

葡萄 *Vv*WRKY13 转录因子的生物信息学分析

侯丽霞, 王文杰, 刘 新

(青岛农业大学 生命科学院, 山东省高校植物生物技术重点实验室, 山东 青岛 266109)

摘 要:利用生物信息学方法对葡萄 *Vv*WRKY13 基因编码蛋白的理化性质、保守结构域、亲水性/疏水性、信号肽、跨膜结构域、磷酸化修饰、二级结构、亚细胞定位和系统进化等进行了预测和分析。结果表明:该蛋白分子量为 25 718.1 Da, 等电点为 9.08, 不稳定系数为 60.78; C 端含有 WRKY 保守域, 属于 WRKY 转录因子家族成员; 不具有信号肽和跨膜域, 是亲水性蛋白; 蛋白质二级结构最主要的结构元件是 α -螺旋和无规则卷曲, 并散布于整个蛋白; 亚细胞定位预测显示该蛋白主要位于细胞核。该研究结果将为进一步研究其生物学功能提供参考。

关键词:葡萄; *Vv*WRKY13; 转录因子; 生物信息学

中图分类号:Q 946.1; S 663.1 **文献标识码:**A **文章编号:**1001-0009(2013)13-0117-04

葡萄(*Vitis vinifera* L.)具有优良的酿酒特性和较高的食用价值,在世界范围内被广泛种植,也是我国的主要经济果树之一。2007 年法国科学家宣布葡萄的基因组测序已经完成,是迄今为止第 4 种基因组被完全破译的开花植物^[1]。这将为利用分子生物学改良葡萄品种提供有利条件。

WRKY 转录因子在植物中广泛存在,是植物中非常重要的一类转录因子。最早是在甘薯中克隆出来,随后相继在皱叶欧芹、拟南芥、烟草和水稻等十多种植物中都克隆出 WRKY 基因^[2-6]。WRKY 家族成员数量较多,共同特征是 N 端的高度保守的 WRKY 结构域和 C 端的锌指结构域。根据 WRKY 结构域的数目和锌指结构的类型,可将 WRKY 蛋白可分成 3 种类型,I 型含 2 个 WRKY 结构域,锌指结构模式为 C-X₄₋₅-C-X₂₂₋₂₃-H-X-H; II 型和 III 型均含 1 个 WRKY 结构域,II 型的锌指结构模式与 I 型相同,III 型的锌指结构域为 C-X₇-C-X₂₂₋₂₃-H-X-C^[7]。其中,WRKY 结构域可以特异性识别 W-box 顺式作用元件,启动调节相关基因的表达来应答多种生理过程^[8]。文献报道 WRKY 转录因子既参与高温、低温、高盐、干旱和渗透等非生物胁迫,又参与植物的生长发育及物质代谢途径,而且在抗细菌、病毒、虫害和机械伤害等方面也起着重要的调节作用,表现出 WRKY 转录因子参与

调控的多样性和重要性^[9-10]。如大豆的 *Gm*WRKY13 基因受到盐胁迫和甘露醇模拟的干旱胁迫后表达量均升高,推测其参与了盐胁迫和干旱胁迫过程^[11];在白叶枯病菌和稻瘟病菌侵染水稻的过程中发现被诱导 OsWRKY13 被强烈诱导表达,过表达 OsWRKY13 的转基因水稻还表现出对白叶枯病菌和稻瘟病菌的明显抗性^[12]。

该研究利用生物信息学方法,对葡萄 *Vv*WRKY13 基因编码的蛋白质进行分析,预测其信号肽、跨膜结构域、保守结构域、亚细胞定位以及与其它物种的进化关系等,以期为进一步研究 *Vv*WRKY13 的生物学功能以及调节机制奠定基础。

1 材料与方法

1.1 试验材料

通过美国国立生物信息中心 NCBI 网站(<http://www.ncbi.nlm.nih.gov>)查找 WRKY 基因,根据 GenBank 登录号为 XM_002278988,查找到 *Vv*WRKY13 全长 cDNA 序列,共 681 bp,编码 226 个氨基酸。

