

牛樟芝菌丝体全长 cDNA 文库构建及 EST 序列分析

李 晶¹, 陈 莉², 刘朋虎¹, 夏舒宁², 刘艳玲¹, 林占熺^{1*}

1. 福建农林大学 国家菌草工程技术研究中心, 福建 福州 350002

2. 福建农林大学生命科学学院, 福建 福州 350002

摘 要: 目的 为了更好地挖掘牛樟芝 *Antrodia cinnamomea* 菌丝体三萜合成和代谢相关基因、筛选互作蛋白和牛樟芝指纹图谱分析, 构建了牛樟芝菌丝体 cDNA 文库并分析其表达序列标签(EST)序列。方法 以牛樟芝菌丝体为材料, 采用 Gateway 法构建其 cDNA 文库, 对部分 EST 序列进行生物信息学分析、功能注释和 EST-SSR 分析。结果 成功构建了牛樟芝菌丝体 cDNA 文库, 经鉴定文库重组率高达 95%, 文库滴度为 6.1×10^6 cfu/mL, 总克隆数为 1.2×10^7 cfu, 插入片段大小为 300~2 000 bp, 平均长度达 1 000 bp。随机挑选单克隆进行测序获得 65 个有效 EST 序列, 其中 1 个重叠群, 64 个单一序列, 比对结果显示有 45 条序列有明确功能注释, 18 条为未知功能基因, GO 功能注释结果表明序列涉及细胞组成、转运、催化活性、调控等功能。所有 EST 序列共含有 271 个 SSR, 牛樟芝核苷酸重复类型丰富, 其中二核苷酸和三核苷酸重复基元占总重复基元的 94.23%。结论 初步明确牛樟芝菌丝体 cDNA 文库及其 EST 序列相关生物信息, 为牛樟芝菌丝体基因组学研究奠定理论基础。

关键词: 牛樟芝; cDNA 文库; 表达序列标签; GO 功能注释; SSR

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2020)04-1052-08

DOI: 10.7501/j.issn.0253-2670.2020.04.032

Construction of full-length cDNA library from *Antrodia cinnamomea* mycelium and EST sequence analysis

LI Jing¹, CHEN Li², LIU Peng-hu¹, XIA Shu-ning², LIU Yan-ling¹, LIN Zhan-xi¹

1. China National Juncao Engineering Research Center, Fujian Agriculture and Forestry University, Fuzhou 350002, China

2. School of Life Science, Fujian Agriculture and Forestry University, Fuzhou 350002, China

Abstract: Objective To excavate the terpenoid synthesis and metabolism-related gene function and screen the interaction protein and fingerprint analysis of *Antrodia cinnamomea* mycelium, a cDNA library from *A. cinnamomea* mycelia was constructed and the EST sequences were analyzed. **Methods** The cDNA library from the *A. cinnamomea* mycelium was constructed by the Gateway technique. A part of EST sequences about the bioinformatics, functional annotation and EST-SSR were analyzed. **Results** The cDNA library of the *A. cinnamomea* mycelium was constructed successfully. The recombinant rate of the cDNA library was 95%, the titer of the library was 6.1×10^6 cfu/mL, the total cloning number was 1.2×10^7 cfu, the length of cDNA was between 300—2 000 bp with an average length of 1 000 bp. The clones were randomly sequenced and 65 valid ESTs were obtained. After being compared in the Genbank database, 45 ESTs had a definite annotation, and 18 ESTs were unnamed and hypothetical protein. The results with GO functional annotation showed that the ESTs involved the cell composition, transport, catalytic activity, regulation functions and etc. It contained 271 SSRs of all the ESTs in total. The nucleotide repeats in *A. cinnamomea* were abundant, among which dinucleotide and trinucleotide repeat units were more common accounting for 94.23%. **Conclusion** The cDNA library from the *A. cinnamomea* mycelium and its ESTs related biological information were preliminarily identified, which will provide a theoretical foundation for research the mycelium genomics of *A. cinnamomea*.

Key words: *Antrodia cinnamomea* Series; cDNA library; expressed sequence tag (EST); GO functional annotation; SSR

收稿日期: 2019-09-06

基金项目: 福建省教育厅中青年科技项目“牛樟芝深层发酵及其三萜成分差异研究”(JT180121)

作者简介: 李 晶(1985—), 女, 湖北武汉人, 博士, 助理研究员, 从事牛樟芝栽培及生物技术研究。E-mail: 13959197195@163.com

