60	Sola l 5	番茄	亲环素
L484_027574	$2.00 \times 10^{-71}$	62	Cat r 1	长春花	亲环素
Pkinase(PF00069)/ Usp(PF00582)/ Pkinase_Tyr (PF07714):蛋白激酶结构域,络氨酸激酶					
L484_008213	$2.00 \times 10^{-150}$	64	Sal k 2	猪毛菜	蛋白激酶
L484_019863	$7.00 \times 10^{-145}$	61	Sal k 2	猪毛菜	蛋白激酶
L484_013843	$5.00 \times 10^{-139}$	60	Sal k 2	猪毛菜	蛋白激酶
L484_013466	$4.00 \times 10^{-157}$	65	Sal k 2	猪毛菜	蛋白激酶
L484_003001	$7.00 \times 10^{-156}$	64	Sal k 2	猪毛菜	蛋白激酶
TIM(PF00121):磷酸丙糖异构酶					
L484_024951	$6.00 \times 10^{-111}$	61	Tri a 31	普通小麦	磷酸丙糖异构酶
L484_025239	$3.00 \times 10^{-124}$	74	Tri a 31	普通小麦	磷酸丙糖异构酶
L484_010743	$3.00 \times 10^{-149}$	78	Tri a 31	普通小麦	磷酸丙糖异构酶
Pectinesterase(PF01095)/ PME1(PF04043):植物转化酶/ 果胶甲基酯酶抑制剂					
L484_025594	$7.00 \times 10^{-136}$	66	Act d 7	美味猕猴桃	果胶酯酶
L484_010467	$1.00 \times 10^{-149}$	60	Sal k 1	猪毛菜	果胶酯酶
L484_000799	$3.00 \times 10^{-142}$	65	Act d 7	美味猕猴桃	果胶酯酶
L484_011120	$5.00 \times 10^{-131}$	75	Act d 7	美味猕猴桃	果胶酯酶
L484_011115	$5.00 \times 10^{-137}$	62	Act d 7	美味猕猴桃	果胶酯酶
L484_011118	$2.00 \times 10^{-131}$	60	Act d 7	美味猕猴桃	果胶酯酶
L484_011119	$1.00 \times 10^{-135}$	63	Act d 7	美味猕猴桃	果胶酯酶
其他 Pfam 家族					
L484_009041	$6.00 \times 10^{-61}$	51	Alt a 7	链格孢	YCP4 蛋白
L484_025134	$6.00 \times 10^{-122}$	74	Tri a 32	普通小麦	1 - Cys 过氧化还原酶
L484_013958	$1.00 \times 10^{-106}$	67	Cand a 3	白念珠菌	1 - Cys 过氧化还原酶
L484_000052	$8.00 \times 10^{-50}$	95	Sal k 3	猪毛菜	不依赖于维生素 B <sub>12</sub> 的甲硫氨酸合成酶
L484_000034	$1.00 \times 10^{-36}$	91	Sal k 3	猪毛菜	不依赖于维生素 B <sub>12</sub> 的甲硫氨酸合成酶
L484_009365	0	89	Sal k 3	猪毛菜	不依赖于维生素 B <sub>12</sub> 的甲硫氨酸合成酶
L484_003436	0	76	Sal k 3	猪毛菜	不依赖于维生素 B <sub>12</sub> 的甲硫氨酸合成酶
L484_019026	0	65	Sola l 2	番茄	酸性 $\beta$ - 呋喃果糖苷酶
L484_013887	$2.00 \times 10^{-109}$	60	Mus a 2	小果野蕉	I 类几丁质酶
L484_004273	$5.00 \times 10^{-36}$	61	Act d 4	美味猕猴桃	巯基蛋白酶抑制剂
L484_023933	0	73	Tri a 34	普通小麦	甘油醛 - 3 - 磷酸脱氢酶
L484_006727	0	87	Tri a 34	普通小麦	甘油醛 - 3 - 磷酸脱氢酶
L484_006912	0	87	Tri a 34	普通小麦	甘油醛 - 3 - 磷酸脱氢酶
L484_017656	$3.00 \times 10^{-157}$	67	Tri a 34	普通小麦	甘油醛 - 3 - 磷酸脱氢酶
L484_009576	$2.00 \times 10^{-23}$	63	Tri a 41	普通小麦	NFKB 1 线粒体泛素连接酶活化子
L484_017068	$1.00 \times 10^{-76}$	80	Cor a 13	欧洲榛	油质蛋白
L484_020553	$6.00 \times 10^{-65}$	63	Cor a 12	欧洲榛	油质蛋白
L484_003956	$2.00 \times 10^{-144}$	61	Hev b 7	橡胶树	马铃薯块茎特异蛋白
L484_003958	$2.00 \times 10^{-154}$	61	Hev b 7	橡胶树	马铃薯块茎特异蛋白
L484_013239	$2.00 \times 10^{-152}$	61	Hev b 7	橡胶树	马铃薯块茎特异蛋白
L484_025502	$4.00 \times 10^{-44}$	60	Cur l 3	新月弯孢菌	细胞色素 C
L484_010676	$7.00 \times 10^{-125}$	78	Pis v 4	阿月浑子	超氧化物歧化酶
L484_011225	$2.00 \times 10^{-131}$	86	Hev b 10	橡胶树	超氧化物歧化酶
L484_003380	0	64	Sola l 2	番茄	$\beta$ - 呋喃果糖苷酶
L484_008196	$2.00 \times 10^{-155}$	70	Pyr c 5	西洋梨	异黄酮还原酶相关蛋白
L484_009016	0	86	Bet v 6	疣皮桦	苯基香豆素苜蓿基还原酶

