

利用基因芯片技术筛选棉花产量性状相关的基因

张利媛^{1,2,3}, 于霁雯¹, 吴嫚¹, 马建辉¹, 范术丽¹, 宋美珍¹, 庞朝友¹, 李兴丽¹, 张金发¹, 喻树迅^{1,2*}

(1. 中国农业科学院棉花研究所 / 棉花生物学国家重点实验室, 河南 安阳 455000; 2. 华中农业大学植物科技学院, 武汉 430070; 3. 中国石油集团工程设计有限责任公司华北分公司, 河北 任丘 062552)

摘要: 根据皮棉产量数据的重复性测验及显著性分析, 从 SG747 (*Gossypium hirsutum* L.) 和 Giza75 (*Gossypium barbadense* L.) 的回交自交系群体中选取高产组和低产组品系各 3 个材料, 取这 2 组材料和其父母本开花后 10 d 的纤维进行芯片分析。比较芯片数据, 获得了 1508 个差异表达基因。对这些差异表达基因进行聚类分析, 结果显示高产组品系和低产组品系分别聚类在一起, 符合基于田间数据分组的预期结果。利用 Blast2GO 软件对 1508 个差异表达基因进行 Blast-Mapping-Annotation-KEGG 分析, 结果表明参与细胞质膜发育、过氧化物酶活性、抗逆和营养物质运输等生物学过程的基因最多。进一步利用 RT-PCR 分析, 得到一系列与纤维品质和产量等性状密切相关的基因。研究结果为棉花高产基因工程和分子育种提供了丰富的基因资源。

关键词: 棉花; 产量; 基因芯片; *GhMKRP2*; *GhPOD*

中图分类号: S562.035.3 **文献标志码:** A

文章编号: 1002-7807(2013)05-0417-09

Identification of Genes Related to Cotton Yield Using Affymetrix Gene Chips

ZHANG Li-yuan^{1,2,3}, YU Ji-wen¹, WU Man¹, MA Jian-hui¹, FAN Shu-li¹, SONG Mei-zhen¹, PANG Chao-you¹, LI Xing-li¹, ZHANG Jin-fa¹, YU Shu-xun^{1,2*}

(1. Institute of Cotton Research, Chinese Academy of Agricultural Sciences / State Key Laboratory of Cotton Biology, Anyang, Henan 455000, China; 2. College of Plant Technology, Huazhong Agricultural University, Wuhan 430070, China; 3. China Petroleum Engineering Huabei Company, Renqiu, Hebei 062552, China)

Abstract: Three lines from high yield and low yield groups were selected from a backcross inbred line (BIL) population, which was developed from a cross between upland cotton SG747 (*Gossypium hirsutum*) and Egyptian cotton Giza75 (*Gossypium barbadense* L.). Fiber RNA of 10 days post anthesis from the six BIL lines and parents lines were extracted for Affymetrix Cotton GeneChips. Through the transcriptome analysis among these lines, 1508 differentially expressed (DE) genes were identified based on a two-fold difference or greater. Clustering analysis of these DE genes showed that the high yield BIL lines and the low yield BIL lines clustered into two distinct groups. The Blast-Mapping-Annotation-KEGG analysis by Blast2GO software was performed on the DE genes, and revealed that many genes were related to plasma membrane, oxidation reduction, response to stress and transport. 50 DE genes were selected for further analysis by RT-PCR analysis, and many fiber- and yield-related genes were identified, which would make a firm foundation for further genetic engineering and molecular breeding studies.

Key words: cotton; yield; gene chip; *GhMKRP2*; *GhPOD*

随着分子生物学技术的不断发展, 利用高通量芯片技术筛选和克隆产量相关基因已经成为分子育种研究的热点。目前, 在水稻中已经克隆得到了一系列与产量性状相关的基因 *Ghd7*^[1]、

DEP1^[2]和 *IPA1* 等^[3]。在玉米中, 也发现并验证了一些与花序、抽穗等产量性状相关基因, 例如 *RAMOSA*^[4]、*REL2*^[5]和 *Sos1*^[6]等。同时, 育种家也对小麦杂种优势的利用和高产相关基因之间的

