

千里光 β -微管蛋白基因结构与功能的生物信息学分析

张林^{1*}, 徐德林^{1*}, 储士润¹, 吴贵英², 沈访², 钱刚^{1**}

(1. 遵义医学院细胞生物学与遗传学教研室, 贵州遵义 563000;

2. 遵义医学院第一临床学院, 贵州遵义 563000)

摘要: β -微管蛋白是影响细胞新陈代谢和行使功能的重要结构物质, 研究 β -微管蛋白基因的序列信息对揭示其蛋白结构与功能具有重要指导意义。从千里光全长 cDNA 文库中分离得到 β -微管蛋白基因, 并采用生物信息学软件进行序列分析。结果显示, 该基因长度为 1750 bp, 编码的蛋白质长度为 448 个氨基酸, 与柚子 β -微管蛋白的同源性最高, 达 96%; 其蛋白质分子量为 50.01 kD, 理论等电点为 4.83。 β -微管蛋白二级结构主要组成为无规则卷曲结构和 α 螺旋结构; 结构域分析发现该蛋白具有两个保守结构域; 三级结构预测为相对稳固的类球形结构; 信号肽分析将该蛋白主要定位于细胞质、过氧化物酶体、线粒体基质等亚细胞器位置。将该基因序列上传至 GenBank 所获得的登录号为 KF887495。本实验结果为千里光 β -微管蛋白的作用机制揭示和应用研究提供了基础数据, 也为植物 β -微管蛋白基因的分子研究提供了理论依据及基础资料。

关键词: 千里光; β -微管蛋白; 结构与功能; 生物信息学

中图分类号: Q75

文献标识码: A

文章编号: 2095-0837(2014)05-0487-06

Bioinformatic Analysis of Tubulin-beta Gene in *Senecio scandens* Buch. -Ham. ex D. Don

ZHANG Lin^{1*}, XU De-Lin^{1*}, CHU Shi-Run¹, WU Gui-Ying², SHEN Fang², QIAN Gang^{1**}

(1. Department of Cell Biology and Genetics, Zunyi Medical University, Zunyi, Guizhou 563000, China;

2. The First Clinical Institute of Zunyi Medical University, Zunyi, Guizhou 563000, China)

Abstract: Beta-tubulin is an essential component in cell metabolism and function implementation. In this article, the *Senecio scandens* Buch. -Ham. ex D. Don beta-tubulin gene sequence was analyzed for its structure and function, as well as the relationship between them, through bioinformatics. The beta-tubulin nucleotide sequence was obtained from a formerly constructed full-length cDNA library of *S. scandens*, and was used as the query for a series of online bioinformatics tools. The *S. scandens* beta-tubulin gene was 1750 bp long, coding 448 amino acid residues, and shared 96% sequence identity with the beta-tubulin gene in *Citrus maxima*. This protein had a predicted molecular weight of 50.01 kD and theoretical isoelectric point of 4.83. The secondary structure of this protein mainly contained random coils and an alpha helix with two conserved domains. Its predicted tertiary structure was a relatively stable spherical structure. Signal peptide analysis revealed that the most probable sub-cellular locations of this protein were the cytoplasm, peroxisome and mitochondrial matrix space. This gene was uploaded to GenBank with the access number KF887495. This research lays a solid foundation

收稿日期: 2013-12-31, 退修日期: 2014-02-24。

基金项目: 贵州省优秀科技教育人才省长基金项目(黔省专合字 2008-61 号); 遵义医学院博士启动基金(F-567); 贵州省中药现代化科技产业研究开发专项(黔科合 ZY 字[2013]3002 号)。

作者简介: 张林(1988-), 男, 硕士研究生, 研究方向为药用动植物资源开发利用(E-mail: zhanglinzhaixian@163.com); 徐德林, (1981-), 男, 博士, 副教授, 研究方向为药用动植物资源开发利用(E-mail: xudelin2000@163.com)。

* 共同第一作者。

** 通讯作者(Author for correspondence. E-mail: pengjiaqiong@163.com)。

for future studies on function mechanism revelation and application for genetically improving *S. scandens*. This study also provides some basic data for studies on plant beta-tubulin.

Key words: *Senecio scandens*; Tubulin-beta; Structure and function; Bioinformatics analyzing

千里光(*Senecio scandens* Buch.-Ham. ex D. Don)为菊科(Compositae)千里光属多年生草本植物,是我国传统中药材,具有清热解毒、杀虫、明目、保肝等功效,主要用于治疗痈肿疮毒、感冒发热、目赤肿痛、泄泻痢疾和皮肤湿疹等^[1-3]。本实验室的前期研究证实,千里光对迟钝爱德菌、金黄色葡萄球菌、铜绿假单胞菌、大肠埃希菌、甲型副伤寒杆菌、福氏痢疾杆菌均有抗菌作用^[4],具有重要的药用价值。