*通信作者 林占熺(1943—), 男, 福建连城人, 研究员, 博士生导师, 从事菌草技术研究及推广工作。E-mail: lzxjuncao@163.com

牛樟芝 *Antrodia cinnamomea* Series 是我国台湾地区特有的珍稀中药资源, 具有重要药用和经济价值。因牛樟芝子实体在自然界生长缓慢, 极大程度阻碍了其对人体健康的研究和应用。探究牛樟芝生长发育及其调控机制是解决牛樟芝产业发展瓶颈的主要方向之一。早在 1994 年就报道了关于 cDNA 文库的研究及应用, 随着分子生物学的发展, 目前常用建库的方法包括 SMART 法^[1-2]、Oligo-capping 法^[3]、Capture 法^[4]、Cap-trapper 法^[5]和 Cap-jumping 法^[6]等, 上述这些方法均以真核生物 mRNA 5'端 CAP 结构进行, 每种方法各有其优缺点。利用 Gateway 技术构建文库, 直接将 cDNA 片段重组至目的载体, 具有高效、快速等特点, 且能够较大程度保留低拷贝基因克隆, 准确反映 mRNA 原始丰度在文库中的表达量^[7]。表达序列标签 (expressed sequenced tag, EST) 是指在构建成功的 cDNA 文库中随机挑选单克隆子, 通过 5'→3'测序得到的 cDNA 序列^[8], 其长度为 150~1 000 bp, 该技术的发展极大程度推进了生物基因组水平的研究, 常被用于新基因分离和鉴定、基因表达谱研究、遗传图谱构建、基因组学研究等, 该方法不仅可以准确、全面的阐述生物的进化水平和联系, 还有助于了解其多样性及系统发育学^[9-11]。高多态性遗传标记的简单重复序列 (simple sequence repeats, SSR) 已广泛用于物种遗传图谱构建, 用于进行物种遗传多样性分析、品种鉴定和基因诊断等方面, 该序列通常由 1~6 个重复单元核苷酸串联组成, 广泛分布于真核生物基因编码区和非编码区^[12]。从文库中直接筛选 EST-SSR 序列操作简单、成本低廉且序列保守, 用于设计引物可大大缩短研发周期、提高结果准确率, 目前该研究已在众多经济作物、保护植物上得到广泛应用, 但在牛樟芝相关研究中尚无报道。

牛樟芝为拟层孔菌科、薄孔菌属, 又称樟芝、牛樟菇等, 于 2015 年 9 月被我国正式公告将其列为中草药进口药典目录, 现代药理研究认为牛樟芝子实体内含有大量多糖^[13-14]、三萜和甾醇等物质^[15], 对保肝^[16]、抗氧化^[17]、治疗肥胖^[18]、癌症^[19-20]和抗炎^[21-22]等各种疾病具有显著疗效。在 NCBI 数据库中检索到较少牛樟芝相关分子信息和 EST 序列, 这表明牛樟芝在基因功能、蛋白功能和代谢途径等研究中还有较大的研究空间。本研究以牛樟芝菌丝体为材料, 构建其 cDNA 文库及分析其 EST 序列, 同时开展其核酸序列、EST 数据库及牛樟芝分子标

记等方面研究, 并为牛樟芝菌丝体药用价值的开发和基因功能的相关研究奠定理论基础。

1 材料

牛樟芝菌株 *Antrodia cinnamomea* T. T. Chang & W. N. Chou AC001 (Genbank 号 KM925002) 由台湾神农真菌生物技术有限公司惠赠, 保存于福建农林大学国家菌草工程技术研究中心。牛樟芝菌种接种至马铃薯葡萄糖琼脂平板 (PDA) 上置于 28 °C 恒温活化后转接至液体培养基中^[23], 28 °C, 120 r/min 黑暗条件下摇床培养, 分别收集培养 14、21、28、35 d 后的菌丝体, 等量均匀混合后用于构建 cDNA 文库。

2 方法

2.1 总 RNA 提取及 mRNA 分离

采用 TRIzol 法^[24]提取牛樟芝菌丝体中总 RNA, 并在 1% 的琼脂糖凝胶上电泳检测完整性, 利用 Nanodrop2000C 进行纯度检测, 对合格的总 RNA 采用 Fast Track MAG Maxi mRNA Isolation Kit 进行 mRNA 分离, 采用 Clone Miner II cDNA 构建试剂盒进行 cDNA 合成。

2.2 文库的构建

采用 Gateway 法酵母双杂进行牛樟芝 cDNA 文库构建, 利用上步分离得到的 mRNA 为模板进行 cDNA 第 1、2 链的合成, 转录得到的 cDNA 连接 3 种不同读框的 adaptor, 通过 BP 重组方法分别克隆至 pDONR222 载体制备成初级 cDNA 文库。然后再提取初级 cDNA 文库混合质粒, 通过 LR 重组到酵母双杂交载体 pGADT7-DEST 上制备成酵母双杂交次级 cDNA 文库。

2.3 文库鉴定、滴度和重组率测定

将所得 cDNA 初级、次级文库均按 1:1 000 体积比稀释后涂布平板, 37 °C 倒置培养 8~12 h 后可见单菌落, 对其进行计数, 按公式计算文库滴度。初级文库检测引物为 pDONR222-F: 5'-TCCCAGT-CACGACGTTGTAAAACGACGGCCAGTCTT-3' 和 pDONR222-R: 5'-AGAGCTGCCAGGAAACAGCTATGACCATGTAATACGACTC-3'

初级文库滴度=单菌落数量×稀释系数×10³/涂板的体积

随机从次级文库中挑选 100 个单菌落斑置于 5 μL 无菌水中作为模板进行 PCR 扩增, 结果进行琼脂糖凝胶电泳检测计算重组率, 次级文库检测引物为 T7: 5'-TAATACGACTCACTATAGGGC-3' 和 3DA: 5'-AGATGGTGCACGATGCACAG-3'。

