

韩长志,刘艳,任雪敏.全基因组预测褐环乳牛肝菌的分泌蛋白[J].江苏农业科学,2017,45(23):45-48.  
doi:10.15889/j.issn.1002-1302.2017.23.013

# 全基因组预测褐环乳牛肝菌的分泌蛋白

韩长志,刘艳,任雪敏

(西南林业大学林学院/云南省森林灾害预警与控制重点实验室,云南昆明 650224)

**摘要:**褐环乳牛肝菌作为针叶树种上重要的外生真菌,为宿主植物提供营养元素和水分,在此过程中,对于该菌分泌蛋白发挥的作用机制尚不清楚。利用 SignalP、ProtComp 等预测程序对褐环乳牛肝菌中 18 303 条蛋白质序列进行分泌蛋白找寻,明确其含有 327 个分泌蛋白,同时,对上述分泌蛋白的氨基酸大小分布、信号肽长度、信号肽切割位点等进行分析,明确该菌分泌蛋白的氨基酸长度集中于 51~450 aa,信号肽长度以 17~21 aa 的序列最为集中,其信号肽切割位点均属于 A-X-A 类型。通过上述生物信息学分析方法有效地实现了褐环乳牛肝菌分泌蛋白的预测,分泌蛋白的信号肽切割位点类型与其他已经报道的致病疫霉、粗糙脉孢霉、禾谷炭疽菌等分泌蛋白信号肽切割位点一致。研究结果为进一步解析分泌蛋白在该菌为宿主植物提供营养过程中发挥的作用机制提供重要的理论支持。

**关键词:**褐环乳牛肝菌;全基因组;分泌蛋白;信号肽;预测程序

**中图分类号:** S182 **文献标志码:** A **文章编号:** 1002-1302(2017)23-0045-04

褐环乳牛肝菌(*Suillus luteus*),别称褐环黏盖牛肝菌、黄黏团子、土色牛肝菌等。目前,对该菌的研究主要集中在发酵生物学<sup>[1]</sup>、发酵条件<sup>[2]</sup>、培养条件<sup>[3]</sup>方面,以及其对樟子松、油松苗的促生作用<sup>[4]</sup>和与木霉协同促进红皮云杉苗木抗逆增效作用<sup>[5]</sup>等方面,此外还有过量钠盐、碱性 pH 值<sup>[6]</sup>和平菇汁<sup>[7]</sup>、接种量<sup>[8]</sup>对其生长影响及 pH 值对其三羧酸循环、草酸分泌及代谢途径中关键酶活性影响<sup>[9-10]</sup>等方面。作为针叶树种的重要外生菌根菌之一,褐环乳牛肝菌在提高根系对磷元素和水分的吸收方面发挥着重要作用<sup>[11]</sup>,然而,在此过程中,该菌中分泌蛋白所发挥的作用机制尚不清楚。

真菌中分泌蛋白一般具有 N-端含有信号肽、无跨膜结构域、无糖基磷脂酰肌醇化蛋白(GPI)锚定位点以及没有将蛋白输送至线粒体或其他胞内细胞器的预测定位信号等 4 个基本特征。前人基于上述特征,对黄孢原毛平革菌<sup>[12]</sup>、结核分枝杆菌<sup>[13]</sup>、酿酒酵母<sup>[14]</sup>、根癌土壤杆菌<sup>[15]</sup>、禾谷镰刀菌<sup>[16]</sup>、粗糙脉孢菌<sup>[17]</sup>、稻瘟菌<sup>[18]</sup>、大丽轮枝菌<sup>[19]</sup>、致病疫霉<sup>[20]</sup>、米曲霉<sup>[21]</sup>、禾谷炭疽菌<sup>[22]</sup>等分泌蛋白进行生物信息学预测。然而,对褐环乳牛肝菌分泌蛋白的研究尚未见报道,随着 *S. luteus* UH-Slu-Lm8-n1 全基因组序列的释放<sup>[23]</sup>,为开展该菌分泌蛋白的预测提供了数据支持。本研究以 *S. luteus* 蛋白数据库中的 18303 条序列为基础,利用 SignalP、ProtComp 等预测程序对分泌蛋白进行找寻,以期进一步解析分泌蛋白为宿主植物提供营养方面的功能研究打下坚实的理论基础。

