

袁娜,戴琛,杨郁文,等.棉花多肽的分离提取方法建立及质谱鉴定[J].江苏农业科学,2019,47(20):95-98.

doi:10.15889/j.issn.1002-1302.2019.20.020

棉花多肽的分离提取方法建立及质谱鉴定

袁娜¹,戴琛²,杨郁文¹,杜建厂¹,张保龙¹

(1.江苏省农业科学院种质资源与生物技术研究所/江苏省农业生物学重点实验室,江苏南京 210014;

2.南京农业大学生命科学学院,江苏南京 210095)

摘要:多肽是一类小分子物质,在植物生长、发育及抗逆等众多生命过程中起着重要的调控作用。旨在探索棉花根部多肽的分离提取方法,建立基于改良尿素法的棉花多肽组样品收集方法。基于该方法,笔者在棉花根部组织中获得809个小分子肽序列,这些多肽的大小在10~70个氨基酸之间。随后,采用液相色谱-质谱联用技术,笔者进一步对获得的多肽进行来源鉴定和序列分析。通过功能注释和亚细胞定位预测,发现这些多肽来源的基因在结合和催化活性、细胞过程、代谢过程功能亚类中所占比例最高,并且这些基因大多分布在细胞质中(52%)。此外由结果还可以看出,与抗病防卫相关的蛋白包括PR10、包含NB-ARC结构域的蛋白、过氧化氢酶、氧化还原酶等来源的多肽还伴有甲酰化、脱酰胺基化、乙基化等10种以上的氨基酸修饰方式。该技术为后续棉花多肽组学研究、棉花多肽数据库的建立,以及挖掘如抗病抗逆相关的多肽应用于生产奠定了理论基础和技术支持。

关键词:棉花;多肽;蛋白质组学;质谱分析

中图分类号:S562.01 **文献标志码:**A **文章编号:**1002-1302(2019)20-0095-04

多肽是一类由10~100个氨基酸组成的、由前体多肽经过加工剪切后形成的成熟、具有特殊功能的肽类信号分子^[1]。在动物、细菌、真菌中,功能活性肽在机体的生长发育过程中具有重要的调控作用,它们作为药物、疫苗、酶抑制剂、导向药物及药物先导药等已在世界范围内被广泛应用,具有非常高的应用价值^[2-3]。随着系统素(systemin)、植物磺胺素(phytosulfokine,简称PSK)、CLE、SCR(SP11)、CLV3等植物多肽分子的陆续鉴定,证实了这类物质可以作为信号分子介导细胞与细胞之间的短距离信息交流,在植物生长、发育及抗逆等众多生命过程中起着重要的调控作用^[1,4]。

植物内源活性多肽不仅可以通过基因工程导入植物中进行过表达,以改善植物的生长发育及抗逆抗病性,还可以用作生长调节剂、杀菌剂,用来直接外源喷施植株,以提高植物的生长和抗逆能力。例如Zhang等分别将*alfAFP*基因及花烟草中的防御素*NaDI*基因转入陆地棉后,植株对黄萎病菌的抗性显著增强^[5-6]。将人工合成的抗菌肽转入棉花中,也同样增强了植株对黄萎病菌的抗性^[7-8]。此外,Huffaker等用ZmPep1多肽体外处理玉米叶片,可以减小由死体营养型病原菌玉米小斑病菌(*Cochliobolis heterostrophus*)和半活体营养型病原菌玉米茎腐病菌(*Colletotrichum graminicola*)引起的死亡细胞数量和坏死斑的大小^[9]。由此可见,植物多肽在植物应对病原菌的侵害过程中可以起到非常重要的作

用,具有重要的应用价值。

随着蛋白质组学的发展,多肽组学在近10年来也获得了较多关注。然而,由于植物材料的特殊性,相较于动物,植物多肽组学的发展仍较为缓慢。在本研究中,笔者参考动物多肽组的纯化过程^[10-11]和植物蛋白质组的提取方法^[12-13],对棉花根部的多肽提取方法进行探索和建立,利用高分辨级质谱质谱联用仪,对所提取的多肽进行NanoLC-MS/MS(纳升级液相色谱串联质谱)分析,利用已发表的棉花蛋白数据库进行搜索比对,并手动验证质谱的分析结果。本研究将为系统研究棉花多肽组的生物学信息和挖掘具有重要应用价值的新多肽提供新的方法和思路。