微管蛋白为细胞骨架的重要结构物质,在动植物所有组织和细胞中普遍存在^[5-7],微管蛋白对细胞的新陈代谢和功能实现具有重要影响。在药物开发研究中,微管蛋白被当作抗肿瘤药物的重要靶点,微管蛋白的组成与状态对抗微管药耐药性起主要作用^[8]。微管蛋白主要由α、β-微管蛋白组成,抗微管药物是通过影响β-微管蛋白的聚合与解聚而发挥抗癌作用的。通过对其它物种中β-微管蛋白基因的研究发现,β-微管蛋白功能具有多样性,主要包括维持细胞形态、参与胞内运输与细胞运动、调节细胞分裂及信号传导等^[9]。目前关于动物β-微管蛋白基因的研究较多,但关于植物β-微管蛋白基因的报道较少,尚未见到千里光β-微管蛋白基因的文献。对千里光β-微管蛋白基因进行核苷酸、氨基酸、蛋白结构、功能位点的信息学分析,对阐明其功能和作用机制、结构与功能的关系具有重要指导意义。此外,基于β-微管蛋白基因序列在动物、植物中具有保守性,也可为千里光抗菌性状的研究和以β-微管蛋白为靶点的抗癌药物的研发提供启示。

本研究从已构建的千里光cDNA文库^[10]中筛选出β-微管蛋白基因的全长序列,利用在线生物信息学分析软件对该基因进行氨基酸序列、理化性质、蛋白二级结构、结构功能域、三级结构、亚细胞定位、功能分类及蛋白质信号肽的分析,推测千里光β-微管蛋白核苷酸和氨基酸序列的结构特点和作用机制,以期为该基因的功能验证、作用机理

阐释和基因的应用研究提供理论指导,也为植物β-微管蛋白基因的分子研究提供基础资料与参考。

1 研究方法

1.1 千里光β-微管蛋白基因的克隆

应用宝生物(大连)工程公司的RNA提取试剂盒提取千里光叶片总RNA,用SMART方法构建叶片组织全长cDNA文库^[10],对测序结果进行NCBI数据库Blast X分析(<http://blast.ncbi.nlm.nih.gov/Blast.cgi>),从注解结果中初步筛选出注解为β-微管蛋白基因的全长序列。

1.2 千里光β-微管蛋白序列的生物信息学分析

1.2.1 氨基酸序列分析

用NCBI的在线分析软件Blast对测序获得的核苷酸序列进行在线比对,ORF finder检测cDNA序列的开放阅读框;利用DNAMAN软件进行多物种同源氨基酸序列的多重比对。

1.2.2 理化性质分析

应用Expasy网站的ProtParam程序(<http://web.expasy.org/protparam/>)对千里光β-微管蛋白进行氨基酸含量与理论分子量及等电点的分析;利用ProtScale程序(<http://web.expasy.org/protscale/>)进行β-微管蛋白的疏水性分析。

1.2.3 结构预测

使用在线软件GOR4(<http://npsa-pbil.ibcp.fr/cgi-bin/npsa-automat.pl>)进行目的序列的二级结构分析;应用SMART服务器(<http://smart.embl-heidelberg.de/>)和GenBank对目的序列进行功能结构域分析;用同源模型服务软件SWISS-MODEL(<http://swissmodel.expasy.org/>)进行三维结构预测^[11,12]。

1.2.4 功能分析

通过在线工具Psort(<http://psort.nibb.ac.jp/form.Html>)进行目的序列的亚细胞定位分析;应用在线工具Profun(<http://www.cbs.dtu.dk/services/ProtFun/>)进行功能分类预测;使用在线工具

SignalP (<http://www.cbs.dtu.dk/servi-ces/SignalP/>)进行蛋白质信号肽分析。

2 结果与分析

2.1 氨基酸序列预测分析

根据 ORF finder 对千里光 β -微管蛋白基因的开放阅读框分析发现,该基因全长 1750 bp,共编

码氨基酸残基 448 个,包含一个完整的开放阅读框。将该基因序列提交至 GenBank,获得注册登录号为 KF887495。应用软件 DNAMAN 进行多物种氨基酸序列比对,发现千里光 β -微管蛋白的氨基酸序列与螺旋狸藻、蒺藜苜蓿、水稻、可可树、陆地棉及乌拉尔图小麦等植物的 β -微管蛋白基因序列同源性很高(图 1),说明 β -微管蛋白基因在植