## 1 材料与方法

### 1.1 褐环乳牛肝菌数据来源

*S. luteus* 的全基因组序列来源于 NCBI (<http://www.ncbi.nlm.nih.gov/genome/?term=Suillus>)。

### 1.2 分泌蛋白确定方法

参照前期对禾谷炭疽菌分泌蛋白的预测方法进行<sup>[22]</sup>。具体而言,利用 SignalP v3.0<sup>[24]</sup>获得 N 端具有信号肽的蛋白序列,并利用 ProtComp v9.0<sup>[25]</sup>实现上述蛋白的定位;再通过 TMHMM v2.0<sup>[26]</sup>对蛋白的跨膜结构域进行预测,并利用 big-PI Fungal predictor 在线分析实现蛋白脂质锚定修饰(glycosylphosphatidylinositol-anchored proteins,简称 GPI)的预测<sup>[27]</sup>;最后,利用 TargetP v1.1<sup>[28]</sup>实现非胞外分泌蛋白的排除,最终明确 *S. luteus* 中的分泌蛋白。

### 1.3 分泌蛋白信号肽酶识别位点预测

利用 LipoP v 1.0<sup>[29]</sup>预测 *S. luteus* 中分泌蛋白的信号肽酶识别位点。

## 2 结果与分析

### 2.1 褐环乳牛肝菌中分泌蛋白的预测

*S. luteus* 中含有 18 303 个蛋白质序列,其中,氨基酸长度最小的蛋白为 49 aa,最大的为 5 054 aa,蛋白多分布在 800 aa 以下,有 16 966 个,比例高达 92.70%,尤以分布在 51~100 aa 的蛋白数量最多,为 4 088 个,比例为 22.34%。而高于 1 000 aa 的蛋白,数量较小,比例为 1.93%(图 1)。

首先,对上述 18 303 个蛋白序列进行信号肽预测,发现 1 083 个蛋白具有信号肽序列;其次,对上述蛋白进行亚细胞定位预测,发现有 378 个蛋白分泌至胞外,705 个蛋白质并不分泌到胞外,其中 299 个蛋白转运至细胞质膜,151 个蛋白传输至线粒体,104 个蛋白转运至内质网,58 个蛋白转运至高尔基体,30 个蛋白转运至液泡中,25 个蛋白转运至溶酶体中,而转运至细胞质、细胞核以及过氧化物酶的蛋白质总共有 38

收稿日期:2016-06-29

基金项目:国家级大学生创新创业训练计划(编号:20160677004);云南省高校优势特色重点学科(生态学)建设项目;真菌学国家重点实验室开放基金;云南省高校林下生物资源保护及利用科技创新团队(编号:2014015)。

作者简介:韩长志(1981—),男,河北石家庄人,博士,副教授,主要从事经济林木病害生物防治与真菌分子生物学的研究。E-mail: hanchangzhi2010@163.com。