1 材料与方法

1.1 试验材料

供试棉花材料为亚洲棉中棉红茎,由江苏省农业科学院种质资源与生物技术研究所张保龙研究员提供。

1.2 试验时间和地点

本试验于2018年3月至2019年4月进行,试验在江苏省农业科学院种质资源与生物技术研究所江苏省重点实验室及南京农业大学生命科学学院完成。

1.3 试验方法

1.3.1 棉花根部多肽的分离提取 将亚洲棉中棉红茎的种子用清水冲洗后,播种于盆钵中,待幼苗长至2叶1心时,采集植物的根部组织进行多肽提取。用液氮研磨叶片至粉末状后,加入3倍体积的多肽提取液[含有7 mol/L尿素,2 mol/L硫代尿素,20 mmol/L DTT(二硫苏糖醇),5 μL Cocktail蛋白酶抑制剂],振荡提取15 min后,离心(4℃,12 000 × g,10 min),取上清。上清液通过Microcon® YM-10超滤管(Millipore, USA)除去分子量大于10 ku的大分子蛋白。在过

收稿日期:2019-05-29

基金项目:江苏省自然科学基金(编号:BK20170604)。

作者简介:袁娜(1989—),女,江苏宿迁人,博士,助理研究员,主要从事棉花抗病育种及生物信息学研究。E-mail: thefuries@163.com。

通信作者:杜建厂,博士,研究员,主要从事生物信息学研究。

E-mail: dujianchang@hotmail.com。

滤后的上清中加入 3 倍体积的预冷丙酮,于 $-20\text{ }^{\circ}\text{C}$ 放置 4 h。离心取沉淀($4\text{ }^{\circ}\text{C}$, $12\ 000\times g$, 10 min),沉淀再用 1 倍体积的预冷丙酮洗 1 次,离心($4\text{ }^{\circ}\text{C}$, $12\ 000\times g$, 5 min),弃上清,将沉淀倒置晾干。将沉淀加入 70% 乙腈[含 0.1% TFA(三氟乙酸)]中,振荡提取 15 min,离心取上清($4\text{ }^{\circ}\text{C}$, $12\ 000\times g$, 5 min)。用 Millipore[®] Ziptips (Merck KGaA, Darmstadt, Germany) 对上清液进行脱盐处理,最终收集后抽真空干燥。将处理后的样品用 0.1% 甲酸复溶,离心($4\text{ }^{\circ}\text{C}$, $12\ 000\times g$, 2 min),取上清液进行 Nano LC-MS/MS 分析。

1.3.2 Nano LC-MS/MS 分析分离鉴定 将多肽样品于 Nano-RSLC 液相系统(Thermo Fisher Scientific, CA, USA) 中进行 60 min 的梯度洗脱,流速为 300 nL/min,进样量为 10 μL 。采用纳升电喷雾离子源,电离模式:电喷雾 ESI(+),喷雾电压为 2.2 kV,毛细管温度为 $200\text{ }^{\circ}\text{C}$ 。一级质谱在 Orbitrap 中扫描,扫描范围是 350 ~ 1 800 m/z ,分辨率为 60 000。所得信息用 PEAKS 软件进行 De Novo(从头)测序,并用测得的序列搜寻棉花蛋白数据库,以得到最合适的鉴定结果。

2 结果与分析

2.1 棉花根部多肽的提取效果

植物内源多肽的丰度通常都非常低,并且植物细胞外的细胞壁中及细胞内含有大量色素,加上存在的次级代谢物等为多肽的获得和分离带来了极大困难。本研究由于后续应用高分辨率的质谱联用仪,因此在提取方法上排除了与色谱不相容的十二烷基硫酸钠(SDS)法,笔者采用对分子量极小的蛋白分辨效果好的改良尿素法进行棉花多肽的提取。为了避免蛋白质和多肽的降解,本试验在样品采集和提取过程中始终保持在低温($4\text{ }^{\circ}\text{C}$)下进行,并且在提取液中加入蛋白酶抑制剂。质谱分析后用 PEAKS studio 软件进行数据处理,将从头测序、数据库搜索 2 种方法结合进行多肽的鉴定。笔者从棉花根部组织共获得 809 个小分子肽序列,详见表 1。这一结果表明,采用的改良尿素法对棉花多肽的分离效果较好。这些多肽的长度均较小,最大不超过 70 个氨基酸,以长度为 10~20 个氨基酸的多肽最多,其次为 20~30 个氨基酸长度的多肽。