收稿日期: 2013-02-25

作者简介: 张利媛 (1985-), 女, 硕士研究生, 工程师, zhangliyuan0225@163.com; * 通讯作者, yu@cricaas.com.cn

基金项目: 国家 863 计划 (2013AA102601-01-14); 国家科技支撑计划 (2011BAD35B05-1-6); 国家 973 计划 (2010CB126006)

互作关系进行了深入探索^[7-10]。这些研究为提高作物产量并进一步应用于大田生产打下了良好的分子基础。由于棉花基因组的复杂性,对其产量性状的研究还停留在 QTL (Quantitative trait locus) 定位上。Shen 等^[11]在陆地棉的重组自交系 (Recombinant inbred lines, RIL) 群体中定位到了 7 个对产量性状有重要作用的 QTL 位点。王志伟等^[12]采用改良的加性-显性(AD)模型对棉花主要性状的遗传效应进行了正反交比较研究,发现子棉产量、皮棉产量、衣分、铃数等产量相关性状均存在显著或极显著的 QTL 加性效应(A1)和显性效应(D1)。Wang 等^[13]通过对 180 个 RILs 遗传作图分析,检测到与产量性状相关的 34 个 QTLs。此外,还有许多其他对产量及其相关性状 QTL 定位的报道^[14-15]。

基因芯片作为一种高效的筛选差异表达基因的方法,已经在棉花中得到广泛应用。朱一超等^[16]利用 cDNA 芯片分析短纤维突变型和野生型纤维发育 4 DPA(Days post anthesis)的基因差异表达情况,得到 15 个差异表达的 ESTs(Expressed sequence tag),并进行了深入分析。Shi 等^[17]利用基因芯片技术筛选到 3 个在纤维细胞的伸长过程中起关键作用的乙烯合成酶基因。Chaudhary 等^[18]通过对海岛棉野生种及驯化种不同时期的纤维 cDNA 芯片研究发现,过氧化物酶基因的表达演化可能影响物种的驯化过程。但是,目前未见利用基因芯片对棉花产量相关基因的研究报道。本研究通过对产量差异较大的 2 组陆海回交近交系棉花进行芯片分析,筛选到一批与棉花产量性状相关的基因,为棉花的高产育种工作奠定了基础。

1 材料和方法

1.1 试验材料

陆地棉 SG 747 和海岛棉 Giza 75 杂交,经连续回交 2 代和自交 5 代后获得了高代回交近交系。根据纤维长度、比强度、马克隆值等性状,从中选择 17 个回交近交系,种植于中国农业科学院棉花研究所试验基地。统计 17 个品系 2006—2008 年间的皮棉产量,并利用 SAS 软件对数据进行 Duncan Grouping 显著性分析。开花当天对

棉铃进行挂牌标记,从长势良好的植株上取 5 DPA、10 DPA、15 DPA、20 DPA、25 DPA 棉铃,小心地剥取棉纤维,并迅速浸入液氮冷冻,保存在 -80 °C 冰箱中备用。

1.2 RNA 的提取和质量检测

利用 CTAB/ 酸酚法提取棉纤维样品总 RNA^[19],并用琼脂糖凝胶电泳检测总 RNA 的完整性及纯度,并利用 DU800 检测 RNA 质量和浓度。

1.3 基因芯片的制备及数据的获得

将 RNA 反转录成 cDNA,与 Affymetrix(<http://www.affymetrix.com>) 的棉花基因组表达谱芯片杂交(由上海晶泰生物技术有限公司完成):首先,利用 GeneArray Scanner 图像分析软件对芯片荧光信号进行图像扫描,然后利用 GeneChip Operating Software (GCOS) 软件对芯片数据进行分析,将图像信息转化为数字信息,最后用 MAS 5.0 对芯片数据进行均一化处理得到数据。

1.4 差异表达基因的筛选及功能分类

选用 Partek 和 R 软件比较高产组和低产组的基因芯片数据,根据比对结果筛选差异表达基因。并利用 Blast2GO 软件对所选择出来的差异表达基因进行 Blast-Mapping-Annotation-KEGG 和 GO 分析。