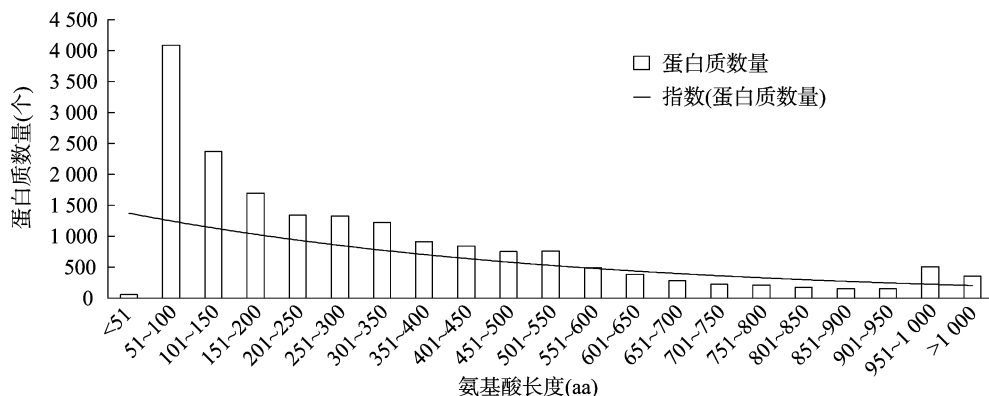


图1 褐环乳牛肝菌蛋白分布情况

个;再次,对上述序列进行跨膜结构域分析,发现有 290 个蛋白不含有跨膜结构域,有 88 个蛋白含有不同数量的跨膜结构域,其中,含有 1 个跨膜结构域的蛋白数量为 77 个,含有 2 个跨膜结构域的蛋白数量为 8 个,含有 3 个跨膜结构域的蛋白数量为 3 个。结合前期分析结果,将 290 个不含跨膜结构域的推测蛋白和 77 个含有 1 个跨膜域的蛋白,合计 367 个蛋白均初步推断为分泌蛋白,继而进行后续 GPI 锚定蛋白的预测;接下来对上述序列进行 GPI 锚定蛋白预测,发现 23 个蛋白是 GPI 锚定蛋白,334 个蛋白序列不具有 GPI 锚定位点,还有 10 个序列小于 55 个氨基酸残基未能进行分析;最后,利用 TargetP 对 344 个非 GPI 锚定的蛋白进行分析,发现 327 个蛋

白含有胞外定位信号,14 个蛋白含有线粒体目标肽,3 个蛋白含有其他定位信号。通过上述分析,最终获得 327 个具有 4 个典型特征的分泌蛋白。

## 2.2 褐环乳牛肝菌分泌蛋白氨基酸长度及信号肽特征

通过对 327 个分泌蛋白进行氨基酸大小分析,发现这些分泌蛋白在 51 ~ 100 aa 的数量最多,比例最大,为 26.30%,其次为 101 ~ 150 aa 的分泌蛋白,比例为 16.51% (图 2)。通过对不同长度的氨基酸进行分析发现,尤以 450 aa 以下的分泌蛋白所占比例较高,为 80.12%。结果表明,分泌蛋白多属于小型蛋白,其所含氨基酸长度一般较小,这与前期研究得出的禾谷炭疽菌<sup>[22]</sup>、樟疫霉<sup>[30]</sup>等病原菌分泌蛋白特征相一致。

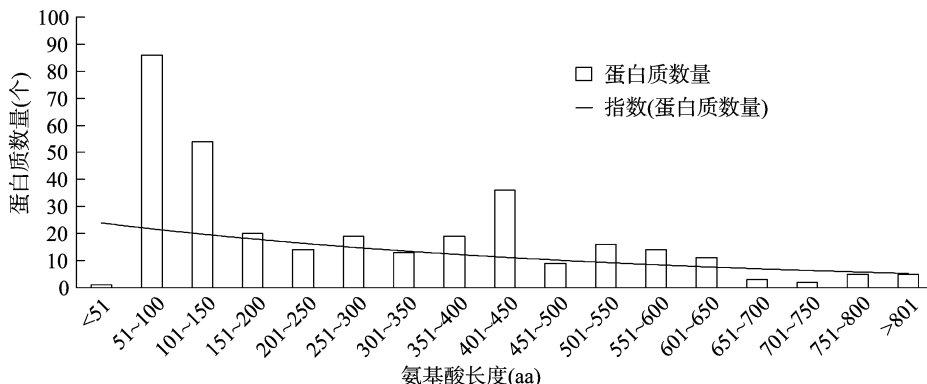


图2 褐环乳牛肝菌分泌蛋白氨基酸长度分布分析结果

对上述分泌蛋白所含信号肽进行氨基酸长度分析,发现含有信号肽长度为 17 ~ 21 aa 的蛋白质数量最多,比例为 63.91%,其中尤以所含信号肽长度为 19 aa 的蛋白序列居多,比例为 16.82% (图 3)。另外,利用 LipoP 对上述分泌蛋