sc-36	MREILHIQGGCGCNIGAKFWEVICDEHGIDQAGK...GDSIQLERLINVYNEASGGRFPRAVLMDLEPGTMDSRSCPFQGIIFRPNDFVFGQSAGNN	100
XP_002330612.1	MREILHIQGGCGCNIGAKFWEVICDEHGIDQAGK...GDSIQLERLINVYNEASGGRFPRAVLMDLEPGTMDSRSCPFQGIIFRPNDFVFGQSAGNN	100
ACM78033.1	MREILHIQGGCGCNIGAKFWEVICDEHGIDQAGK...GDSIQLERLINVYNEASGGRFPRAVLMDLEPGTMDSRSCPFQGIIFRPNDFVFGQSAGNN	100
A GH08242.1	MREILHIQGGCGCNIGAKFWEVICDEHGIDQAGK...GDSIQLERLINVYNEASGGRFPRAVLMDLEPGTMDSRSCPFQGIIFRPNDFVFGQSAGNN	100
ABS50668.1	MREILHIQGGCGCNIGAKFWEVICDEHGIDQAGK...GDSIQLERLINVYNEASGGRFPRAVLMDLEPGTMDSRSCPFQGIIFRPNDFVFGQSAGNN	100
EPS63676.1	MREILHIQGGCGCNIGAKFWEVICDEHGIDQAGK...GDSIQLERLINVYNEASGGRFPRAVLMDLEPGTMDSRSCPFQGIIFRPNDFVFGQSAGNN	100
XP_003594596.1	MREILHIQGGCGCNIGAKFWEVICDEHGIDQAGK...GDSIQLERLINVYNEASGGRFPRAVLMDLEPGTMDSRSCPFQGIIFRPNDFVFGQSAGNN	100
JC2511	MREILHIQGGCGCNIGAKFWEVICDEHGIDQAGK...GDSIQLERLINVYNEASGGRFPRAVLMDLEPGTMDSRSCPFQGIIFRPNDFVFGQSAGNN	100
EOY25532.1	MREILHIQGGCGCNIGAKFWEVICDEHGIDQAGK...GDSIQLERLINVYNEASGGRFPRAVLMDLEPGTMDSRSCPFQGIIFRPNDFVFGQSAGNN	100
ABY86660.1	MREILHIQGGCGCNIGAKFWEVICDEHGIDQAGK...GDSIQLERLINVYNEASGGRFPRAVLMDLEPGTMDSRSCPFQGIIFRPNDFVFGQSAGNN	100
EMS61990.1	MREILHIQGGCGCNIGAKFWEVICDEHGIDQAGK...GDSIQLERLINVYNEASGGRFPRAVLMDLEPGTMDSRSCPFQGIIFRPNDFVFGQSAGNN	100
Consensus	mreilhiqggcgcnigakfwevicdehgidi g y gds lqler nvy nea ggr vpravlmdlepgtmdsrscpfqgiifrpnfvfgqsagnn	100
sc-36	WAKGHYTEGAELIDSVLDVVRKEAENCDCLGQFQVCHSLGGCTGSGNTILLISKIREEYPDRMLLTFSVFPSPKVSDTVVEPYNATLSVHQLVENADECM	200
XP_002330612.1	WAKGHYTEGAELIDSVLDVVRKEAENCDCLGQFQVCHSLGGCTGSGNTILLISKIREEYPDRMLLTFSVFPSPKVSDTVVEPYNATLSVHQLVENADECM	200
ACM78033.1	WAKGHYTEGAELIDSVLDVVRKEAENCDCLGQFQVCHSLGGCTGSGNTILLISKIREEYPDRMLLTFSVFPSPKVSDTVVEPYNATLSVHQLVENADECM	200
A GH08242.1	WAKGHYTEGAELIDSVLDVVRKEAENCDCLGQFQVCHSLGGCTGSGNTILLISKIREEYPDRMLLTFSVFPSPKVSDTVVEPYNATLSVHQLVENADECM	200
ABS50668.1	WAKGHYTEGAELIDSVLDVVRKEAENCDCLGQFQVCHSLGGCTGSGNTILLISKIREEYPDRMLLTFSVFPSPKVSDTVVEPYNATLSVHQLVENADECM	200
EPS63676.1	WAKGHYTEGAELIDSVLDVVRKEAENCDCLGQFQVCHSLGGCTGSGNTILLISKIREEYPDRMLLTFSVFPSPKVSDTVVEPYNATLSVHQLVENADECM	200
XP_003594596.1	WAKGHYTEGAELIDSVLDVVRKEAENCDCLGQFQVCHSLGGCTGSGNTILLISKIREEYPDRMLLTFSVFPSPKVSDTVVEPYNATLSVHQLVENADECM	200
JC2511	WAKGHYTEGAELIDSVLDVVRKEAENCDCLGQFQVCHSLGGCTGSGNTILLISKIREEYPDRMLLTFSVFPSPKVSDTVVEPYNATLSVHQLVENADECM	200
EOY25532.1	WAKGHYTEGAELIDSVLDVVRKEAENCDCLGQFQVCHSLGGCTGSGNTILLISKIREEYPDRMLLTFSVFPSPKVSDTVVEPYNATLSVHQLVENADECM	200
ABY86660.1	WAKGHYTEGAELIDSVLDVVRKEAENCDCLGQFQVCHSLGGCTGSGNTILLISKIREEYPDRMLLTFSVFPSPKVSDTVVEPYNATLSVHQLVENADECM	200
EMS61990.1	WAKGHYTEGAELIDSVLDVVRKEAENCDCLGQFQVCHSLGGCTGSGNTILLISKIREEYPDRMLLTFSVFPSPKVSDTVVEPYNATLSVHQLVENADECM	200
Consensus	wakghytegaelid vldvvrkeaeencdcglqfqvchslggctgsgntlliskireeypdrmlltfsvfspkvsdtvvepynatlsvhqlvenadecm	200
sc-36	VLDNEALYDICTRLKLATPSFGDNLNHLISATMSGVICCLRLPQGQLNSDLRKLAVNLIPPLRHFMMVGFAPILTSGRSQCYRALIVPELTQQMWDAKNMM	300
XP_002330612.1	VLDNEALYDICTRLKLATPSFGDNLNHLISATMSGVICCLRLPQGQLNSDLRKLAVNLIPPLRHFMMVGFAPILTSGRSQCYRALIVPELTQQMWDAKNMM	300
ACM78033.1	VLDNEALYDICTRLKLATPSFGDNLNHLISATMSGVICCLRLPQGQLNSDLRKLAVNLIPPLRHFMMVGFAPILTSGRSQCYRALIVPELTQQMWDAKNMM	300