1.5 RT-PCR 和 qRT-PCR

利用 PrimeScript RT reagent Kit (Perfect Real Time)反转录试剂盒将 RNA 合成 cDNA(宝生物工程(大连)有限公司)。从差异表达基因中筛选上调基因 26 个、下调基因 24 个,使用 Oligo 6.0 软件设计引物(表 1),以看家基因 18S rRNA 为内参对照,利用 Premix Taq 酶(the Premix Ex Taq Version 2.0)进行 RT-PCR(宝生物工程(大连)有限公司)。以 18S rRNA 作为内参对照,利用荧光定量 PCR 仪(Bio-Rad 公司)进行 qRT-PCR 分析。所用试剂盒为 SYBR Premix Ex Taq II (Perfect Real Time)(宝生物工程(大连)有限公司),利用 CFX Manager 1.6 软件进行数据分析。

2 结果与分析

2.1 田间数据分析结果和芯片制备

根据回交自交系后代的产量性状,将材料分

表 1 RT-PCR 所选基因及其引物序列

Table 1 Primers of the selected genes

Probe set	Accession number	Gene name		Sequence
		18S rRNA	F	5'-AACCAAACATCTCACGACAC-3'
			R	5'-GCAAGACCGAAACTCAAAG-3'
GhiAffx.52906.1.S1_at	DW224876.1	<i>GhFAH</i>	F	5'-ACCCTTCCAAAGACCAAAGCAGA-3'
			R	5'-CCTCACAATCCCAACCACCATTA-3'
Ghi.10183.1.S1_at	DT562914	<i>GhMKRP2</i>	F	5'-TCCAGCAGGAGGAATAATGAGC-3'
			R	5'-TCCACCAGGCGTTTCAGGGTAC-3'
Gra.2682.2.S1_s_at	CO092077	<i>GhPAL</i>	F	5'-TTTGGTGCTACTTCTCATCG-3'
			R	5'-TCTAATGCCAGAGTATCCTTGT-3'
Ghi.319.1.S1_at	DT548678	Unknown	F	5'-AGCACAGTTTCAGGCTGGAATAA-3'
			R	5'-GGGTAAGCAGAAGATGATGGGTG-3'
GraAffx.29883.2.S1_s_at	CO127441	<i>GhAPGL</i>	F	5'-CTGTTCCAGTCGGAGGATGTTA-3'
			R	5'-ATATGTACGAGCAAGGTGACGA-3'
Ghi.10647.1.S1_s_at	DV849739	<i>GhRPL16</i>	F	5'-GGAGCTGGATGAGAAGAAACT-3'
			R	5'-GCTTGAAGAGCATATCTACCG-3'
Ghi.1171.1.S1_s_at	DT458994	Unknown	F	5'-AGCCACATTACCACATCCTTAT-3'
			R	5'-TGAAACAGAACTGATCCAGCA-3'
?Ghi.8388.1.S1_x_at	CD486400	<i>GhATLP</i>	F	5'-AAGCGAAACCCAAGTACCATT-3'
			R	5'-GTCACAGTTCCTTTGCCATT-3'
GhiAffx.3823.1.S1_at	DW232575.1	Unknown	F	5'-CTGCCGTTTGGTTACATTGAGA-3'
			R	5'-TGGAATACATTGGGATGGAGAA-3'
Ghi.5630.1.S1_s_at	DT047105	<i>GhGBF</i>	F	5'-AGTGGTGAAGTTGATGGTGGTT-3'
			R	5'-AGGCGGTCTAGGATGAGGAA-3'
Ghi.9158.1.S1_at	DT462880	<i>GhmetRS</i>	F	5'-CTGTCCAGATGGAGGGTGTA-3'
			R	5'-GAAGCTCAAGAAATAAGTGCC-3'
Ghi.2089.1.S1_at	AI727311	<i>GhMAP</i>	F	5'-TACCAGCGAAAGTTCAGGAG-3'
			R	5'-GCCATTGTTTTACGGGCATCA-3'
GhiAffx.25988.1.S1_s_at	DW519529.1	Unknown	F	5'-CGGCGAACACGACTAGAACAA-3'
			R	5'-CGAAGAATCCGACCACCTCAA-3'
Ghi.89.1.S1_x_at	CD486576	<i>GhXYL</i>	F	5'-AGCCAAAGGCAACCTCTATCTCG-3'
			R	5'-CACTTCCACTCAATCCCAACT-3'
Ghi.3081.1.S1_s_at	DT527367	<i>GhALDH</i>	F	5'-GGTATCCACCCTGCAATGCTA-3'
			R	5'-CAAACCCTGAAATGGGAAATGA-3'
Ghi.3212.1.A1_at	DT467489	<i>GhPOD</i>	F	5'-TGAGGGAGAAATGATGGGTTGA-3'
			R	5'-AAATTGCCTGGCTTTACGGACT-3'
Ghi.885.1.S1_s_at	AI728693	Unknown	F	5'-GCTCATCACCCAAAGAATCGG-3'
			R	5'-GCTCTTCGCAGTGTTCAGT-3'
Gra.636.2.S1_s_at	CO089538	<i>GhTFP</i>	F	5'-CGATGGGACTTCATTCTGGCACT-3'
			R	5'-AATCCTTTCCCTCAACGGCTC-3'
Ghi.5410.1.S1_s_at	DT049778	Unknown	F	5'-CAAACGTGGGACAGTATGAC-3'
			R	5'-CATCGGAGCTAAGGATTGGT-3'
Gra.2239.1.S1_s_at	CO092078	<i>GhPAL</i>	F	5'-GAATGCGTGTTTCGACAAGGAC-3'
			R	5'-CTCCACCCATCGAGATTCTGC-3'
Ghi.750.1.A1_s_at	DR455227	Unknown	F	5'-TATCCTCCTCTTCTCACCTTC-3'
			R	5'-ATGCCAGTTCTCCACTCAACC-3'
Ghi.1076.2.S1_s_at	DT459205	Unknown	F	5'-GAATCTTCATACCTTCCCTCAT-3'
			R	5'-TGCTGCCAACATTATACCAAC-3'
GhiAffx.53145.1.A1_s_at	DW502919.1	<i>GhCAX</i>	F	5'-GCCTGGCGGAAAGCTGACAAT-3'
			R	5'-TACGGGCATAACTGAGAAACA-3'
Gra.1083.1.A1_at	CO121246	<i>GhAPL</i>	F	5'-AGTCTGCTCCCAACATTACCG-3'
			R	5'-CAAATGCAGGATTAAGATGC-3'
Ghi.659.1.S1_at	DR456379	<i>GhAGPL</i>	F	5'-TACACCCTACCGCCGAAATCC-3'
			R	5'-ATAGCCGCAAGCCATCACCAC-3'
Ghi.732.1.S1_at	DR455437	Unknown	F	5'-GGACGATTGAGAATATGACCG-3'
			R	5'-ACCAAAGCCGATTGATAAGA-3'