1.2 试验方法

利用在线软件 ProtParam(<http://web.expasy.org/protparam/>)对 *Vv*WRKY13 氨基酸序列进行分析;用 NCBI 网站在线软件 Conserved Domains(<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>)预测保守结构域;利用在线软件 NetPhos(<http://www.cbs.dtu.dk/services/NetPhos/>)分析磷酸化位点;利用在线软件 SPOMA (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html)预测二维结构;利用 <http://www.cbs.dtu.dk/services/TMHMM/>预测跨膜结构域;利用 <http://www.cbs.dtu.dk/services/TMHMM/>预测跨膜结构域;利用 <http://www.cbs.dtu.dk/services/TMHMM/>

第一作者简介:侯丽霞(1978-),女,博士,讲师,研究方向为葡萄分子细胞生物学。E-mail:houliaxia78@163.com.

责任作者:刘新(1966-),女,博士,教授,研究方向为植物逆境生理与分子生物学。E-mail:liuxin6080@yahoo.com.cn.

基金项目:青岛市自然科学基金资助项目(12-1-4-5-(12)-jch;10-3-4-5-5-jch);青岛农业大学高层次人才启动基金资助项目(630722)。

收稿日期:2013-03-04

dtu.dk/services/SignalP 预测信号肽;利用 ProtScale (<http://web.expasy.org/protscale/>)分析氨基酸序列的疏水性/亲水性;利用 <http://wolfsort.org/>进行亚细胞定位的预测;利用软件 MEGA4.0 采用 N-J 法构建进化树。

2 结果与分析

2.1 葡萄 VvWRKY13 蛋白质理化性质分析

通过在线软件 ProtParam 对葡萄转录因子 VvWRKY13 进行分析,由图 1 可知,VvWRKY13 的分子式为 $C_{1127}H_{1767}N_{329}O_{338}S_{12}$,分子量约为 25 718.1 Da,等

电点为 9.08,不稳定系数为 60.78(40 以下为稳定蛋白),推测其为不稳定蛋白。该蛋白中甘氨酸、精氨酸、亮氨酸、赖氨酸和丝氨酸含量相对较多,而半胱氨酸、色氨酸和酪氨酸含量相对较少。

2.2 葡萄 VvWRKY13 保守结构域分析

利用 Conserved Domains 程序对葡萄 VvWRKY13 进行结构域分析,从图 2 可以看出,第 149~208 位氨基酸是典型的 WRKY 转录因子家族成员的保守结构域。而且第 175~206 位氨基酸符合 WRKY 家族 II 型成员的特征,因此 VvWRKY13 属于 II 型 WRKY 转录因子。