续表 1

基因编号	<i>E</i> 值	相似度 (%)	致敏原	物种	蛋白注释
L484_010255	$1.00 \times 10^{-32}$	81	Pru p 7	巴旦木	赤霉素调节蛋白
L484_011426	$2.00 \times 10^{-52}$	83	Tri a 45	普通小麦	延伸因子 I
L484_011500	$4.00 \times 10^{-28}$	73	Pru p 7	巴旦木	赤霉素调节蛋白
L484_013172	$5.00 \times 10^{-56}$	73	Hum j 1	葎草	未知蛋白
L484_018564	$2.00 \times 10^{-180}$	80	Pyr c 5	西洋梨	异黄酮还原酶相关蛋白
L484_018565	$4.00 \times 10^{-180}$	79	Pyr c 5	西洋梨	异黄酮还原酶相关蛋白

的氨基酸残基具有参与致敏活性的可能。I 型致敏原具有多样的生理生化特征、多样的初级结构和三维结构。蛋白质只有正确折叠,才具有生物学活性,为充分了解桑树潜在致敏原的致敏潜力,笔者以部分蛋白质为例,构建了蛋白质三维模型,以便更准确地了解蛋白质的致敏性。

Bet v 1 成员属于 PR10 家族,在植物的花粉、种子和果实中高表达,在多种植物的花粉和果实中具有过敏原活性<sup>[10]</sup>。有研究发现,桑葚中的 1 个 17 ku 蛋白疑似 PR10 家族的 Bet v 1,是主要的过敏原<sup>[3]</sup>。Bet v 1 的二级结构元件构成 1 个大的疏水口袋,具有储存或载体蛋白的功能<sup>[11]</sup>,能够结合黄酮类化合物<sup>[12]</sup>。在本研究中,4 个已被初步鉴定为致敏原的成员属于该家族,具有引起患者致敏反应的潜力。L484\_001696 的 3-D 结构模型具有典型的 Bet v 1 三级结构,进一步说明该家族桑树潜致敏原具有致敏潜力,值得深入研究。