白进行信号肽酶识别位点的预测分析,发现 244 个蛋白序列含有 SP I 型信号肽识别位点,5 个含有 SP II 型信号肽识别位点,比例分别为 74.62%、1.53%,说明 *S. luteus* 中的分泌蛋白大部分是由 SPI 型信号肽酶进行识别,从而切除信号肽的。

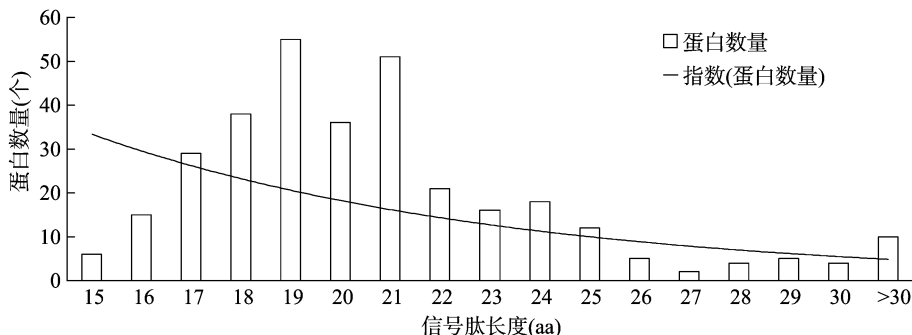


图3 褐环乳牛肝菌分泌蛋白信号肽长度分布分析结果



此外,通过对 *S. luteus* 分泌蛋白信号肽的切割位点 -3 ~ +3 位进行统计分析,发现位于 -3 位、-2 位、-1 位、1 位、2 位、3 位最多的氨基酸分别为 A(丙氨酸)、S(丝氨酸)、A、A、P(脯氨酸)、V(缬氨酸)/T(苏氨酸)(表 1)。因此,位于信号肽切割位点之前的氨基酸组成为 ASA,属于 AXA 类型,为 SP I 型信号肽识别位点,与通过 LipoP 预测的结果相一

表 1 褐环乳牛肝菌分泌蛋白信号肽切割位点 -3 ~ +3 氨基酸分布情况

氨基酸种类	-3		-2		-1		1		2		3	
	数量	比例(%)	数量	比例(%)	数量	比例(%)	数量	比例(%)	数量	比例(%)	数量	比例(%)
A	113	34.56	22	6.73	244	74.62	45	13.76	15	4.59	27	8.26
V	100	30.58	20	6.12	0	0	18	5.50	28	8.56	40	12.23
I	34	10.40	9	2.75	0	0	20	6.12	18	5.50	33	10.09
T	30	9.17	16	4.89	4	1.22	33	10.09	20	6.12	40	12.23
S	23	7.03	77	23.55	25	7.65	28	8.56	42	12.84	29	8.87
C	10	3.06	5	1.53	9	2.75	5	1.53	13	3.98	4	1.22
G	10	3.06	6	1.83	32	9.79	15	4.59	13	3.98	6	1.83
L	4	1.22	41	12.54	4	1.22	23	7.03	7	2.14	34	10.40
R	2	0.61	14	4.28	0	0	12	3.67	8	2.45	5	1.53
N	1	0.31	27	8.26	0	0	5	1.53	6	1.83	11	3.36
P	0	0	1	0.31	5	1.53	0	0	95	29.05	18	5.50
M	0	0	4	1.22	1	0.31	6	1.83	5	1.53	6	1.83
K	0	0	6	1.83	0	0	11	3.36	3	0.92	2	0.61
E	0	0	6	1.83	0	0	13	3.98	9	2.75	8	2.45
D	0	0	8	2.45	0	0	17	5.20	13	3.98	14	4.28
H	0	0	8	2.45	1	0.31	8	2.45	5	1.53	13	3.98
W	0	0	11	3.36	0	0	2	0.61	1	0.31	5	1.53
Q	0	0	12	3.67	1	0.31	34	10.40	8	2.45	5	1.53
Y	0	0	15	4.59	1	0.31	12	3.67	9	2.75	8	2.45
F	0	0	19	5.81	0	0	20	6.12	9	2.75	19	5.81