表 1 部分棉花根部的多肽信息

多肽序列	分子质量 (u)	长度 (氨基酸)	质荷比	基因名称	来源蛋白质
AKRLIPEFNRLVEKIHPSTNSGILLPEK	3 513.081	31	703.628 7	Cotton_A_01756	Chaperonin 10
VAKKSVVPIASGRQIVQSPT	2 064.206	20	689.075 3	Cotton_A_04099	Eukaryotic aspartyl protease family protein
NLDFLVNLVRD	1 316.709	11	659.364 1	Cotton_A_15196	Oxidoreductase, zinc - binding dehydrogenase family protein
GVVTYDYENTSPVAPARLF	2 141.043	19	1 071.530 3	Cotton_A_36744	PR protein class 10
SPDDVPKYWIAF	1 436.698	12	719.357 5	Cotton_A_30826	Glycosyl hydrolase family 38 protein
SVLLLPHSFTSF	1 346.723	12	674.384 5	Cotton_A_22880	Alpha - L - arabinofuranosidase I
AKLTKLTKDQADYISVPEVGPY	2 435.295	22	1 218.677 0	Cotton_A_07391	S - adenosyl - L - homocystein hydrolase
KAFVAVDIENLMPKVAPQ	1 870.002	17	936.026 8	Cotton_A_21816	VdI2
PPPAPLIDMEELTKW	1 735.886	15	868.950 4	Cotton_A_04273	Plasma membrane intrinsic protein 2,4
DDNFNITDDTRIRA AVPTIKYLMGHGSKVILS	3 559.846	32	890.969 5	Cotton_A_33273	phosphoglycerate kinase
MTIGGQPAGRIVMELF	1 718.885	16	860.468 4	Cotton_A_08910	cyclophilin
AGVTIAHGCVLPNPNVPL	1 740.989	18	871.515 5	Cotton_A_21432	Histone H2A 12
GVFSYDYETSPVAPARLF	2 119.026	19	1 060.545 0	Cotton_A_35117	pCPR10 - 16
VESQGFSGIEWSDGVHLVRSPIAVRWIQG	3 208.642	29	803.177 9	Cotton_A_18889	Subtilase family protein
VVVAADGSGNFKTVSEAVAKAPE	2 245.159	23	1 123.589 0	Cotton_A_18959	hypothetical protein PHAVU_010G123100g
VTYMPRPDGDAPRPVTLIPGDGIGPLVTNA VEQVM	3 616.875	35	1 206.652 0	Cotton_A_08855	Isocitrate dehydrogenase 1
KDEAMPDVKSLDIRNFI	1 990.019	17	664.347 8	Cotton_A_10892	probable ATP synthase 24 ku subunit
EFEVTSPIAPARLFKAF	1 922.030	17	962.034 9	Cotton_A_35116	ribonuclease - like PR - 10

2.2 棉花根部多肽的功能分类

由图 1 可见,通过比对棉花蛋白数据库,发现 737 个多肽可以匹配到已知蛋白上。通过对这些已知蛋白进行 GO 注释,笔者发现这些多肽来源的基因在结合和催化活性、细胞过程、代谢过程功能亚类中所占的比例最高(图 1)。从未来生产应用角度,本研究还发现了与抗病防卫相关的蛋白包括病程相关蛋白 10(PR10)、包含 NB-ARC 结构域的蛋白、过氧化氢酶、氧化还原酶等来源的多肽。例如,本研究发现病程相关蛋白 10 来源多肽的多个多肽,这些多肽除了序列较短,还伴有甲酰化、脱酰胺基化、乙基化等 10 种以上的氨基酸修饰方式,详见图 2。

2.3 棉花根部多肽的定位

笔者将多肽的来源基因进行亚细胞定位预测。图 3 的预测结果显示,根部多肽来源的基因大多分布在细胞质中(52%),其次分别分布在细胞核(23%)、质体(20%)中,此外,还有少量多肽分布于细胞质膜、内质网膜上、过氧化物酶上等。