重组率=电泳阳性转化子数/总挑选单菌落数

2.4 阳性克隆片段测序与序列生物信息学分析

牛樟芝基因组序列暂未见报道。本研究中随机挑选文库阳性克隆转化子送铂尚生物技术有限公司进行测序, 测序结果中通过 Vecscreen tool (www.ncbi.nlm.nih.gov/tools/vecscreen) 去除载体序列; 采用 CAP3 软件去除序列的不准确碱基、poly-A 尾及低质量序列 (<100 bp) 等片段^[25], 最后从而获得较准确测序。

利用 NCBI 的 BLAST 程序对拼接后的单拷贝序列与核酸库 (NT) 和蛋白 (NR) 数据库进行对比 (www.ncbi.nlm.nih.gov/), 完成 BLASTN 及 BLASTX 的同源性比对, 若数据库检索无结果在 EST-others 数据库比对, 寻找相似性高的 EST 片段, 设置 E 值 $<1 \times 10^{-5}$ ^[26]。采用 BLAST2GO 软件 (www.blast2go.com/) 对 EST 序列数据进行分子功能 (molecular function)、细胞组成 (cellular component) 及生物学过程 (biological process) 等方面的功能分类注释^[27]。

2.5 EST-SSR 序列分析

采用 GRAMENE SSRtool 在线工具 (<https://archive.gramene.org/db/markers/ssrtool>) 对牛樟芝菌丝体 EST 进行 SSR 筛选。搜索 SSR 的长度标准为: 二核苷酸、三核苷酸、四核苷酸、五核苷酸、六核苷酸的最少重复次数为 7、5、4、4、3 次, 其重复基元为完全重复、不完全重复, 对 SSR 频率和长度进行统计和分析。

3 结果与分析

3.1 RNA 制备与 mRNA 分离

提取获得的牛樟芝菌丝体总 RNA 经检测其质量浓度为 963 ng/ μ L、总量为 481 μ g, $A_{260}/A_{280} = 2.14$ 。通过 1% 琼脂糖凝胶电泳检测其完整性, 并对总 RNA 分离进一步得到 mRNA, 见图 1。

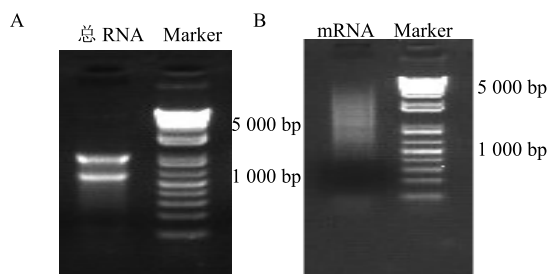
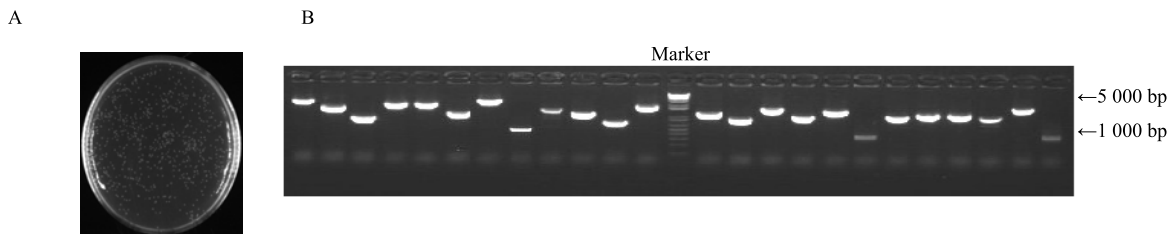


图 1 牛樟芝菌丝体总 RNA (A) 和 mRNA (B) 凝胶电泳图
Fig. 1 Electrophoresis of total RNA (A) and mRNA (B) of *A. cinnamomea* mycelium

3.2 牛樟芝菌丝体 cDNA 文库建立及滴度、重组率鉴定

牛樟芝菌丝体的全长 cDNA 初级文库转化平板稀释度 1:1 000, 文库滴度为 5.2×10^6 cfu/mL, 总克隆数为 1.0×10^7 cfu (图 2); 次级文库重组率高达 95%, 转化平板稀释度 1:1 000, 文库滴度为 6.1×10^6 cfu/mL, 总克隆数为 1.2×10^7 cfu, 插入片段大小为 300~2 000 bp, 平均长度达 1 000 bp (图 3), 说明该文库包含了相对较长的 cDNA, 符合文库库容量要求。



各泳道为随机挑选的转化子, 图 3 同

Randomly selected transformers were in each electrophoresis lane, same as fig. 3

图 2 初级文库容量鉴定 (A) 及片段长度分析 (B)