A GH08242.1	VLDNEALYDICTRLKLATPSFGDNLNHLISATMSGVICCLRLPQGQLNSDLRKLAVNLIPPLRHFMMVGFAPILTSGRSQCYRALIVPELTQQMWDAKNMM	300
ABS50668.1	VLDNEALYDICTRLKLATPSFGDNLNHLISATMSGVICCLRLPQGQLNSDLRKLAVNLIPPLRHFMMVGFAPILTSGRSQCYRALIVPELTQQMWDAKNMM	300
EPS63676.1	VLDNEALYDICTRLKLATPSFGDNLNHLISATMSGVICCLRLPQGQLNSDLRKLAVNLIPPLRHFMMVGFAPILTSGRSQCYRALIVPELTQQMWDAKNMM	300
XP_003594596.1	VLDNEALYDICTRLKLATPSFGDNLNHLISATMSGVICCLRLPQGQLNSDLRKLAVNLIPPLRHFMMVGFAPILTSGRSQCYRALIVPELTQQMWDAKNMM	300
JC2511	VLDNEALYDICTRLKLATPSFGDNLNHLISATMSGVICCLRLPQGQLNSDLRKLAVNLIPPLRHFMMVGFAPILTSGRSQCYRALIVPELTQQMWDAKNMM	300
EOY25532.1	VLDNEALYDICTRLKLATPSFGDNLNHLISATMSGVICCLRLPQGQLNSDLRKLAVNLIPPLRHFMMVGFAPILTSGRSQCYRALIVPELTQQMWDAKNMM	300
ABY86660.1	VLDNEALYDICTRLKLATPSFGDNLNHLISATMSGVICCLRLPQGQLNSDLRKLAVNLIPPLRHFMMVGFAPILTSGRSQCYRALIVPELTQQMWDAKNMM	300
EMS61990.1	VLDNEALYDICTRLKLATPSFGDNLNHLISATMSGVICCLRLPQGQLNSDLRKLAVNLIPPLRHFMMVGFAPILTSGRSQCYRALIVPELTQQMWDAKNMM	300
Consensus	vldnealyd ictrlkl tp fgdlnhlisatmsgvicclrlfpqqlnsdlrklnlnipfprlhffmgfapltsrgsq yr ltvpeitqqmwda knmm	300
sc-36	CAADPRHGRYLTASAFRGKMSKTEVDEQINVNQNKNSYYFVEWIPNVKSSVCDIPPEGLKMASTIFGNSTIQEMFRRVSEQFTAFRRKAFLHWITG	400
XP_002330612.1	CAADPRHGRYLTASAFRGKMSKTEVDEQINVNQNKNSYYFVEWIPNVKSSVCDIPPEGLKMASTIFGNSTIQEMFRRVSEQFTAFRRKAFLHWITG	400
ACM78033.1	CAADPRHGRYLTASAFRGKMSKTEVDEQINVNQNKNSYYFVEWIPNVKSSVCDIPPEGLKMASTIFGNSTIQEMFRRVSEQFTAFRRKAFLHWITG	400
A GH08242.1	CAADPRHGRYLTASAFRGKMSKTEVDEQINVNQNKNSYYFVEWIPNVKSSVCDIPPEGLKMASTIFGNSTIQEMFRRVSEQFTAFRRKAFLHWITG	400
ABS50668.1	CAADPRHGRYLTASAFRGKMSKTEVDEQINVNQNKNSYYFVEWIPNVKSSVCDIPPEGLKMASTIFGNSTIQEMFRRVSEQFTAFRRKAFLHWITG	400
EPS63676.1	CAADPRHGRYLTASAFRGKMSKTEVDEQINVNQNKNSYYFVEWIPNVKSSVCDIPPEGLKMASTIFGNSTIQEMFRRVSEQFTAFRRKAFLHWITG	400
XP_003594596.1	CAADPRHGRYLTASAFRGKMSKTEVDEQINVNQNKNSYYFVEWIPNVKSSVCDIPPEGLKMASTIFGNSTIQEMFRRVSEQFTAFRRKAFLHWITG	400
JC2511	CAADPRHGRYLTASAFRGKMSKTEVDEQINVNQNKNSYYFVEWIPNVKSSVCDIPPEGLKMASTIFGNSTIQEMFRRVSEQFTAFRRKAFLHWITG	400
EOY25532.1	CAADPRHGRYLTASAFRGKMSKTEVDEQINVNQNKNSYYFVEWIPNVKSSVCDIPPEGLKMASTIFGNSTIQEMFRRVSEQFTAFRRKAFLHWITG	400
ABY86660.1	CAADPRHGRYLTASAFRGKMSKTEVDEQINVNQNKNSYYFVEWIPNVKSSVCDIPPEGLKMASTIFGNSTIQEMFRRVSEQFTAFRRKAFLHWITG	400
EMS61990.1	CAADPRHGRYLTASAFRGKMSKTEVDEQINVNQNKNSYYFVEWIPNVKSSVCDIPPEGLKMASTIFGNSTIQEMFRRVSEQFTAFRRKAFLHWITG	400
Consensus	caadprhgryltasa frgkmsktkevdeqm nvqnknssyfveip nvkssvcd pp glkm tf gnst iqemfrrvseqfta frrkaflhw tg	400
sc-36	EGMDMEFTEAESNMNDLVAEYQQYQDADIEYDDNDIEVAHDGVSL	448
XP_002330612.1	EGMDMEFTEAESNMNDLVAEYQQYQDADIEYDDNDIEVAHDGVSL	445
ACM78033.1	EGMDMEFTEAESNMNDLVAEYQQYQDADIEYDDNDIEVAHDGVSL	445
A GH08242.1	EGMDMEFTEAESNMNDLVAEYQQYQDADIEYDDNDIEVAHDGVSL	445
ABS50668.1	EGMDMEFTEAESNMNDLVAEYQQYQDADIEYDDNDIEVAHDGVSL	445
EPS63676.1	EGMDMEFTEAESNMNDLVAEYQQYQDADIEYDDNDIEVAHDGVSL	445
XP_003594596.1	EGMDMEFTEAESNMNDLVAEYQQYQDADIEYDDNDIEVAHDGVSL	445
JC2511	EGMDMEFTEAESNMNDLVAEYQQYQDADIEYDDNDIEVAHDGVSL	445
EOY25532.1	EGMDMEFTEAESNMNDLVAEYQQYQDADIEYDDNDIEVAHDGVSL	446
ABY86660.1	EGMDMEFTEAESNMNDLVAEYQQYQDADIEYDDNDIEVAHDGVSL	445
EMS61990.1	EGMDMEFTEAESNMNDLVAEYQQYQDADIEYDDNDIEVAHDGVSL	445
Consensus	egmdemefteaeasnndmlvaeyqqydat e e e	445