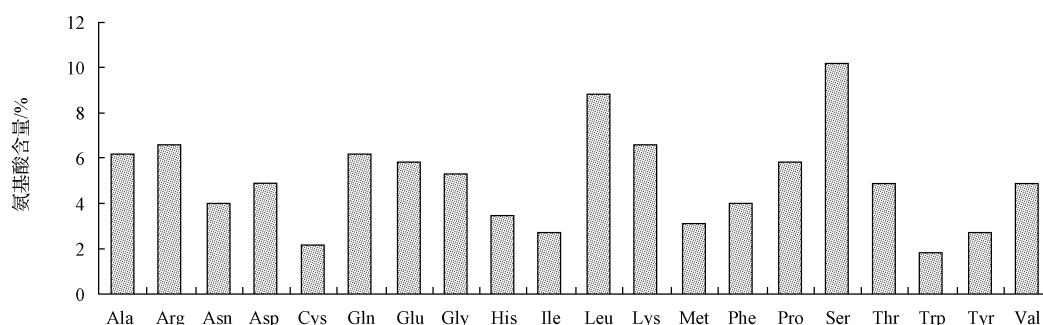


图 1 VvWRKY13 的氨基酸含量分布

Fig. 1 The distribution graphs of amino acid content of VvWRKY13

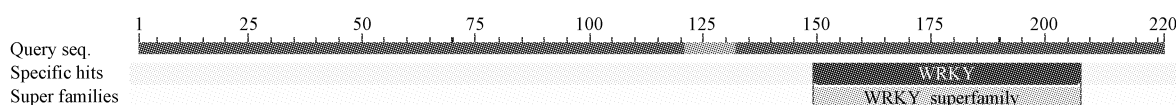


图 2 VvWRKY13 保守结构域分析

Fig. 2 The analysis of conserved domain prediction of VvWRKY13

2.3 葡萄 VvWRKY13 氨基酸序列疏水性/亲水性分析

蛋白质结构的特征是疏水/亲水间的平衡,了解氨基酸序列疏水性/亲水性对蛋白质的结构和功能预测有一定的作用。用 ProtScale 分析 VvWRKY13 氨基酸序列的疏水性/亲水性,根据氨基酸分值越高疏水性越强得分越低亲水性越强的规律,从图 3 可以看出,

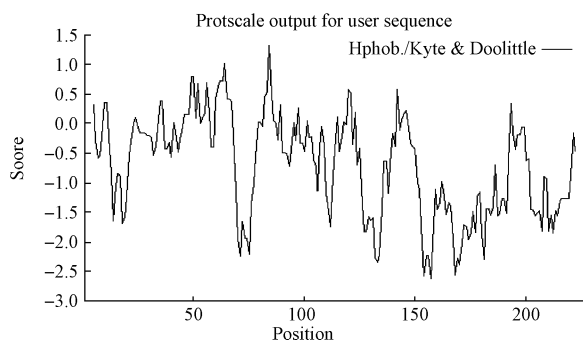


图 3 VvWRKY13 氨基酸序列疏水性/亲水性预测

Fig. 3 Predication of hydrophobicity/hydrophilicity for VvWRKY13

VvWRKY13 在 70 个氨基酸附近包含 1 个亲水头部,130 个氨基酸后也表现出一定的亲水性,属于亲水性蛋白。

2.4 葡萄 VvWRKY13 的信号肽和跨膜结构域分析

利用 SignalP 和 TMHMM 对 VvWRKY13 的信号肽和跨膜域进行预测,由图 4A 可知,C 值最大切割点在第 54 个氨基酸位置,分值为 0.17,综合剪切点分值(Y 值)最高也在第 54 个氨基酸位置,为 0.16,信号肽最大分值(S 值)在第 53 个氨基酸位置,为 0.16,因此判断 VvWRKY13 没有信号肽,属于非分泌性蛋白。由图 4B 可知,VvWRKY13 整个蛋白都在膜外,因此不具有跨膜结构域。

2.5 葡萄 VvWRKYs 氨基酸序列翻译后的磷酸化修饰预测

细胞内蛋白质磷酸化在信号传导中发挥着重要的作用,研究报道一共有 3 种主要的磷酸化部位-丝氨酸、苏氨酸和酪氨酸结合位点。大多数蛋白质是在丝氨酸和苏氨酸残基上磷酸化,而许多与信号转导有关的蛋白质还在酪氨酸位置上被磷酸化。利用在线软件 NetPhos