Fig. 2 Identity (A) and length analysis (B) of primary cDNA library

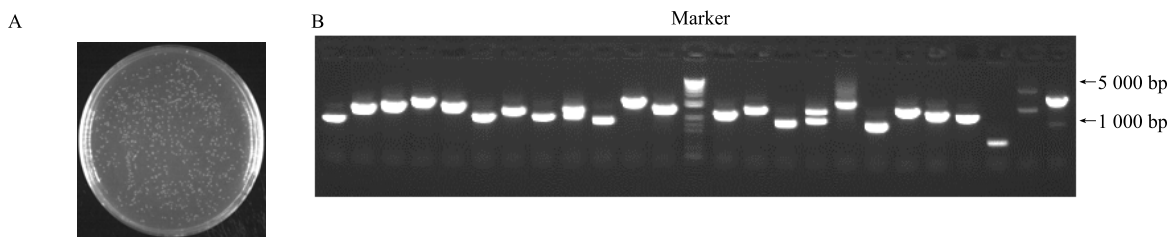


图 3 次级文库容量鉴定 (A) 及片段长度分析 (B)

Fig. 3 Identity (A) and length analysis (B) of secondary cDNA library

3.3 EST 序列注释及生物信息学分析

从文库中随机挑取若干单克隆进行测序, 去掉载体序列和低质量序列后获得 65 条有效单一序列, 并检测其长度 (图 4)。采用 CAP3 程序拼接后获得包括 1 个重叠群和 64 个单一序列, 即非重复的唯一序列, 文库具有良好的代表性。采用 NCBI 的 BLASTX 程序对拼接后的序列与 NR 数据库进行同源性比对和基因功能注释, 结果见表 1。64 条单一序列除了 1 条 (1.56%) 无显著性同源性外; 其余 63 条 (98.44%) 均具有显著同源序列, 包括 45 条已知功能蛋白和 18 条未知蛋白, 其同源蛋白来自牛樟芝、灰树花、茯苓和硫磺多

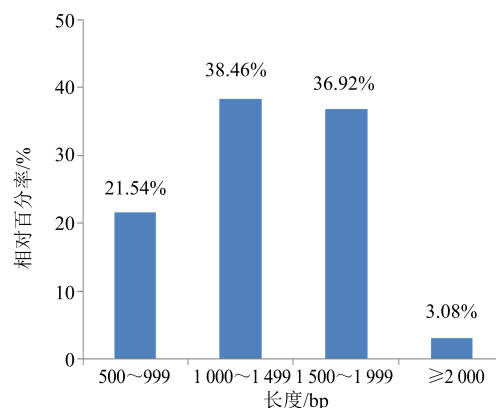


图 4 ESTs 长度分布图

Fig. 4 Distribution of ESTs length

表 1 牛樟芝菌丝体 EST 序列蛋白库比对结果

Table 1 Comparison between EST and protein database of *A. cinnamomea* mycelium

名称	BLASTX 比对结果	最大分值	E 值	相似度/%	EST Genbank 序列
AC2	NAD-dependent formate dehydrogenase	702	0	86.12	MN205326
AC3	mitochondrial carrier (硫磺多孔菌, KZT01 943.1)	484	5×10^{-168}	84.75	MN205327
AC4	1-cys peroxiredoxin (牛樟芝, AAX59 894.1)	462	7×10^{-164}	100.00	MN205328
AC5	NAD(P)-binding protein (茯苓, PCH39 423.1)	253	6×10^{-92}	66.84	MN205329
AC9	phosphatidylserine synthase 2 (茯苓, PCH39 820.1)	781	0	87.95	MN205330
AC12	carboxylic acid transporter (茯苓, PCH39 927.1)	751	0	74.79	MN205331
AC13	predicted protein (褐腐菌, XP_012178 603.1)	103	3×10^{-21}	47.37	MN205332
AC14	alpha-ketoacid dehydrogenase kinase (拟蜡菌, OCH94 093.1)	646	0	77.18	MN205333
AC15	acetyl-coenzyme a synthetase (硫磺多孔菌, KZT08 091.1)	1 205	0	90.94	MN205334
AC17	dipeptidyl-peptidase 5 (绣球菌, XP_027615 982.1)	441	0	85.04	MN205335
AC18	Zn-dependent exopeptidase (茯苓, PCH40 884.1)	590	0	75.79	MN205336
AC21	methionine adenosyltransferase (拟蜡菌, OCH88 795.1)	722	0	93.62	MN205337
AC22	E3 ubiquitin-protein ligase UHRF1 (灰树花, OBZ72 511.1)	201	3×10^{-56}	53.37	MN205338
AC23	sorbitol dehydrogenase (灰树花, OBZ67 832.1)	709	0	91.01	MN205339
AC27	hypothetical protein LAESUDRAFT_672451 (硫磺多孔菌, KZT10 884.1)	372	9×10^{-124}	67.48	MN205340
AC28	hypothetical protein LAESUDRAFT (硫磺多孔菌, KZT03 671.1)	218	3×10^{-68}	68.92	MN205341
AC29	cytochrome P450 monooxygenase (牛樟芝, AEB40 226.1)	1 075	0	94.93	MN205342
AC30	hypothetical protein SCP_0904850 (绣球菌, XP_027617 519.1)	114	2×10^{-26}	43.46	MN205343
AC31	UMP-CMP kinase (拟蜡菌, OCH95 225.1)	329	1×10^{-109}	71.03	MN205344
AC33	vacuolar amino acid transporter 2 (绣球菌, XP_027611 999.1)	356	1×10^{-116}	87.50	MN205345
AC34	U3 small nucleolar RNA-associated protein MPP10 (灰树花, OBZ79 155.1)	564	0	66.34	MN205346
AC35	hypothetical protein OBBRIDRAFT (拟蜡菌, OCH84 396.1)	181	7×10^{-47}	62.22	MN205347
AC36	DUF1909-domain-containing protein (冬生多孔菌, RDX56 099.1)	131	8×10^{-35}	85.14	MN205348
AC37	NAP-domain-containing protein (硫磺多孔菌, KZT00 903.1)	290	2×10^{-110}	80.93	MN205349
AC38	predicated protein (褐腐菌, XP_012185 683.1)	339	3×10^{-112}	64.48	MN205350
AC39	hypothetical protein POSPLADRAFT (绵腐卧孔菌, XP_024336 598.1)	545	0	74.93	MN205351