TLP 属于植物病程相关蛋白第 5 家族,能被多种胁迫诱导表达,在多种水果和作物中具有过敏原活性<sup>[8,13-14]</sup>。TLP 蛋白 3D 结构中的功能域 II 表现出非常保守的构象,预测这些氨基酸序列为主要的 IgE 结合表位<sup>[15]</sup>。L484\_022590 蛋白具有保守结构域 THN,其三级结构由 3 个功能域组成: Domain I 为 N 端核心功能域,由 2 个反向平行的 β 片层组成; Domain II 由半胱氨酸二硫键形成的 3 个较短的 α 螺旋构成; Domain III 由 2 个 β 折叠和 1 个大环构成<sup>[15]</sup>。桑树中有 4

个蛋白与致敏蛋白 Pru p 2 具有 60% 以上的序列一致性。预测的 L484\_022590 三维结构(图 3)与典型 TLP 非常相似,其具体致敏活性值得研究。

RlpA-like\_DPBB 家族包含保守的 DPBB (double-psi beta-barrel) 折叠,DPBB 功能域的 N 端常常连接花粉致敏原<sup>[16]</sup>。Kiwellin 是猕猴桃中重要的富含半胱氨酸的 RlpA-like\_DPBB 家族致敏原蛋白,来自绿色猕猴桃的 Act d 5 蛋白大约为 22 ku,由 N 端 kissper 功能域和 C 端 KiTH 功能域组成,其中后者又由具有 2 个扩展环的 6 个 β 片层构成的筒体构成<sup>[17]</sup>。本研究中搜索到 5 个桑树潜在致敏原与猕猴桃 Act d 5 具有高度一致性。L484\_007509 的三维结构具有典型的 kissper 和 KiTH 功能域,具有致敏潜力。

几丁质结合蛋白家族属于碳水化合物结合结构域家族 18(carbohydrate-binding module family 18,简称 CBM18),具有几丁质识别或结合活性,存在于多种植物与真菌中,能结合 N-乙酰葡萄糖胺<sup>[18]</sup>。该家族成熟蛋白 N 端的 Chitin bind 1 (PF00187)功能域能识别或结合几丁质氨基。本研究中发现 4 个该家族桑树潜在致敏原蛋白,与小果野蕉 Mus a 2 或巴西橡胶 Hev b 6 序列一致性均达 70% 以上,并构建了 L484\_024568 的三维结构(图 3)。