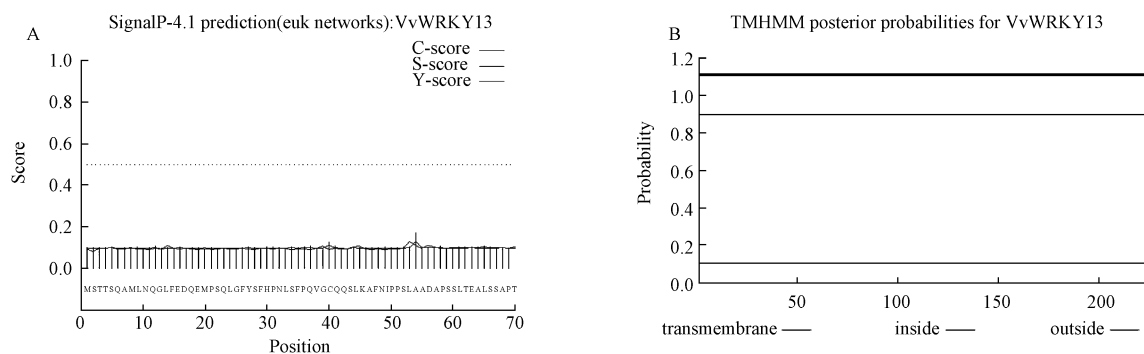


图4 VvWRKY13 的信号肽(A)和跨膜结构域(B)预测

Fig.4 Prediction of signal peptides and trans-membrane regions

分析可看出,在 VvWRKY13 中含有 7 个丝氨酸磷酸化位点,2 个苏氨酸磷酸化位点,1 个酪氨酸磷酸化位点。WRKY 蛋白磷酸化后,改变蛋白质的活性,这种改变可能为激活,也可能是抑制作用,从而参与植物的应答反应。

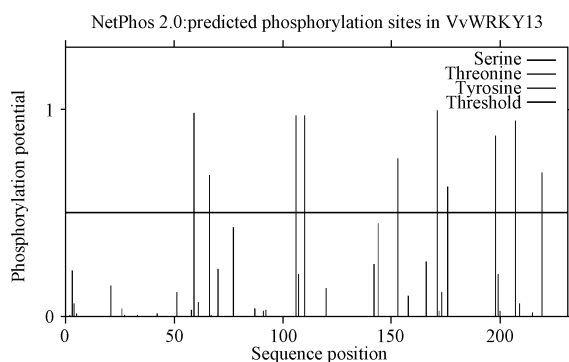


图5 VvWRKY13 氨基酸序列翻译后的磷酸化修饰预测

Fig.5 Phosphorylation site prediction of VvWRKY13

2.6 葡萄 VvWRKY13 二级结构预测

多肽链借助氢键排列成沿一维方向而呈现有规则的重复构象的二级结构,是氨基酸顺序与三维构象之间的桥梁。二级结构借助范德华力、氢键、静电和疏水等相互作用形成蛋白质的三级结构,从而发挥正常的生物学功能。利用 SPOMA 对 VvWRKY13 转录因子的氨基酸序列的二维结构进行预测,由图 6 可知,VvWRKY13 含有无规则卷曲、 α -螺旋、 β -折叠延伸链和 β -转角结构,其中无规则卷曲部分较多,约占 54.87%,其次为 α -螺旋结构,占 26.99%, β -折叠片层结构占 13.72%,最少的是 β -转角,占 4.42%。可见无规卷曲和 α -螺旋是 VvWRKY13 的最主要二级结构,分布于整条肽链中。

2.7 葡萄 VvWRKY13 的亚细胞定位预测

了解蛋白的亚细胞定位可以为蛋白功能预测提供基础。利用在线软件 WoLF PSORT 对 VvWRKY13 的亚细胞定位进行预测,结果表明,此蛋白在细胞核中的定位预测值为 10.0,而在叶绿体、细胞质和细胞外的预测值都为 1.0,推测 VvWRKY13 主要存在于细胞核中,这与转录因子主要在细胞核中发挥功能吻合。