续表 1 Continued table 1

Probe set	Accession number	Gene name	Sequence	
Ghi.6204.1.A1_x_at	CO493382	Unknown	F	5'-CCAAGCACTCTATACCCTACTATGA-3'
			R	5'-TAGGTTGTATGGGAGTTCGTTTC-3'
GhiAffx.26845.1.S1_at	DW229234.1	α -Gh EXP4	F	5'-TCTTTCCCGTGAAGGTTTGG-3'
			R	5'-TTTGGTGGGTCAGGCACTCT-3'
Ghi.6102.1.S1_at	DR458734	Unknown	F	5'-TGCTCGCAAGGGATTGAAGT-3'
			R	5'-TGCCTGGCTTTCCCATGTAC-3'
GhiAffx.4305.1.S1_at	DW225591.1	Unknown	F	5'-CAGGCGCAAATAGCAAACAAC-3'
			R	5'-CCCAAACATCGGATACTCACA-3'
Ghi.4418.1.S1_s_at	DR176775	GhFLA	F	5'-AGGACTCGAAACCAAACCA-3'
			R	5'-ATGCGACGGTGGAGAATAG-3'
Ghi.5892.1.A1_at	CO497111.1	Unknown	F	5'-CACTCAAATCACTCGCTCACAT-3'
			R	5'-GTGTTGTTTCGGGAAGAAGACG-3'
GhiAffx.41577.1.S1_s_at	DW503697.1	GhLMWHSP	F	5'-GGCTGATCTTCCAGGGCTTAA-3'
			R	5'-TGCCAGGTGTCGTTCTTGCT-3'
GraAffx.23958.2.S1_s_at	CO126474	Unknown	F	5'-GAGATTTGAAATATGCCCATG-3'
			R	5'-AACCCCTTAGTTTGCTCTTG-3'