续表 1

名称	BLASTX 比对结果	最大分值	E 值	相似度	Genbank 序列号
AC40	hypothetical protein HYDPIDRAFT_29335 (松甲藻, KIJ63 539.1)	89	1×10^{-17}	70.31	MN205352
AC41	kynureninase (茯苓, PCH38 103.1)	701	0	77.85	MN205353
AC43	cytochrome P450 (硫磺多孔菌, KZT01 907.1)	624	0	62.18	MN205354
AC46	mitochondrial genome maintenance MGM101 (冬生多孔菌, RDX53 803.1)	282	1×10^{-91}	90.41	MN205355
AC49	NADH dehydrogenase (茯苓, PCH40 004.1)	595	0	78.36	MN205356
AC52	predicated protein (褐腐菌, XP_012181 539.1)	558	0	83.20	MN205357
AC53	indoleamine 2, 3-dioxygenase (茯苓, PCH39 352.1)	597	0	72.12	MN205358
AC55	pyruvate kinase (红栓菌, OSD03 555.1)	938	0	87.55	MN205359
AC57	predicated protein (褐腐菌, XP_012185 683.1)	282	2×10^{-91}	64.66	MN205360
AC58	subtilisin like protease (绣球菌, XP_027611 768.1)	760	0	88.00	MN205361
AC59	tubulin alpha-1a chain (灰树花, OBZ76 459.1)	795	0	98.07	MN205362
AC61	14-3-3-1 protein (栎迷孔菌, KZT71 517.1)	508	3×10^{-180}	98.02	MN205363
AC63	dynein light chain, flagellar outer arm (真姬菇, RDB23 343.1)	193	4×10^{-60}	91.89	MN205364
AC73	ATP synthase F1 gamma (茯苓, PCH36 766.1)	543	0	88.18	MN205365
AC75	kinase-like protein (茯苓, PCH43 934.1)	664	0	81.00	MN205366
AC76	alpha/beta-hydrolase (茯苓, PCH40 344.1)	600	0	74.35	MN205367
AC77	glycoside hydrolase family 5 protein (绵腐卧孔菌, XP_024338 264.1)	323	3×10^{-98}	67.50	MN205368
AC78	cytochrome P450 (绵腐卧孔菌, BAK09 377.1)	590	0	62.20	MN205369
AC79	hypothetical protein MPER_04035 (念珠菌球虫, EEB96 775.1)	74	9×10^{-13}	85.17	MN205370
AC83	predicated protein (褐腐菌, XP_012180 973.1)	607	0	70.13	MN205371
AC84	calreticulin domain containing protein (栎迷孔菌, KZT72 961.1)	794	0	84.07	MN205372
AC86	dicarboxylic acid transporter (硫磺多孔菌, KZT06 709.1)	525	0	89.61	MN205373
AC88	putative monooxygenase (干朽菌, XP_007324 373.1)	598	0	61.07	MN205374
AC90	myo-inositol transporter 1 (绣球菌, XP_027614 396.1)	707	0	82.95	MN205375
AC91	hypothetical protein SCP_1001120 (绣球菌, XP_027617 783.1)	707	0	71.19	MN205376
AC93	actin-related protein (绣球菌, XP_027618 707.1)	777	0	95.13	MN205377
AC95	subtilisin-like protease (绣球菌, XP_027611 768.1)	730	0	74.90	MN205378
AC96	glutamate-ammonia liqase (赭黄齿耳菌, TCD69 661.1)	697	0	94.02	MN205379
AC98	hypothetical protein pospladraft_1071673 (绵腐卧孔菌, XP_024334 193.1)	802	0	85.78	MN205380
AC99	adenosine kinase (绣球菌, XP_027608 215.1)	552	0	83.97	MN205381
AC100	hypothetical protein PHMEG_00017208 (疫病菌, OWZ10 005.1)	78	6×10^{-15}	82.98	MN205382
AC101	oxysterol-binding protein (茯苓, PCH43 009.1)	680	0	84.03	MN205383
AC102	eukaryotic translation initiation factor 5 (硫磺多孔菌, KZT08 930.1)	591	0	87.76	MN205384
AC104	hypothetical protein SCP_1302980 (绣球菌, XP_027619 395.1)	546	0	75.50	MN205385
AC106	alpha/beta-hydrolase (茯苓, PCH37 718.1)	505	6×10^{-172}	68.29	MN205386
AC107	hypothetical protein BN946_scf184 934.g8scf184934 (朱砂菌, CDO75 793.1)	182	1×10^{-55}	67.61	MN205387
AC108	serine hydroxymethyltransferase 3 (灰树花, XP_027620 033.1)	915	0	90.59	MN205388
AC109	未知				MN205389
AC110	predicted protein (褐腐菌, XP_012179 294.1)	669	0	83.67	MN205390