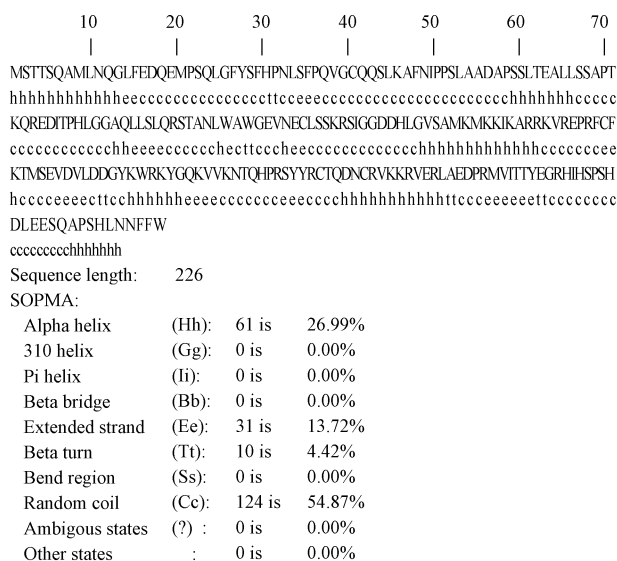


图6 VvWRKY13 的二级结构预测

注: Hh: α -螺旋; Ee: β -折叠延伸链; Tt: β -转角; Cc: 无规则卷曲。

Fig.6 The prediction of secondary structure of VvWRKY13

Note: Hh: α -helix; Ee: β -folding extension chain; Tt: β -sheet; Cc: Random coil.

2.8 葡萄 VvWRKY13 的系统进化树分析

通过 NCBI 的 BlastP 检索到与 VvWRKY13 同源性较高的蛋白,如毛果杨(*Populus trichocarpa*: EEE90494.1)、拟南芥(*Arabidopsis thaliana*: AEE87071.1)、水稻(*Oryza sativa*: DAA05078.1)、玉米(*Zea mays*: ACN34879)、黄瓜(*Cucumis sativus*: ADU52523.1)、高粱(*Sorghum bicolor*: XP_002455053)、蓖麻(*Ricinus communis*: XP_002534045.1)、苹果(*Malus domestica*: ADL36863.1)、苜蓿(*Medicago truncatula*: XP_003602502.1)、二穗短柄草(*Brachypodium distachyon*: XP_003581484.1)和西洋参(*Panax quinquefolius*: AEQ29015.1),对这 11 种植物的 WRKY 进行蛋白序列比对并构建系统发育树,由图 7 可知,葡萄的 VvWRKY13 与毛果杨、蓖麻和苹果的 WRKY 亲缘关系较近,聚为一类,同属于木本植物。而与其它植物相对亲缘关系较远。

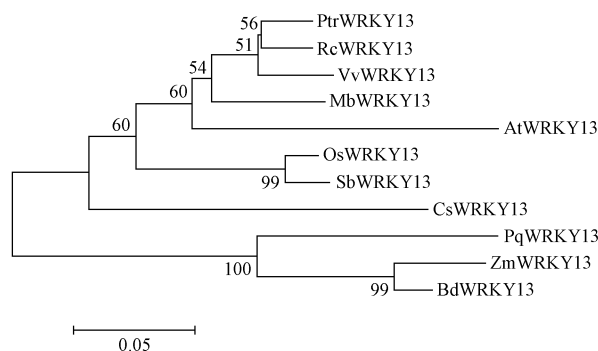


图7 植物 VvWRKY13 蛋白同源序列的系统发育树

Fig. 7 Phylogenetic tree of PR1 proteins homolog sequences from plants

3 结论与讨论

WRKY 转录因子超家族成员众多,如在模式植物拟南芥中发现 74 个成员,在水稻中至少有 100 多个成员。它们参与了植物的生物胁迫、非生物胁迫、生长发育和器官衰老等过程^[13]。该试验利用生物信息学方法预测了葡萄 VvWRKY13 转录因子的理化性质、结构域和亚细胞定位等基本信息,结果表明,VvWRKY13 的 C 端具有一个 WRKYGQK 的保守结构域,属于 II 型的 WRKY 家族成员,这与仇玉萍等^[14] 研究结果一致。VvWRKY13 没有信号肽和跨膜结构域,属于亲水性蛋白。此外,VvWRKY13 具有多个丝氨酸、苏氨酸和酪氨酸的磷酸化修饰位点,可能与其调节下游基因表达的功能相适应。二级结构预测表明,VvWRKY13 以无规卷曲和 α -螺旋为主,推测与其活性位点形成有关。亚细胞定位结果显示其定位于细胞核,这是作为转录因子发挥功能的必要条件之一。该研究将为利用试验手段验证其功能提供依据,并为进一步利用基因工程改良葡萄品种奠定基础。

Bioinformatics Analysis of VvWRKY13 Transcription Factor in *Vitis vinifera* L.