成高产和低产 2 组(高产组:SG747,NMGA-017,NMGA-144,NMGA-026; 低产组:NMGA-023,NMGA-066,NMGA-095,Giza75), 经 Duncan

极差分析显示 2 组材料的产量性状存在显著性差异(表 2)。然后,取 8 个材料 10 DPA 的纤维样品,并提取总 RNA 点制芯片。

表 2 产量性状的方差分析

Table 2 The variance analysis of yield trait

材料 Genotype	Giza 75	NMGA-066	NMGA-095	NMGA-023	NMGA-144	NMGA-017	NMGA-026	SG747
皮棉产量 Lint yield/(kg·hm ²)	258.0	352.5	358.5	403.5	795.0	800.0	888.0	988.5
t _{0.05}	G	FG	EFG	EFG	ABC	AB	AB	A

注: t_{0.05}, 在 0.05 水平上进行 Duncan Grouping 聚类分析。

Note: t_{0.05}, Duncan Grouping at 0.05 level.

2.2 基因芯片数据的可靠性验证

李龙云等^[20-21]曾对基因芯片数据进行可靠性验证。其验证结果表明基因芯片的数据是可靠的。

2.3 芯片比对

根据 ratio>2.0 和 ratio<0.5 的标准, 采用 Partek 软件筛选得到 1508 个差异表达基因。对所筛选出来的差异表达基因进行聚类分析(图 1)显示: 2 个产量较高的系聚在一起, 2 个产量较低的系聚在一起, 而产量介于中间的 2 个材料聚在一起。所以, 根据材料的田间产量数据筛选与产量相关差异表达基因的方法是可靠的。

2.4 差异基因的功能分析

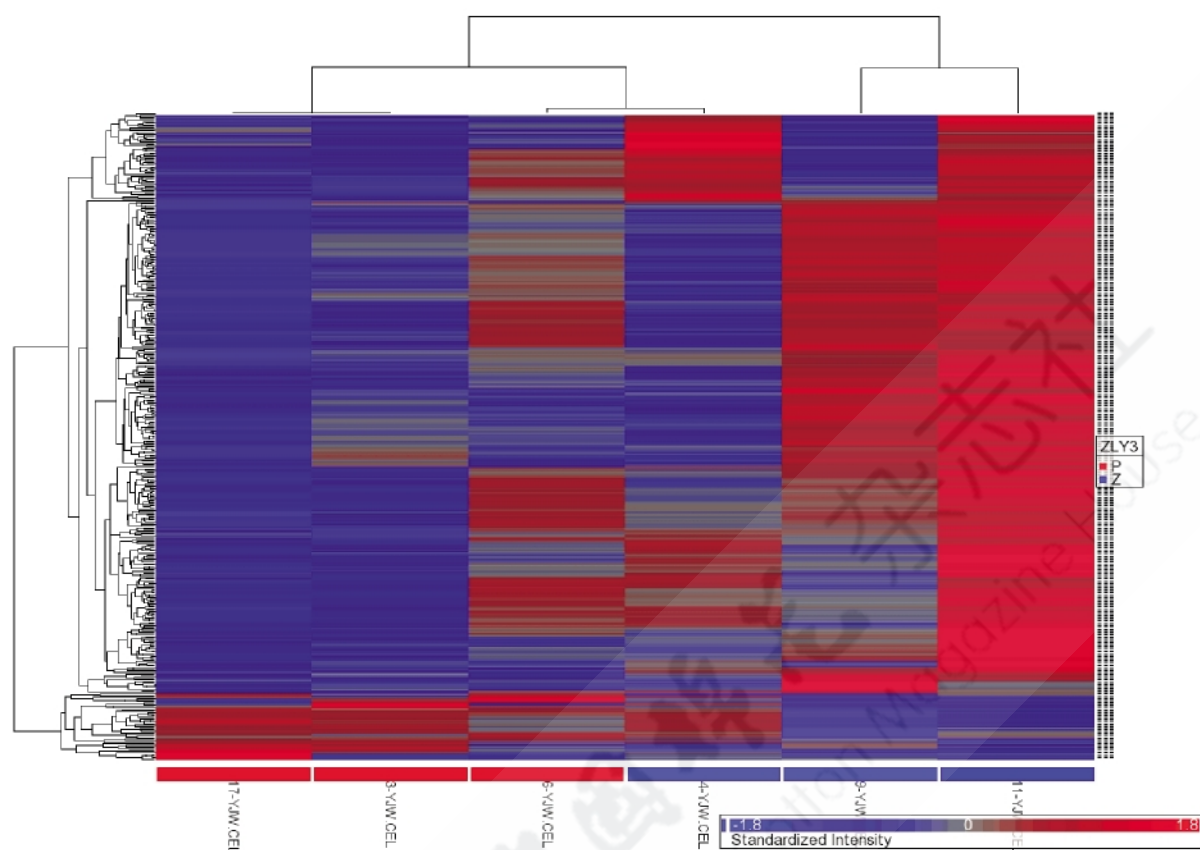
利用 Blast2GO 软件, 对 1508 个差异表达基因进行 Blast-Mapping-Annotation-KEGG 分析, 根据分析结果分别从生物学过程、细胞组分和分子功能 3 个方面对这些基因进行分析(图 2-图 4)。