孔菌等物种。这也从一定程度上反映了牛樟芝作为中药新资源的研究还有待于进一步深入。对已有功能注释的单一序列在 BLAST2GO 软件中进行功能学的分类,共获得 GO 术语 283 个,

分为生物学过程 136 个、GO 术语所占比例为 48.06%,细胞组成 60 个、GO 术语所占比例为 21.20%和分子功能 87 个、GO 术语所占比例为 30.74% (图 5)。

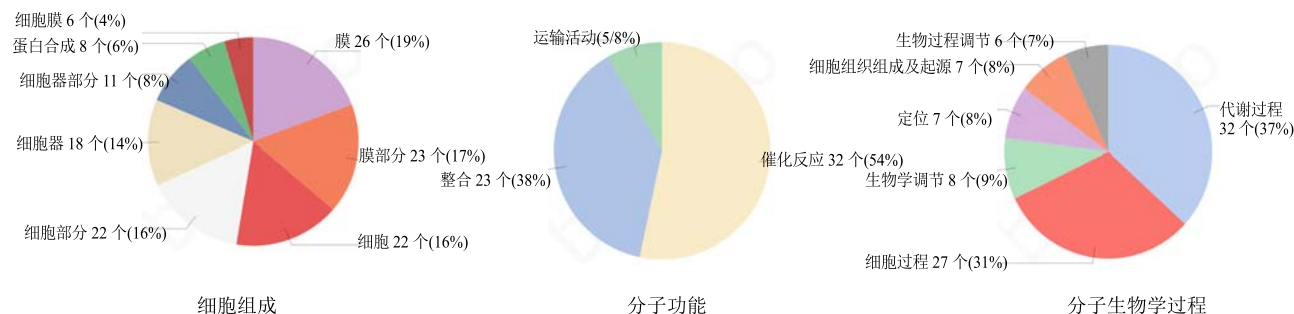


图 5 牛樟芝菌丝体 EST 序列 GO 功能注释

Fig. 5 GO functional annotations for EST sequences of *A. cinnamomea* mycelium

3.4 牛樟芝 EST-SSR 特点分析

对牛樟芝 64 条 EST 序列进行 SSR 检索发现含有 SSR 的 EST 共计 63 条,占全部 EST 总数的 96.92%,共含有 271 个 SSR 序列。在 63 条含有 SSR 的 EST 中只含有 1 个 SSR 的有 7 条、含有 2 个 SSR 的有 12 条、含有 3 个及 3 个以上的共计

44 条。牛樟芝核苷酸重复类型丰富,包括二核苷酸、三核苷酸、四核苷酸、六核苷酸和七核苷酸,从表 2 结果看出二核苷酸所占全部 SSR 比例最大,其次为三核苷酸,四核苷酸、六核苷酸和七核苷酸均只有 1 个,由此可牛樟芝 EST-SSR 中二核苷酸重复占主导地位。

表 2 牛樟芝 EST-SSR 的重复基元

Table 2 Repeated motifs in EST-SSR of *A. cinnamomea* mycelium

重复类型	基元种类	重复基元
二核苷酸	12	AT、AG、AC、TG、TA、TC、CT、CA、CG、GC、GT、GA AAG、ATC、ACA、AGC、TCA、TCC、TTC、TCG、TTG、TTA、TTG、TGG、TGC、
三核苷酸	37	CAT、CAA、CAG、CAC、CTG、CTC、CCT、CGG、CGC、CGA、CGT、GAC、GAG、 GAT、GCT、GGA、GTG、GGC、GTC、GTT、GCC、GGT、GCG、GAA
四核苷酸	1	ATAC
五核苷酸	0	—
六核苷酸	1	CTGGGG
其他	1	CGTGAGG

在牛樟芝 EST 中二核苷酸和三核苷酸重复基元占总重复基元的 94.23%,其中二核苷酸重复基元 GC、CG、CT、TC 出现的次数较频繁,分别占二核苷酸重复序列的 18.77%、16.90%、12.21%和 11.74%;而 AG、GT 和 TA 出现频率较少,其所占比例均不超过 5%。在三核苷酸重复基元中,AAG、GAA 和 TTC 等出现频率略高,均占三核苷酸重复的 5.55%,而 CAC、CGC、AGC、ATC 等均出现 1 次。四核苷酸、六核苷酸和七核苷酸出现频率都偏低,见表 3。