Profilin 是一类参与细胞骨架运动和重组的激动蛋白结合蛋白<sup>[19]</sup>,在包括豚草、橡胶、苹果、樱桃、桃在内的多种植物

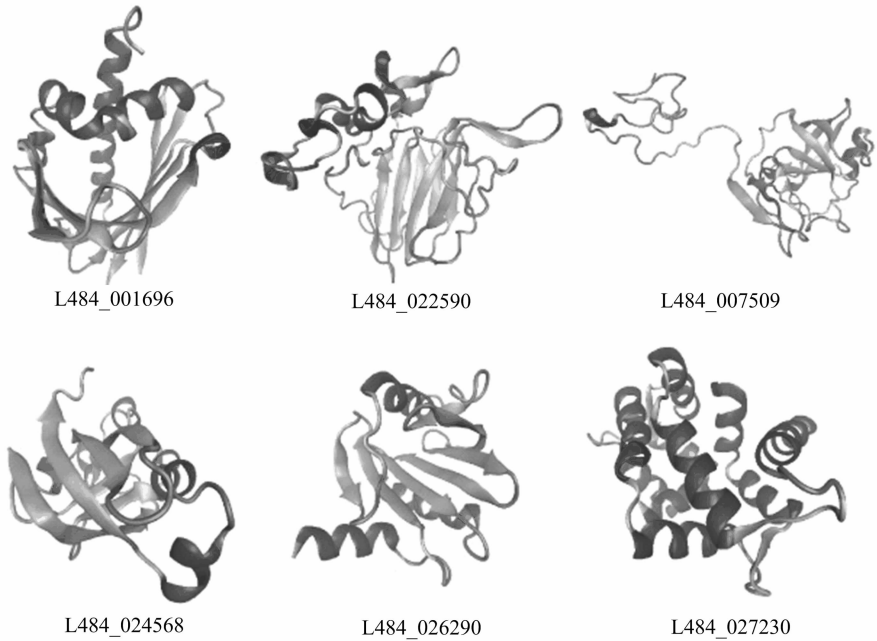


图3 桑树主要潜在致敏原三级结构预测

的花粉、果实或种子中具有致敏原活性。典型的 Profilin 蛋白具有 7 个反向的  $\beta$  片层和 2 个位于两端的  $\alpha$  螺旋(图 3 中的 I484\_026290),在桑树中发现了 3 个该家族成员。

EF-hand 家族属于钙结合蛋白,EF-hand 基序包含螺旋-环-螺旋拓扑结构,能在环上结合钙离子。源自藜、狗牙根、油橄榄和猫尾草等植物花粉中的 EF-hand 家族成员具有致敏原活性。本研究发现,2 个桑树潜在过敏原属于该家族,构建了 I484\_027230 蛋白三维结构(图 3),这些成员具有典型 EF-hand 家族蛋白的类似结构,很可能在桑树的花粉中具有致敏原活性。

### 3 讨论与结论

过敏性疾病是临床上的常见病症,因此,过敏反应性疾病的诊断和治疗的关键是分离和鉴定出引发过敏反应的致敏原。传统鉴定和获得致敏蛋白的方法需要经过表达、纯化、结晶、磁共振或 X 射线衍射等复杂的试验步骤,在实际操作中会遇到诸多问题。目前,随着基因组和蛋白质组数据的增加,以及生物信息学的发展,出现了多种比对和结构建模的软件,这些预测结果大大缩减了科研工作量,使研究工作更有针对性。

植物花粉和果实过敏的发病率都很高,目前在很多植物中已经鉴定出多种植物源致敏原。桑树花粉和果实均能引起人的过敏反应<sup>[2-3]</sup>,而对桑树致敏原的研究较少。全面了解桑树内源性致敏原,包括致敏原的种类和数量、致敏原的时空表达特征等,对于桑树过敏症和交叉反应的诊断具有重要的参考价值。本研究以已知致敏原序列为探针,搜索桑树蛋白质数据库,共获得 143 条潜在致敏原序列。GO 注释显示,它们属于细胞部分,参与细胞过程,说明它们主要为细胞建成分,参与了重要的细胞生物学功能,具有结合活性,参与代谢过程和应激反应,说明这些蛋白能与其他组分结合,可能具有抗原抗体结合活性,是这些蛋白具有致敏性的重要前提。

桑树致敏原蛋白结构多样,本研究共查询到 77 个 pfam 保守结构域,主要包括 HSP70、Profilin、Tubulin、Tubulin\_C、Bet\_v\_1、Chloro\_a\_b-bind 和 RlpA-like\_DPBB 结构域等。其中 Bet\_v\_1 家族、Profilin 家族、Thaumatococcus 家族、kiwellin 家族是较常见的致敏原类型,在多种植物的花粉和果实中具有过敏原活性<sup>[10,13-14,17]</sup>,有些蛋白具有抗真菌活性。本研究发现了多个桑树潜在致敏原成员属于这些家族,其蛋白与已鉴定的致敏原蛋白具有较高的序列一致性,其三级结构与已鉴定的致敏原蛋白结构类似,其致敏潜力值得深入研究。本研究也发现,部分桑树潜在致敏原蛋白属于糖基水解酶、糖苷水解酶和果胶酯酶家族,这些酶类参与了重要的细胞进程,关于其致敏原活性的报道较少,主要在植物的生长发育过程中发挥作用。

### 参考文献:

- [1] Cianferoni A, Spergel J M. Food allergy: review, classification and diagnosis[J]. Allergy International, 2009, 58(4): 457-466.
- [2] Navarro A M, Orta J C, Sanchez M C, et al. Primary sensitization to *Morus alba*[J]. Allergy, 1997, 52(11): 1144-1145.
- [3] Choi J H, Sim J K, Oh J Y, et al. An IgE-mediated allergic reaction caused by mulberry fruit[J]. Allergy Asthma & Immunology

Research, 2015, 7(2): 195-198.