sc-36: 千里光(*Senecio scandens*); XP_002330612.1: 毛果杨(*Populus trichocarpa*); ACM78033.1: 柚子(*Citrus maxima*); AGH08242.1: 钻天柳(*Salix arbutifolia*); ABS50668.1: 大桉树(*Eucalyptus grandis*); EPS63676.1: 螺旋狸藻(*Genlisea aurea*); XP_003594596.1: 荟藜苜蓿(*Medicago truncatula*); JC2511: R2242-水稻(R2242-rice); EOY25532.1: 可可树(*Theobroma cacao*); ABY86660.1: 陆地棉(*Gossypium hirsutum*); EMS61990.1: 乌拉尔图小麦(*Triticum urartu*)。Consensus: 序列同源性。同源性≥33%用灰色表示; 同源性≥50%用黄色表示; 同源性≥75%用蓝色表示; 同源性为100%用黑色背景表示。红色标线分别表示氨基酸位点第1~4位、第99~106位、第140~146位、第391~394位、第390~410位。

Amino acid homologs in at least 33% are highlighted in gray, amino acid homologs in at least 50% are highlighted in light yellow, amino acid homologs in at least 75% are highlighted in blue. Identical amino acids are highlighted in black. Red line refers to the amino acid sites as 1~4, 106~106, 140~146, 391~394, 390~410 in β -tubulin protein.