从图 2 可以看出, 在 cutoff=40.0 的水平上,

这些基因共被分为 21 个不同的生物学过程, 其中参与逆境反应过程的基因最多, 有 143 个; 其次是运输过程, 有 89 个基因。植物抗逆生长和营养物质的运输是每个作物各个生育时期都必需的生物学过程, 通过分析发现其在维持纤维发育方面也同样起重要作用。

图 3 表明, 在 cutoff=20.0 的水平上, 这些基因被分到 24 个不同的分子功能种类中, 其中具有 ATP 结合活性、氧化还原活性和转录因子活性的基因最多, 这些功能基因可能与产量性状密切相关。

图 4 显示, 在 cutoff=20.0 水平上, 这些基因被归为 18 个不同的细胞组分, 其中与质体膜组分、线粒体膜组分、液泡膜组分相关的基因最多。由于决定棉花皮棉产量主要因素就是纤维细胞壁的质膜发育, 这些基因也可能与产量性状密切相关。



注:红色:上调基因;蓝色:下调基因。Red: up regulated genes; Blue: down regulated genes.

图1 差异基因聚类

Fig. 1 Clustering of the differential genes

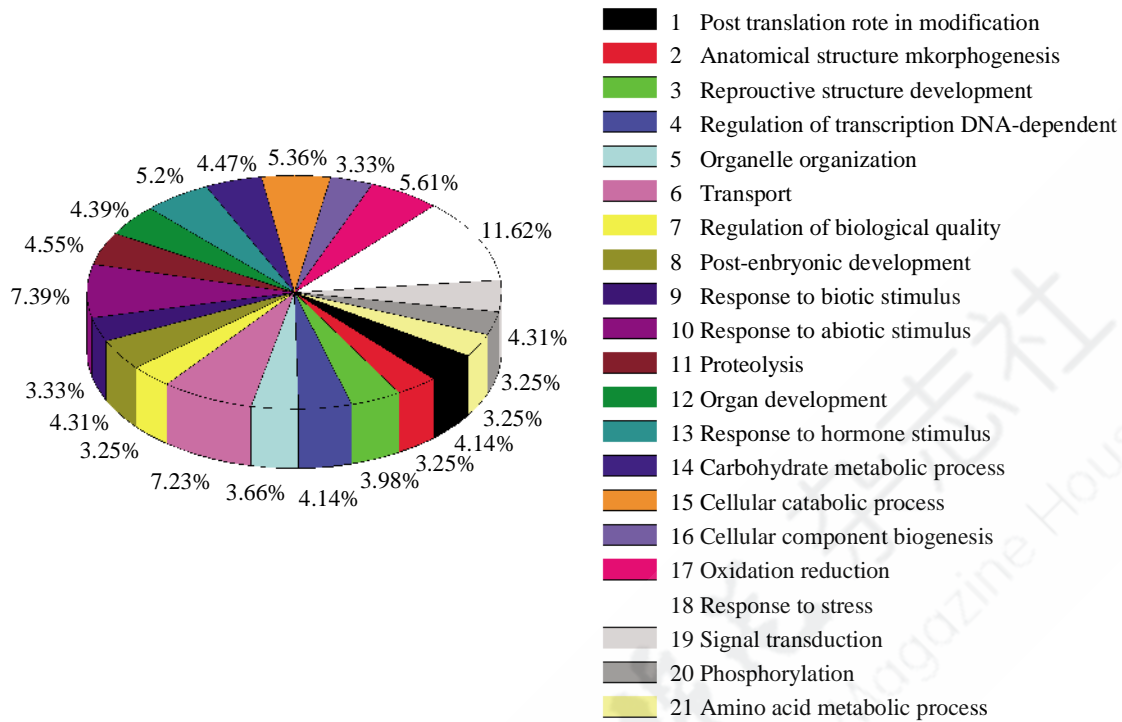
2.5 RT-PCR 分析

根据基因差异表达的显著性分析,结合 Blast2GO 分析和文献报道,筛选到 50 个纤维发育相关基因。并提取高产组、低产组和 2 个亲本共 8 个材料的 10 DPA、15 DPA、20 DPA、25 DPA 纤维的 RNA 进行 RT-PCR 检测。从图 6 可以看出, *GhMKRP2* 基因在纤维发育前期变化规律不明显,而纤维发育中后期在高产材料及陆地棉亲本中的表达量明显比低产材料和海岛棉亲本中高,推测该基因可能与纤维产量性状密切相关。*GhPOD* 基因在纤维发育早期的部分高产材料中表达量较高,推测其在纤维发育前期的高调表达很可能会提高产量。而其他的 6 个基因,在纤维发育前期的表达情况没有明显的变化规律,但在纤维发育后期,各自在不同材料中的表达量与产量性状呈负相关,因此推测这些基因的表达可能抑制棉纤维次生壁纤维素的沉积,进而影响棉花产量的形成。

3 讨论

目前对产量相关基因的研究已经成为各个作物研究热点,本研究以多年皮棉产量数据为依据,结合表达芯片数据,筛选到 1508 个差异表达基因。结合 Blast2GO 的结果和已报道的纤维发育相关基因,选取 50 个基因进行表达模式的验证,结果表明这些基因与产量性状密切相关。这些差异表达基因的挖掘,为以后棉花高产基因工程和分子育种提供了丰富的基因资源。