注:A 表示丙氨酸;V 表示缬氨酸;I 表示异亮氨酸;T 表示苏氨酸;S 表示丝氨酸;C 表示半胱氨酸;G 表示甘氨酸;L 表示亮氨酸;R 表示精氨酸;N 表示天冬酰胺;P 表示脯氨酸;M 表示蛋氨酸;K 表示赖氨酸;E 表示谷氨酸;D 表示天冬氨酸;H 表示组氨酸;W 表示色氨酸;Q 表示谷氨酰胺;Y 表示酪氨酸;F 表示苯丙氨酸。

同时,对组成蛋白质的 20 种氨基酸在信号肽中的分布情况进行统计分析,发现 L(亮氨酸)比例最大,为 17.75%,其次为 A,比例为 16.74%,再次为 S、V、F(苯丙氨酸)、I(异亮氨酸)、M(蛋氨酸)、T、G(甘氨酸)(图 4)。

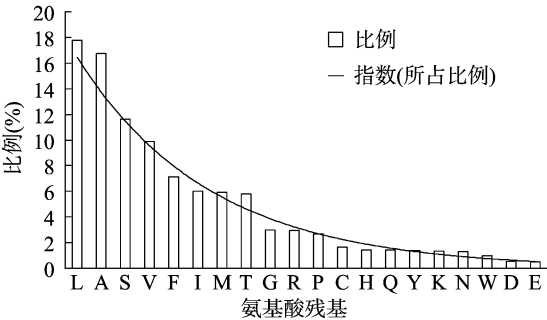


图4 20种氨基酸在分泌蛋白信号肽中比例分布

3 讨论与结论

随着 *S. luteus* 全基因组序列的公布,为深入解析其基因功能的研究提供了重要的数据基础。云南松(*Pinus yunnanensis*)是我国西南地区尤其是云南省境内重要的建群树种,也是该地区的荒山绿化造林先锋树种,其面积、蓄积量

致,同时,也与前人通过对粗糙链孢菌<sup>[17]</sup>、马铃薯晚疫病菌<sup>[20]</sup>、禾谷炭疽菌<sup>[22]</sup>等分泌蛋白具有的信号肽的酶切位点类型报道一致。上述结果表明,无论是植物病原菌,还是植物共生菌,对于分泌蛋白进行信号肽序列识别、切割的信号肽酶类型多为 I 型。

分别占云南省林地面积、蓄积量的 29.2%、15.8%<sup>[31-34]</sup>。作为该树种上重要的外生菌根菌,*S. luteus* 在营养元素利用和水分吸收方面发挥着重要作用。据报道,宿主植物中约 20% 的净初级生产力将传递给外生菌根<sup>[35]</sup>,其生物量在温带森林中约占土壤微生物总生物量的 30%<sup>[36]</sup>。

植物与病原菌互作过程中,病原菌为了实现在植物中定殖、扩展等过程,利用分泌蛋白实现其致病、生长等生理过程。然而,对于外生菌根菌分泌蛋白在营养元素利用和水分吸收方面所具有的功能,尚不明确。本研究发现 *S. luteus* 中存在 327 个具有典型特征的分泌蛋白,与前期获得的禾谷炭疽菌<sup>[22]</sup>、樟疫霉<sup>[30]</sup>的分泌蛋白数量相比,明显较少,是否是由于该菌不是致病菌,致使其所具有的参与植物体内定殖、扩展相关的分泌蛋白退化所造成的,有待于今后进一步得以明确。就 *S. luteus* 分泌蛋白所具有的氨基酸长度而言,多集中于 51 ~ 450 aa 之间,比例为 80.12%;就分泌蛋白所具有的信号肽长度而言,多集中于 15 ~ 28 aa 之间,比例高达 94.19%,上述特征与致病菌分泌蛋白所具有的特征相似。