3 讨论与结论

在动物细胞及动物组织中,多肽和多肽组的提取方法较为成熟,然而在植物中,由于植物细胞结构的特殊性,目前植物多肽组的提取仍然是多肽组学研究中的关键步骤。在蛋白

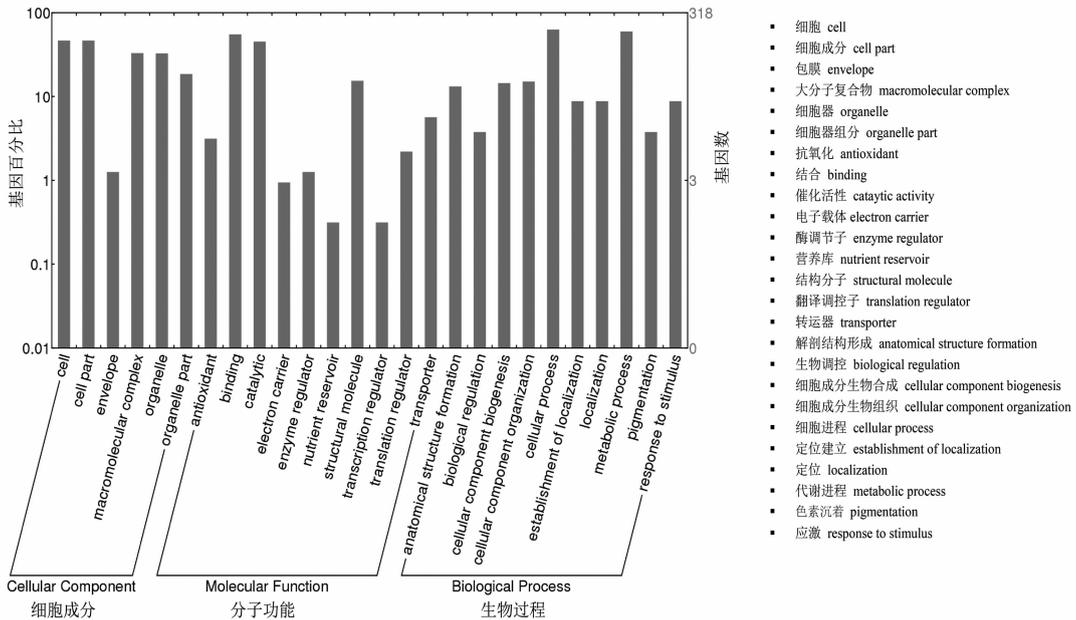


图1 棉花多肽基因的 GO 功能注释

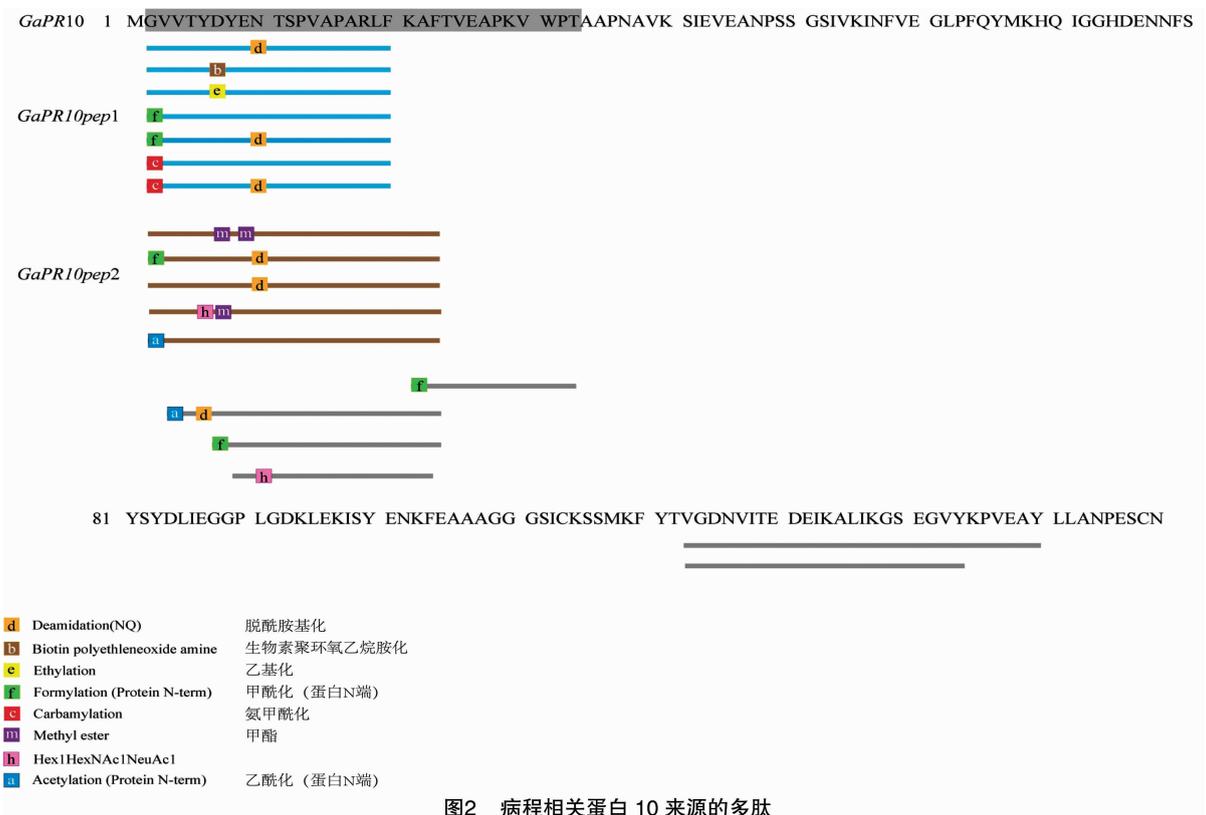


图2 病程相关蛋白 10 来源的多肽

质组的提取过程中,常使用的去垢剂如 SDS、Trixton 等可以快速高效地提取样本的蛋白质,然而由于这些去垢剂与色谱不相容,因此本研究采用了对分子量极小的蛋白分辨率好的改良尿素法对棉花多肽进行提取。孙婷婷等利用尿素法,基于 Nano LC - MS/MS 在水稻叶片和根部共鉴别出约 110 个多肽,通过功能分析发现,这些多肽参与了水稻重要的生理活动^[13]。在本研究中,笔者在提取液中加入蛋白酶抑制剂,

首先减少了内源降解的干扰,其次使得提取过程保持在低温环境下进行,另外,提取过程中的无机盐提取剂在最后的除盐步骤中被去除,极大地提高了植物多肽组的提取纯度和效率。通过高分辨率的质谱检测和数据整理,笔者发现这种方法的提取效果较好,在棉花根部组织中鉴定出 809 个多肽。

通过对这些多肽的前体蛋白功能进行功能注释后发现,它们参与到细胞组成、分子调控和生物代谢各个方面。例如

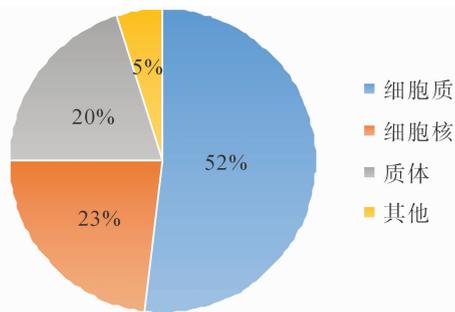


图3 棉花根部多肽的亚细胞分布