- [4] Nair S, Kukreja N, Singh B P, et al. Identification of B cell epitopes of alcohol dehydrogenase allergen of *Curcularia lunata*[J]. PLoS One, 2011, 6(5): e20020.
- [5] Hileman R E, Silvanovich A, Goodman R E, et al. Bioinformatic methods for allergenicity assessment using a comprehensive allergen database[J]. International Archives of Allergy and Immunology, 2002, 128(4): 280-291.
- [6] Bordoli L, Kiefer F, Arnold K, et al. Protein structure homology modeling using SWISS-MODEL workspace[J]. Nature Protocols, 2008, 4(1): 1-13.
- [7] Morano K A. New tricks for an old dog[J]. Annals of the New York Academy of Sciences, 2007, 1113(1): 1-14.
- [8] Breiteneder H. Thaumatin-like proteins - a new family of pollen and fruit allergens[J]. Allergy, 2004, 59(5): 479-481.
- [9] Faeste C K, Jonscher K R, Sit L, et al. Differentiating cross-reacting allergens in the immunological analysis of celery (*Apium graveolens*) by mass spectrometry[J]. Journal of AOAC International, 2010, 93(2): 451-461.
- [10] Soeria - Atmadja D, Önell A, Kober A, et al. Multivariate statistical analysis of large-scale IgE antibody measurements reveals allergen extract relationships in sensitized individuals[J]. Journal of Allergy and Clinical Immunology, 2007, 120(6): 1433-1440.
- [11] Fernandes H, Michalska K, Sikorski M, et al. Structural and functional aspects of PR-10 proteins[J]. The FEBS Journal, 2013, 280(5): 1169-1199.
- [12] von Loetzen C S, Jacob T, Hartl - Spiegelhauer O, et al. Ligand recognition of the major birch pollen allergen Bet v 1 is isoform dependent[J]. PLoS One, 2015, 10(6): e0128677.
- [13] Esteve C, D'Amato A, Marina M L, et al. In-depth proteomic analysis of banana (*Musa spp.*) fruit with combinatorial peptide ligand libraries[J]. Electrophoresis, 2013, 34(2): 207-214.
- [14] Hegde V L, Ashok K H G, Sreenath K, et al. Identification and characterization of a basic thaumatin-like protein (TLP 2) as an allergen in sapodilla plum (*Manilkara zapota*) [J]. Molecular Nutrition & Food Research, 2014, 58(4): 894-902.
- [15] Leone P, Menu - Bouaouiche L, Peumans W J, et al. Resolution of the structure of the allergenic and antifungal banana fruit thaumatin-like protein at 1.7-Å[J]. Biochimie, 2006, 88(1): 45-52.
- [16] Jorgenson M A, Chen Y, Yahashiri A, et al. The bacterial septal ring protein RlpA is a lytic transglycosylase that contributes to rod shape and daughter cell separation in *Pseudomonas aeruginosa* [J]. Molecular Microbiology, 2014, 93(1): 113-128.
- [17] Offermann L R, Giangrieco I, Perdue M L, et al. Elusive structural, functional, and immunological features of Act d 5, the green kiwifruit kiwellin[J]. Journal of Agricultural and Food Chemistry, 2015, 63(29): 6567-6576.
- [18] Wright H T, Sandrasegaram G, Wright C S. Evolution of a family of N-acetylglucosamine binding proteins containing the disulfide-rich domain of wheat germ agglutinin[J]. Journal of Molecular Evolution, 1991, 33(3): 283-294.
- [19] Valenta R, Duchene M, Ebner C, et al. Profilins constitute a novel family of functional plant pan-allergens[J]. The Journal of Experimental Medicine, 1992, 175(2): 377-385.