HOU Li-xia, WANG Wen-jie, LIU Xin

(Key Lab of Plant Biotechnology in Universities of Shandong Province, College of Life Science, Qingdao Agricultural University, Qingdao, Shandong 266109)

Abstract: The characteristic of VvWRKY13 transcription factor in grape was analyzed and predicted by bioinformatic methods in the following aspects: physical and chemical characters, conserved domain, hydrophobicity or hydrophilicity, signal peptide, transmembrane structure, phosphorylation in amino acid, secondary and tertiary structure of protein, subcellular localization and molecular phylogenetic evolution. The results showed that VvWRKY13 had molecular weight 25 718.1 Da, isoelectric point 9.08 and instability coefficient 60.87; it possessed one conserved WRKY domain, so it was II type WRKY transcription factor; VvWRKY was a hydrophilicity protein without a signal peptide and transmembrane structure; random coil and α -helix were the major structural elements of polypeptide chain; subcellular localization result indicated that VvWRKY13 was located in nucleus very likely. This research would be of great information and values for its function study.

Key words: *Vitis vinifera* L.; VvWRKY13; transcription factor; bioinformatic

参考文献

- [1] 房经贵,刘洪,杨光. 基因组学研究时代的葡萄酿造明星-黑比诺[J]. 中外葡萄与葡萄酒,2008(4):57-59.
- [2] Rushton P J, Macdonald H, Huttly A K, et al. Members of a new family of DNA-binding proteins bind to a conserved cis-element in the promoters of α -Amy2 genes[J]. Plant Mol Biol, 1995, 29(4): 691-702.
- [3] Rushton P J, Torres J T, Parniske M, et al. Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley PR1 genes[J]. Embo J, 1996, 15(20): 5690-5700.
- [4] dePater S, Greco V, Pham K, et al. Characterization of a zinc-dependent transcriptional activator from Arabidopsis[J]. Nucleic Acids Res, 1996, 24: 4624-4631.
- [5] Kim C Y, Zhang S. Activation of a mitogen-activated protein kinase cascade induces WRKY family of transcription factors and defense genes in tobacco[J]. Plant J, 2004, 38(1): 142-151.
- [6] Xu Y H, Wang J W, Wang S, et al. Characterization of GaWRKY1, a cotton transcription factor that regulates the sesquiterpene synthase gene (+)- δ -cadinene synthase-A[J]. Plant Physiol, 2004, 135(1): 507-515.
- [7] Eulgem T, Rushton P J, Robatzek S, et al. The WRKY superfamily of plant transcription factors[J]. Trends in Plant Sci, 2000(5): 199-206.
- [8] 刘戈字,胡莺雷,祝建波. 植物 WRKY 蛋白家族的结构及其功能[J]. 生命的化学, 2006, 26(3): 231-233.
- [9] Pandey S P, Somssich I E. The role of WRKY transcription factors in plant immunity[J]. Plant Physiol, 2009, 150(4): 1648-1655.
- [10] Chen L G, Song Y, Li S J, et al. The role of WRKY transcription factors in plant abiotic stresses[J]. Biochim et Biophys Acta, 2012, 1819(2): 120-128.
- [11] Zhou Q Y, Tian A G, Zou H F, et al. Soybean WRKY-type transcription factor genes, GmWRKY13, GmWRKY21, and GmWRKY54, confer differential tolerance to abiotic stresses in transgenic Arabidopsis plants[J]. Plant Biotechnol J, 2008, 6: 486-503.
- [12] Qiu D Y, Xiao J, Ding X H, et al. OsWRKY13 mediates rice disease resistance by regulating defense related genes in salicylate- and jasmonate-dependent signaling[J]. Mol Plant Microbe Interact, 2007, 20(5): 492-499.
- [13] 于延冲,乔孟,刘振华,等. WRKY 转录因子功能的多样化[J]. 生命科学, 2010, 22(4): 345-351.
- [14] 仇玉萍,荆邵娟,付坚,等. 13 个水稻 WRKY 基因的克隆及其表达谱分析[J]. 科学通报, 2004(18): 1860-1869.