笔者发现,病程相关蛋白来源的多肽可以调控植物的胁迫反应。病程相关蛋白是植物在受到生物、非生物胁迫后产生的一种小分子蛋白,具有广泛的功能,可以硬化细胞壁、传导信号以及抗菌等^[14]。前人研究发现,病程相关蛋白 10 在参与植物的生长发育、次生代谢及应对外界生物和非生物胁迫时发挥着重要的作用^[15]。在对棉花的研究中发现,黄萎病菌侵染可以诱导 PR10 蛋白的表达^[16-17]。转化烟草、拟南芥的试验结果表明,PR10 可以提高植株对黄萎病菌的抗性,并且能提高植物抗菌素黄酮类化合物的水平。此外,PR10 的表达也能改变烟草、拟南芥的表型,如转基因植株生长延缓、转基因拟南芥叶片倾向于横向生长而呈近圆形、顶端优势减弱、侧生茎数量显著增多等现象,证实 PR10 也参与调控植物的生长发育^[18]。尽管 PR10 的部分重要功能已被鉴定,然而其调控机制到目前为止还不清晰。在本研究中,笔者通过质谱分析发现,PR10 可以产生多个不同修饰的多肽分子,因此推测 PR10 可能通过不同的多肽去调控不同的生长并应对环境胁迫反应。在后续研究中,笔者将聚焦这些重要多肽进行功能验证。

植物天然多肽组的提取是植物多肽组学发展的一切基础。通过探索高效的多肽分离方法,可以挖掘植物体中的天然活性多肽,后续将其作为生长调节剂或者抗菌药剂人工合成应用于生产,对于棉花增产抗病具有重要的理论与实践意义。

参考文献:

[1] Tavormina P, de Coninck B, Nikonorova N, et al. The plant peptidome: an expanding repertoire of structural features and biological functions[J]. *Plant Cell*, 2015, 27(8): 2095 - 2118.