ROS (Reactive oxygen species)(活性氧自由基)在棉纤维纵向伸长过程是不可缺少的,而过氧化物酶在棉纤维伸长过程中对 ROS 的调节发挥着重要的作用^[2]。*GhPOD* 基因作为一个过氧化物酶基因,在清除活性氧的过程中发挥作用。同时, Blast2GO 结果也证明,该基因参与过氧化氢的异化过程和氧化还原等生物学过程。本研究表明 *GhPOD* 基因在高产组 10 DPA 棉纤维快速伸长期表现上调,推测该基因在棉纤维伸长时期的



注:1. Post translation rote in modification: 翻译后修饰作用;2. Anatomical structure mkorphogenesis: 形态发生结构;3. Reprouctive structure development: 生殖结构发育;4. Regulation of transcription DNA-dependent: DNA 依赖的调节转录;5. Organelle organization: 器官组织;6. Transport: 运输;7. Regulation of biological quality: 生物质量调控;8. Post-embryonic development: 后胚胎发育;9. Response to biotic stimulus: 对生物刺激的反应;10. Response to abiotic stimulus: 对非生物刺激的反应;11. Proteolysis: 蛋白质水解;12. Organ development: 器官发生;13. Response to hormone stimulus: 对激素刺激的反应;14. Carbohydrate metabolic process: 碳水化合物代谢过程;15. Cellular catabolic process: 细胞降解过程;16. Cellular component biogenesis: 细胞组分起源;17. Oxidation reduction: 氧化还原;18. Response to stress: 对压力的反应;19. Signal transduction: 信号转导;20. Phosphorylation: 磷酸化;21. Amino acid metabolic process: 氨基酸代谢过程。