4 讨论

目前进行微生物文库构建的方法繁多,本研究

采用 Gateway 法构建了牛樟芝菌丝体 cDNA 文库,该方法以 mRNA 为模板直接进行 cDNA 双链的合成,没有经过任何扩增的过程,保证了文库的质量和忠实性;同时,高效率的片段和载体重组连接方法,可以大大提高原始文库的库容量,是其他方法所不能达到的。所获得的 cDNA 文库不仅可以提供构建分子标记连锁图谱的探针,还可以用于分离和挖掘全长基因,从而开展基因功能研究。但在文库的构建过程中应注意以下方面。首先,库容量即为该文库所包含的所有的阳性克隆的数量,库容量越大代表所建文库中包含的 cDNA 片段的总数越多,它是衡量文库质量的重要指标,它体现了文库的代

表 3 牛樟芝菌丝体中二核苷酸和三核苷酸部分重复基元
Table 3 Dinucleotide and trinucleotide repeated motifs in ESTs of *A. cinnamomea* mycelium

重复类型	重复基元	数量	所占全部 EST 比例/%
二核苷酸	GC	40	14.76
	CG	36	13.28
	CT	26	9.59
	TC	25	9.22
	CA	18	6.64
	TG	16	5.90
	AT	13	4.80
	GA	11	4.06
	AC	10	3.69
	GT	8	2.95
	AG	6	2.21
	TA	4	1.48
三核苷酸	AGG	3	1.11
	GTC	3	1.11
	GAA	3	1.11
	CAG	2	0.74
	CGT	2	0.74
	GCG	2	0.74
	GAT	2	0.74
	GTG	2	0.74
	TCC	2	0.74
	TCG	2	0.74
	TGC	2	0.74
	TTG	2	0.74

表性,即文库中包含的样本中表达信息(mRNA)的完整性;其次,平均插入片段长度即将随即的菌落 PCR 得到的产物片段长度取平均值,其体现了文库中重组的 cDNA 片段的序列完整性,片段有足够长才能尽可能地反映出基因的天然结构,从而更容易得到文库中目的基因的完整序列和功能信息^[28];最后,重组率反映的是文库中所有克隆的阳性概率,即包含了重组 cDNA 片段的克隆的比率,以上都是在文库建立时所必须考虑的指标^[29]。随着基因组技术的发展,研究蛋白之间相互作用已成为阐述生物过程、研究蛋白功能、细胞信号转导、代谢网络调控等重要研究方法^[30-31],cDNA 文库的建立是研究互作蛋白筛选的重要前提,从 cDNA 文库中筛选与已知蛋白互作的未知蛋白是研究两种或多种蛋白间互作的重要途径,与基因组相比,cDNA 文库中不含内含子,因其更易克隆和大量表达被广泛使用^[32-33]。

随着生物信息学在微生物高通量测序方面的广泛使用,通过生物信息学对当量生物数据进行资源整合、分析和利用,构建牛樟芝基因数据库,能从中筛选出与牛樟芝三萜合成相关的基因及互作

蛋白。本研究将测序得到的牛樟芝 EST 序列提交到 NCBI 数据库进行比对,并利用 BLAST2GO 软件对基因功能注释、GO 分析和 EST-SSR 研究,为今后牛樟芝三萜类化合物合成途径解析、关键基因挖掘、互作蛋白筛选均具有重要意义。

参考文献

- [1] Dai Z, Wang B, Liu Y, *et al.* Producing aglycons of ginsenosides in bakers' yeast [J]. *Sci Rep*, 2014, 4: 3698-3705.
- [2] Wellenreuther R, Schupp I, Poustka A, *et al.* SMART amplification combined with cDNA size fractionation in order to obtain large full-length clones [J]. *BMC Genom*, 2004, 5(1): 36-45.
- [3] Suzuki Y, Sugano S. *Construction Of Full-Length-Enriched Cdna Libraries* [M] New York: Humana Press: Genomics Protocols, 2001.
- [4] Zhumabayeva B, Chenchik A, Siebert P D. RecA-mediated affinity capture: A method for full-length cDNA cloning [J]. *Biotechniques*, 1999, 27(4): 834-845.
- [5] Sugahara Y, Carninci P, Itoh M, *et al.* Comparative evaluation of 5'-end-sequence quality of clones in CAP trapper and other full-length-cDNA libraries [J]. *Gene*, 2001, 263(1/2): 93-102.
- [6] Efimov V, Chakhmakhcheva O, Archdeacon J, *et al.* Detection of the 5'-cap structure of messenger RNAs with the use of the cap-jumping approach [J]. *Nucleic Acids Res*, 2001, 29(22): 4751-4759.
- [7] 王舟,宗俊勤,郭海林,等. Gateway 技术构建结缕草低温和干旱诱导 cDNA 文库及其质量鉴定 [J]. *草业科学*, 2012, 29(6): 911-917.
- [8] 谢德金,沈少炎,姚旺,等. 草珊瑚叶片全长 cDNA 文库构建及 EST 序列分析 [J]. *中草药*, 2016, 47(23): 4235-4241.
- [9] 崔永兰. 石蒜属植物 cDNA 文库构建与表达序列标签(EST)分析 [D]. 南京: 南京林业大学, 2004.
- [10] Hatey F, Tosser-klopp G, Cloucard-martinato C, *et al.* Expressed sequence tags for genes: A review [J]. *Genet Sel Evol*, 1998, 30(6): 521-529.
- [11] Schena M. Genome analysis with gene expression microarrays [J]. *Bioessays*, 1996, 18(5): 427-431.
- [12] Gupta P, Balyan H, Sharma P, *et al.* Microsatellites in plants: A new class of molecular markers [J]. *Current Sci*, 1996, 12: 45-54.
- [13] Lee I H, Huang R L, Chen C T, *et al.* *Antrodia camphorata* polysaccharides exhibit anti-hepatitis B virus effects [J]. *FEMS Microb Let*, 2002, 209(1): 63-67.