[2] Brivanlou A H, Darnell J E Jr. Signal transduction and the control of gene expression[J]. *Science*, 2002, 295(5556): 813 - 818.

[3] Dave L A, Montoya C A, Rutherford S M, et al. Gastrointestinal endogenous proteins as a source of bioactive peptides - an in silico study[J]. *PLoS One*, 2014, 9(6): e98922.

[4] Czyzewicz N, Yue K, Beeckman T, et al. Message in a bottle: small signalling peptide outputs during growth and development [J]. *Journal of Experimental Botany*, 2013, 64(17): 5281 - 5296.

[5] Zhang H P, Wang X D, Shao M Y, et al. Expression of alfalfa antifungal peptide gene and enhance of resistance to *Verticillium dahliae* in upland cotton[J]. *Acta Agriculturae Scandinavica*, 2010, 60(1): 95 - 100.

[6] Gaspar Y M, McKenna J A, McGinness B S, et al. Field resistance to *Fusarium oxysporum* and *Verticillium dahliae* in transgenic cotton expressing the plant defensin NaD1 [J]. *Journal of Experimental Botany*, 2014, 65(6): 1541 - 1550.

[7] Rajasekaran K, Cary J W, Jaynes J M, et al. Disease resistance conferred by the expression of a gene encoding a synthetic peptide in transgenic cotton (*Gossypium hirsutum* L.) plants [J]. *Plant Biotechnology Journal*, 2005, 3(6): 545 - 554.

[8] 赵亦静, 倪密, 诺林, 等. 人工合成抗菌肽对棉花黄萎病菌的抑菌效果[J]. *浙江大学学报(农业与生命科学版)*, 2013, 39(1): 11 - 17.

[9] Huffaker A, Pearce G, Veyrat N, et al. Plant elicitor peptides are conserved signals regulating direct and indirect antiherbivore defense [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2013, 110(14): 5707 - 5712.

[10] Zhang X, Che F Y, Berezniuk I, et al. Peptidomics of Cpe (fat/fat) mouse brain regions: implications for neuropeptide processing [J]. *Journal of Neurochemistry*, 2008, 107(6): 1596 - 1613.

[11] Ma J, Ward C C, Jungreis I, et al. Discovery of human sORF - encoded polypeptides (SEPs) in cell lines and tissue [J]. *Journal of Proteome Resources*, 2014, 13(3): 1757 - 1765.

[12] Ge P, Ma C, Wang S, et al. Comparative proteomic analysis of grain development in two spring wheat varieties under drought stress [J]. *Analytical and Bioanalytical Chemistry*, 2012, 402(3): 1297 - 1313.

[13] 孙婷婷, 张乐乐, 王倩, 等. 基于 Nano LC - MS/MS 的水稻多肽组学研究 [J]. *植物生理学报*, 2015, 51(7): 1173 - 1178.

[14] Van Loon L C, Rep M, Pieterse C M. Significance of inducible defense - related proteins in infected plants [J]. *Annual Review of Phytopathology*, 2006, 44: 135 - 162.

[15] 卢一鹏, 李伟, 孙楠, 等. 植物病程相关蛋白 PR10 结构、功能及表达调控的研究进展 [J]. *湖北农业科学*, 2016, 55(2): 273 - 279, 284.

[16] 梁杉. 棉花抗黄萎病相关基因类 Pr10 的分离及功能分析 [D]. 保定: 河北大学, 2009: 19 - 34.

[17] 张桂荣. 棉花 PR 蛋白家族分析及 PR10、PR17 的功能研究 [D]. 保定: 河北农业大学, 2015: 11 - 46.

[18] 陈捷胤. 棉花重要抗黄萎病相关基因的表达特性分析及 GhPR10 功能初步研究 [D]. 北京: 中国农业科学院, 2007: 1 - 18.