本研究基于 *S. luteus* 蛋白质数据库中的 18 303 条序列,利用 SignalP、ProtComp、TMHMM、big-PI Fungal Predictor 和 TargetP 等预测程序总共获得 327 个分泌蛋白,明确该数量明显低于致病菌的分泌蛋白数量,同时,通过对上述分泌蛋白进



行氨基酸长度、信号肽长度以及信号肽酶切位点等方面的研究,发现上述分泌蛋白属于小型蛋白,其信号肽长度多集中于 17~21 aa,同时发现真菌信号肽的酶切位点类型具有保守性特点。本研究为深入开展 *S. luteus* 中的分泌蛋白功能研究打下了一定的理论基础。

#### 参考文献:

- [1] 李 敏. 褐环乳牛肝菌 [*Suillus luteus* (L.; Fr.) Gray] 发酵生物学研究[D]. 呼和浩特:内蒙古农业大学,2007.
- [2] 李 敏,闫 伟,堵国成,等. 褐环乳牛肝菌发酵条件的优化[J]. 食品与生物技术学报,2009,28(3):390-396.
- [3] 刘 杨. 褐环黏盖牛肝菌培养条件的研究[D]. 呼和浩特:内蒙古农业大学,2007.
- [4] 钟 帅,闫 伟,张瑞霞. 褐环乳牛肝菌对轻基质营养包樟子松和油松苗的促生效应[J]. 林业科技开发,2014,28(2):122-124.
- [5] 宋小双,邓 勋,尹大川,等. 褐环乳牛肝菌 N94 与绿木霉 T43 复合接种下红皮云杉苗木的生理响应[J]. 吉林农业大学学报,2015,37(1):37-42,46.
- [6] 刘 强,闫 伟. 过量钠盐、碱性 pH 值对褐环乳牛肝菌生长的影响[J]. 内蒙古农业大学学报(自然科学版),2009,30(2):52-56.
- [7] 段 慧. 平菇汁对褐环黏盖牛肝菌生长影响的研究[D]. 呼和浩特:内蒙古农业大学,2010.
- [8] 刘淑清,陈有君,包 健,等. pH 值条件和接种量对褐环乳牛肝菌生长影响[J]. 内蒙古科技与经济,2008(22):29-31,42.
- [9] 孙 琳. pH 值对褐环黏盖牛肝菌三羧酸循环中关键酶活性的影响[D]. 呼和浩特:内蒙古农业大学,2014.
- [10] 王明慧. pH 值对褐环黏盖牛肝菌草酸代谢几个关键酶活性的影响[D]. 呼和浩特:内蒙古农业大学,2013.
- [11] 刘润进,李晓林. 丛枝菌根及其应用[M]. 北京:科学出版社,2000.
- [12] Wymelenberg A V, Sabat G, Martinez D, et al. The phanerochaete chrysosporium secretome: database predictions and initial mass spectrometry peptide identifications in cellulose-grown medium[J]. Journal of Biotechnology, 2005, 118(1):17-34.
- [13] Gomez M, Johnson S, Gennaro M L. Identification of secreted proteins of *Mycobacterium tuberculosis* by a bioinformatic approach[J]. Infection and Immunity, 2000, 68(4):2323-2327.
- [14] 杨 静,李成云,王云月,等. 酿酒酵母分泌蛋白组的计算机分析[J]. 中国农业科学,2005,38(3):516-522.
- [15] 范成明,李成云,赵明富,等. 根癌土壤杆菌 C58 Cereon 中分泌蛋白信号肽分析[J]. 微生物学报,2005,45(4):561-566.
- [16] 于钦亮,马 莉,刘 林,等. 禾谷镰刀菌基因组中含寄主靶向模体分泌蛋白功能的初步分析[J]. 生物技术通报,2008(1):160-165,180.
- [17] 周晓罡,李成云,赵之伟,等. 粗糙脉孢菌基因组分泌蛋白的初步分析[J]. 遗传,2006,28(2):200-207.
- [18] 陈继圣,郑士琴,郑 武,等. 全基因组预测稻瘟菌的分泌蛋白[J]. 中国农业科学,2006,39(12):2474-2482.
- [19] 田 李,陈捷胤,陈相永,等. 大丽轮枝菌 (*Verticillium dahliae* VdLs. 17) 分泌组预测及分析[J]. 中国农业科学,2011,44(15):3142-3153.
- [20] 周晓罡,侯思名,陈铎文,等. 马铃薯晚疫病病菌全基因组分泌蛋白的初步分析[J]. 遗传,2011,33(7):785-793.
- [21] 任 楠,李俊星,沈徐凯. 米曲霉分泌组的预测及分析[J]. 安徽农业科学,2010(25):13622-13625.
- [22] 韩长志. 全基因组预测禾谷炭疽菌的分泌蛋白[J]. 生物技术,2014(2):36-41.