- [14] Ker Y B, Peng C C, Chang W L, *et al.* Hepatoprotective bioactivity of the glycoprotein, antrodan, isolated from *Antrodia cinnamomea* mycelia [J]. *PLoS One*, 2014, 9(4): e93191.
- [15] Yu Y L, Chen I H, Shen K Y, *et al.* A triterpenoid methyl antcinat K isolated from *Antrodia cinnamomea* promotes dendritic cell activation and Th2 differentiation [J]. *Eur J Immunol*, 2009, 39(9): 2482-2491.
- [16] Huang C H, Chang Y Y, Liu C W, *et al.* Fruiting body of *Niuchangchih* (*Antrodia camphorata*) protects livers against chronic alcohol consumption damage [J]. *J Agric Food Chem*, 2010, 58(6): 3859-3866.
- [17] Huang G J, Huang S S, Lin S S, *et al.* Analgesic effects and the mechanisms of anti-inflammation of ergostatrien-3 β -ol from *Antrodia camphorata* submerged whole broth in mice [J]. *J Agric Food Chem*, 2010, 58(12): 7445-7452.
- [18] Peng C H, Yang M Y, Yang Y S, *et al.* *Antrodia cinnamomea* prevents obesity, dyslipidemia, and the derived fatty liver via regulating AMPK and SREBP signaling [J]. *Am J Chin Med*, 2017, 45(1): 67-83.
- [19] Peng C C, Lin Y T, Chen K C, *et al.* *Antrodan*, a β -glucan obtained from *Antrodia cinnamomea* mycelia, is beneficial to benign prostate hyperplasia [J]. *Food Funct*, 2015, 6(2): 635-645.
- [20] Wu T R, Huang T T, Martel J, *et al.* Pinicolol B from *Antrodia cinnamomea* induces apoptosis of nasopharyngeal carcinoma cells [J]. *J Ethnopharmacol*, 2017, 201: 117-122.
- [21] Huang T T, Wu S P, Chong K Y, *et al.* The medicinal fungus *Antrodia cinnamomea* suppresses inflammation by inhibiting the NLRP3 inflammasome [J]. *J Ethnopharmacol*, 2014, 155(1): 154-164.
- [22] Chen Y T, Shen Y C, Chang M C, *et al.* Precursor-feeding strategy on the triterpenoid production and anti-inflammatory activity of *Antrodia cinnamomea* [J]. *Process Biochem*, 2016, 51(8): 941-949.
- [23] 李 晶, 林雄杰, 童金华, 等. 牛樟芝 ITS 序列分析及液体发酵菌丝体总三萜含量测定 [J]. 西南农业学报, 2015, 28(6): 2649-2654.
- [24] Lee J T Y, Tsang W H, Chow K L. Simple modifications to standard TRIzol[®] protocol allow high-yield RNA extraction from cells on resorbable materials [J]. *J Biomater Nanobiotechnol*, 2011, 2(1): 41-45.
- [25] Huang X, Madan A. CAP3: A DNA sequence assembly program [J]. *Genom Res*, 1999, 9(9): 868-877.
- [26] Seong E S, Yoo J H, Choi J H, *et al.* Expressed sequence tags analysis and design of simple sequence repeats markers from a full-length cDNA library in *Perilla frutescens* (L.) [J]. *Int J Genom*, 2015, 2015(6): 679548.
- [27] Conesa A, Götz S, García-gómez J M, *et al.* Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research [J]. *Bioinformatics*, 2005, 21(18): 3674-3676.
- [28] Clarke L, Carbon J. A colony bank containing synthetic CoI EI hybrid plasmids representative of the entire *E. coli* genome [J]. *Cell*, 1976, 9(1): 91-99.
- [29] Gao X, Jia R, Wang M, *et al.* Construction and identification of a cDNA library for use in the yeast two-hybrid system from duck embryonic fibroblast cells post-infected with duck enteritis virus [J]. *Mol Biol Rep*, 2014, 41(1): 467-475.
- [30] Soellick T R, Uhrig J F. Development of an optimized interaction-mating protocol for large-scale yeast two-hybrid analyses [J]. *Genome Biol*, 2001, doi: 10.1186/gb-2001-2-12-research0052
- [31] Dittrich M, Birschmann I, Mietner S, *et al.* Platelet protein interactions: Map, signaling components, and phosphorylation groundstate [J]. *Arter Thromb Vasc Biol*, 2008, 28(7): 1326-1331.
- [32] Kohchi T, Fujishige K, Ohyama K. Construction of an equalized cDNA library from *Arabidopsis thaliana* [J]. *Plant J*, 1995, 8(5): 771-776.
- [33] Fields S, Song O K. A novel genetic system to detect protein-protein interactions [J]. *Nature*, 1989, 340(6230): 245-253.