图 1 β -微管蛋白氨基酸序列特征及比对结果

Fig. 1 Amino acid sequence characteristics and multiple alignment results of β -tubulin

物中的序列具有非常高的保守性。同源比对还发现, 千里光 β -微管蛋白的氨基酸序列与柚子 (*Citrus maxima*) β -tubulin 氨基酸序列 (GenBank 登录号: ACM78033.1) 的同源性最高, 达 96%。

2.2 理化性质预测分析

经在线软件分析, 千里光 β -微管蛋白序列长度为 448 个氨基酸, 蛋白分子量为 50.01 kD, 理论等电点 pI 为 4.83, 原子组成为 $C_{2185} H_{3375} N_{603} O_{683} S_{31}$ 。不稳定系数为 37.72, 表明该蛋白状态稳定。正电荷氨基酸残基(Arg + Lys)38 个, 负电荷残基(Asp + Glu)61 个。疏水性结果分析表明, 最大疏水值为 2.167, 在该多肽中说明该处的疏水性最强; 亲水峰最大值为 -3.256, 整个蛋白质表现出高度的亲水性, 其平均疏水值为 -0.320。

2.3 结构预测

2.3.1 二级结构预测分析

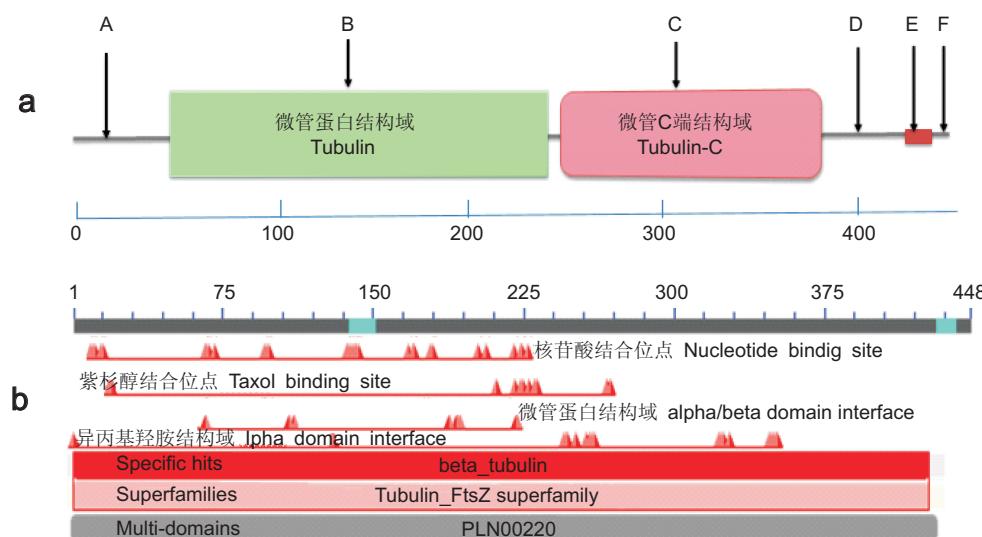
对千里光的 β -微管蛋白氨基酸序列进行二级结构分析, 结果显示其二级结构中 135 个氨基酸残基为 α 融合结构 (alpha helix, 30.13%), 221 个氨基酸残基为无规则卷曲结构 (random coil, 49.33%), 92 个氨基酸残基为线状结构 (extended strand, 20.54%), 表明该蛋白质二级结构的主要组成部分为无规则卷曲结构和 α 融合结构。

2.3.2 结构域预测分析

用 SMART 软件对 β -微管蛋白氨基酸序列进行结构域预测结果显示 (图 2: a), 该蛋白的 N 端 A 处为 β -微管蛋白未知结构位点, 包含 1~46 位氨基酸残基; B 处为微管蛋白结构域 (47~245), 属于微管蛋白/细菌 FtsZ 家族, 是 GTPase 结构域, 该区域可以与 GTP 进行结合; C 处为微管 C 端结构域 (246~383), 属于微管蛋白/细菌 FtsZ 家族; D 处为未知结构位点 (384~430) 结构域; E 处为微复杂结构位点 (431~440); F 处为未知结构位点 (441~448)。应用 GenBank 对该氨基酸序列进行保守结构域分析, 其结果与 SMART 软件分析结果一致 (图 2: b), 均含有 β 结构域和 Ipha N-异丙基羟胺 2 个结构域, 这 2 个结构域均属于微管蛋白/细菌 FtsZ 家族。

2.3.3 三维结构预测分析

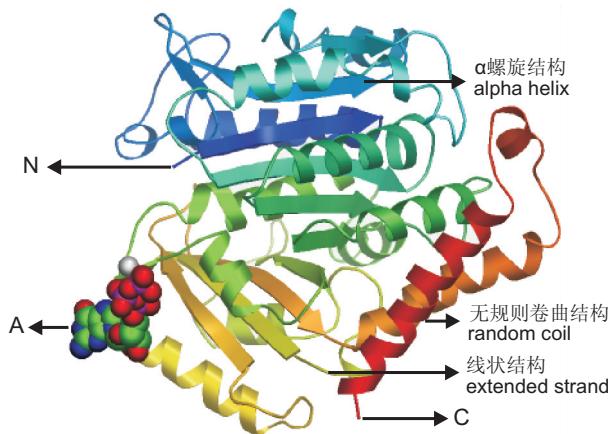
对 β -微管蛋白三维结构的预测结果显示 (图 3), 该蛋白包含 2 个结构域, 与结构域预测的结果一致。图 3 中该蛋白微管结构域为类球形结构, α 融合结构位于模型内部, 表面被无规则卷曲结构与线状结构环绕, 能清晰地辨别出蓝色端为 N 端、红色端为 C 端; A 处为微管 C 端结构域, 该处可能为蛋白微管结合蛋白结构域。