- [23] Grigoriev I V, Nikitin R, Haridas S, et al. MycoCosm portal: gearing up for 1000 fungal genomes[J]. Nucleic Acids Research, 2014, 42: D699-D704.
- [24] Dyrlov Bendtsen J, Nielsen H, von Heijne G, et al. Improved prediction of signal peptides: SignalP 3.0[J]. Journal of Molecular Biology, 2004, 340(4):783-795.
- [25] Emanuelsson O, Nielsen H, Brunak S, et al. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence[J]. Journal of Molecular Biology, 2000, 300(4):1005-1016.
- [26] Krogh A, Larsson B, von Heijne G, et al. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes[J]. Journal of Molecular Biology, 2001, 305(3):567-580.
- [27] Eisenhaber B, Schneider G, Wildpaner M, et al. A sensitive predictor for potential GPI lipid modification sites in fungal protein sequences and its application to genome-wide studies for *Aspergillus nidulans*, *Candida albicans*, *Neurospora crassa*, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*[J]. Journal of Molecular Biology, 2004, 337(2):243-253.
- [28] Emanuelsson O, Brunak S, von Heijne G, et al. Locating proteins in the cell using TargetP, SignalP and related tools[J]. Nature Protocols, 2007, 2(4):953-971.
- [29] Juncker A S, Willenbrock H, Von Heijne G, et al. Prediction of lipoprotein signal peptides in Gram-negative bacteria[J]. Protein Science, 2003, 12(8):1652-1662.
- [30] 韩长志. 全基因组预测樟疫霉的候选效应分子[J]. 南京林业大学学报(自然科学版), 2015(2):69-74.
- [31] 金振洲,彭 鉴. 云南松[M]. 昆明:云南科技出版社,2004.
- [32] 陈 诗,李根前. 一年生云南松不同家系苗木生物量研究[J]. 江苏农业科学,2016,44(6):311-313.
- [33] 程燕芳,王嘉学,许路艳,等. 云南高原喀斯特山原红壤退化中的表层土壤水分变异[J]. 江苏农业科学,2015,43(11):433-437.
- [34] 侯秀丽,吴晓妮,王定康. 滇中不同群落的土壤侵蚀及土壤肥力对比研究[J]. 江苏农业科学,2015,43(12):331-335.
- [35] Kainulainen P, Holopainen J, Palomäki V, et al. Effects of nitrogen fertilization on secondary chemistry and ectomycorrhizal state of scots pine seedlings and on growth of grey pine aphid[J]. Journal of Chemical Ecology, 1996, 22(4):617-636.
- [36] Olsson P A. Signature fatty acids provide tools for determination of the distribution and interactions of mycorrhizal fungi in soil[J]. FEMS Microbiology Ecology, 1999, 29(4):303-310.