a: SMART 预测结果; b: GenBank 数据库分析结果。A、B、C、D、E、F 分别代表氨基酸位点第 1~46 位、第 47~245 位、第 246~383 位、第 384~430 位、第 431~440 位、第 441~448 位。

a: By online server SMART prediction; b: By homologous alignment from GenBank database. A, B, C, D, E and F refer to the amino acid sites 1~46, 47~245, 246~383, 384~430, 431~440, 441~448, respectively.

图 2 结构域预测结果
Fig. 2 Predicted domains of Tubulin-beta



A: 微管 C 端结构域; C: 碳端; N: 氨端。
A refers to C terminal domain of Tubulin-beta, C refers to the carbon terminal, N refers to the nitrogen terminal.

图 3 β -微管蛋白的三维结构模型

Fig. 3 Three-dimensional structure model of Tubulin-beta

2.4 预测分析亚细胞定位、功能及信号肽

对目的基因的功能定位进行 Psort 分析,结果显示 β -微管蛋白主要定位于细胞的细胞质、过氧化物酶体、线粒体基质等亚细胞器位置。通过 CBS 在线 Profun 软件分析其功能,结果发现该蛋白可能具有信号传导、生长因子、胁迫应答、运载体、结构蛋白、阳离子通道等功能(表 1)。对 β -微管蛋白进行信号肽预测,得分最高的是第 145 位氨基酸残基,分值达 0.153,即该处为最有可能的原始剪切位点;具有最高信号肽预测得分的是第

276 位氨基酸残基,分值为 0.284;最有可能的综合剪切位点是第 280 位,分值为 0.134。判断是否为分泌蛋白的标准是 mean S-score > 0.5, β -微管蛋白未达到该数值,说明该蛋白不是分泌蛋白,也表明该蛋白是在细胞内表达后直接在胞内起生理作用的蛋白。

3 讨论

对千里光 β -微管蛋白的核苷酸和氨基酸序列分析表明,该蛋白是相对保守的结构蛋白。千里光 β -微管蛋白的氨基酸序列与其它植物的同源性均在 90%以上,与柚子(*Citrus maxima*) β -tubulin 氨基酸序列(ACM78033.1)的同源性高达 96%。图 1 红色标记部分表明该蛋白质的氨基酸序列一级结构 N 端 1~4 位为 MREI, 是 β -微管蛋白转录后调控信号片段, 第 140~146 位含有 GGGTGSG 结构域, 具有微管蛋白标志信号片段, 该结构域也是 GTP 核苷酸结合位点, 是 α 、 β 两种微管蛋白聚合的必要结构^[13]; 其含有的第 391~394 位 RCAF 在 β -微管蛋白多肽折叠中也起着重要作用^[7]; 研究表明 α 、 β -微管蛋白都含有微管蛋白 2 个保守区, 第 99~106 位 NNWAKGHY 与第 390~410 位 RKAFL-HWYTGEGMDEMEFTE, 其第 Y₃₉₇ 位在千里光 β -微管蛋白中为 C₃₉₇ 位, 验证了 β -微管蛋白基因在不同生物类群之间既有很高的保守性又存在一定的变异性^[14~16]。

蛋白质结构是功能的基础,根据分析结果,千里光 β -微管蛋白二级结构含有无规则卷曲结构,理化性质表现出高度亲水性,推测该蛋白可能有以下功能:(1)作为水分结合蛋白减少水分丧失;(2)作为水合缓冲剂调节水状态;(3)作为水基质或替代物同大分子表面反应,阻止脱水组织中蛋白质的失活^[17]。

对千里光 β -微管蛋白的保守结构域预测表明,该蛋白有 2 个高度保守的结构域,即微管结构域和 C 端结构域,均属于微管蛋白/细菌 FtsZ 家族。该蛋白微管结构域含有 GTP 核苷酸结合位点,是微管蛋白保守较强结构域;其 C 端结构域在三级结构中由小球状体组成不规则结构(图 3 中 A、C 处),据报道,很多微管结合蛋白是通过结合微管

表 1 β -微管蛋白功能预测结果

Table 1 Functional predictions of tubulin-beta protein in *S. scandens*

功能分类 Gene ontology category	预测值 Probability	概率 Odds.
生长因子 Growth factor	0.007	0.520
胁迫应答 Stress response	0.032	0.368
免疫应答 Immune response	0.012	0.137
结构蛋白 Structural protein	0.010	0.342
运载体 Transporter	0.025	0.233
离子通道 Ion channel	0.010	0.169
阳离子通道 Cation channel	0.010	0.215
电压门控离子通道 Voltage gated ion channel	0.005	0.206
受体 Receptor	0.003	0.019
荷尔蒙 Hormone	0.001	0.206
信号传导 Signal transducer	0.052	0.241
金属离子转移 Metal ion transport	0.009	0.020

末端起微调控作用的^[18]，推测该处为微管结合蛋白功能行使的重要位点。植物微管蛋白行使功能，除需要其自身组成微管蛋白，还需要与微管蛋白相结合的微管结合蛋白的参与^[19]，微管结合蛋白能结合调配微管装配，进行数次聚合—解聚循环，仍能结合于微管蛋白^[20]，可见其与微管蛋白关系紧密。

本研究首次对千里光 β -微管蛋白基因进行了生物信息学分析，为进一步阐明微管蛋白的作用机制及功能应用奠定了理论基础，为植物 β -微管蛋白基因的分子研究提供了理论依据及基础资料。

参考文献：

- [1] 国家药典委员会. 中华人民共和国药典：一部(2010年版)[S]. 北京：中国医药科技出版社，2010；32.
- [2] 全国中草药汇编编写组. 全国中草药汇编[M]. 上海：上海科技出版社，1999；1390.
- [3] 孟凡君，张雪君，谢卫东. 中草药千里光研究进展[J]. 东北农业大学学报, 2010, 41(9)：156–160.
- [4] 钱刚，敖弟书，段鹏敏，蒋思恩，李有. 千里光抗菌作用的数量性状分析[J]. 武汉植物学研究, 2010, 28(1)：67–71.
- [5] 闫硕，朱家林，朱威龙，潘李隆，张青文，刘小侠. 棉铃虫 α -微管蛋白基因的克隆、序列分析及表达模式检测[J]. 中国农业科学, 2013, 46(9)：1808–1817.
- [6] Kawasaki H, Sugaya K, Quan GX, Nohata J, Mita K. Analysis of α -and β -tubulin genes of *Bombyx mori* using an EST database [J]. *Insect Biochem Molec*, 2003, 33(1)：131–137.
- [7] Song L, Liu XX, Zhang YA, Zhang QW, Zhao ZW. The cloning and expression of α -tubulin in *Monochamus alternatus* [J]. *Insect Mol Biol*, 2008, 17(5)：495–504.
- [8] 李艳，郭其森. β -tubulin 在 NSCLC 中的表达与抗微管类药物耐药相关性的研究进展[J]. 中华肺部疾病杂志(电子版), 2013, 6(2)：179–183.
- [9] 闫硕，刘小侠，韩娜娜，倪慧，孙洁茹，张青文. 棉铃虫 β -微管蛋白基因 cDNA 序列的克隆与序列分析 [J]. 应用昆虫学报, 2012, 49(1)：130–137.
- [10] 平军娇，张珍，蔡振峰，汤贤春，钱刚. 千里光全长 cDNA 文库的构建与分析 [J]. 中草药, 2012, 43(3)：557–561.
- [11] Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modeling [J]. *Bioinformatics*, 2006, 22(2)：195–201.
- [12] Kiefer F, Arnold K, Kunzli M, Bordoli L, Schwede T. The SWISS-MODEL repository and associated resources [J]. *Nucleic Acids Res*, 2009, 37:D387–D392.
- [13] Kirschner M, Mitchison T. Beyond self-assembly: from microtubules to morphogenesis [J]. *Cell*, 1986, 45: 329–342.
- [14] Yen TJ, Machlin PS, Cleveland DW. Auto regulated instability of beta-tubulin mRNAs by recognition of the nascent amino terminus of beta-tubulin [J]. *Nature*, 1988, 334(6183)：580–585.
- [15] Cleveland DW. Auto regulated control of tubulin synthesis in animal cells [J]. *Curr Opin Cell Biol*, 1989, 1(1)：10–14.
- [16] Aroca A, Raposo R, Lunello P. A biomarker for the identification of four *Phaeoacremonium* species using the β -tubulin gene as the target sequence [J]. *Appl Microbiol Biot*, 2008, 80(6)：1131–1140.
- [17] Soulages JL, Kim K, Walters C, Cushman JC. Temperature induced extended helix/random coil transitions in a group 1 late embryogenesis abundant protein from soybean [J]. *Plant Physiol*, 2002, 128: 822–32.
- [18] 黄聪聪，吴忠义，陈洁，于荣. 植物微管结合蛋白的研究进展 [J]. 植物学通报, 2008, 25(3)：354–362.
- [19] Hamada T. Microtubule associated proteins in higher plants [J]. *Plant Res*, 2007, 120(1)：79–98.
- [20] Ehrhardt DW, Shaw SL. Microtubule dynamics and organization in the plant cortical array [J]. *Annu Rev Plant Biol*, 2006, 57(1)：859–875.