图 2 差异基因在生物学过程上的聚类

Fig. 2 Clustering of the differential expressed genes based on biological processes

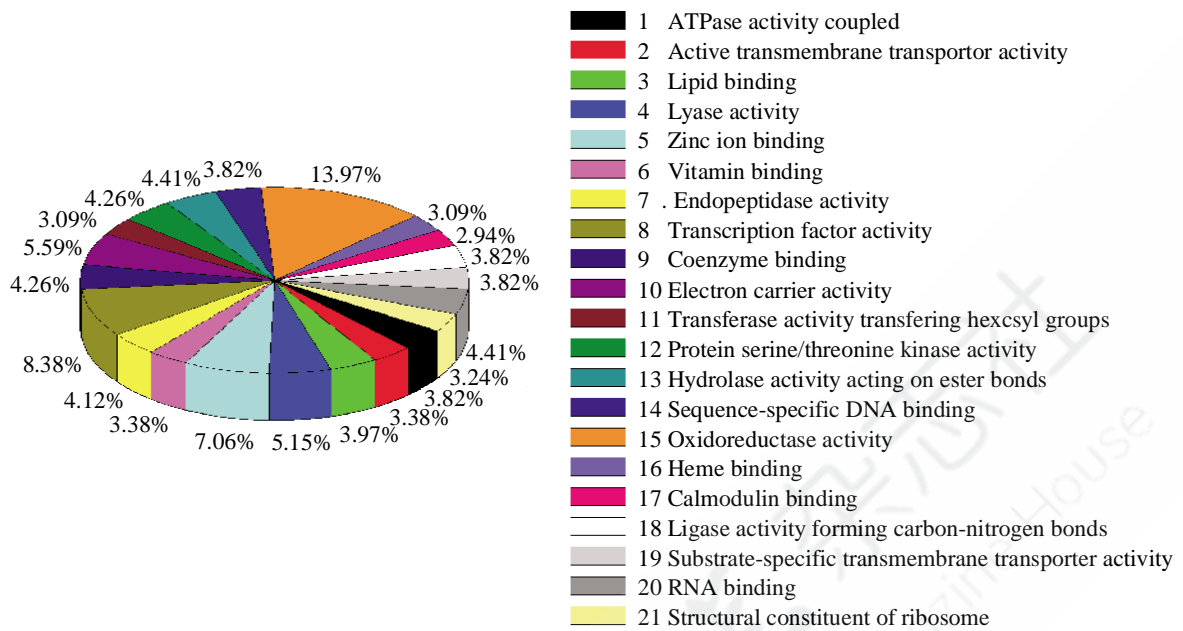
清除活性氧过程中起重要作用,并通过促进棉纤维的伸长来最终促进棉产量的提高。

Expansin 蛋白是非疏水性蛋白,可以通过断裂细胞壁结构大分子之间的非共价键而使细胞壁松弛,从而促进纤维细胞的扩展。Wilkins 等^[23]通过 Microarray 比较分析纤维正常基因型和短粗纤维基因型的纤维基因表达谱,也发现 expansin 表达水平与纤维长度呈正相关。Harmer 等^[24]也从棉花中分离了 6 个膨胀素基因 (*GhExp1*—*GhExp6*)。其中 *GhExp1* 和 *GhExp2* 在纤维细胞中特异性表达。因此,作者认为 *GhExp1* 可能是调节纤维细胞伸长的关键基因。本研究发现 α -*Gh-EXP4* 基因在棉纤维发育前期 (10 DPA 和 15 DPA) 即伸长期高量表达,同时 Blast2GO 结果也显示该基因参与植物细胞壁组织的形成,推测该

基因可能与纤维伸长密切相关。

另外,2 组材料中 *GhRPL16* 和 *GhAPL* 基因在棉纤维发育前期没有明显变化,而到后期 2 个基因在低产材料中的表达量明显高于高产材料。根据 Blast2GO 的分析结果推测:*GhRPL16* 蛋白是核糖体蛋白的一部分,也是构成线粒体功能的重要组成部分,*GhAPL* 参与糖元的生物合成途径,并具有葡萄糖腺苷转移酶的活性,推测其在棉纤维次生壁合成过程中发挥着重要作用,并影响着皮棉产量的形成。

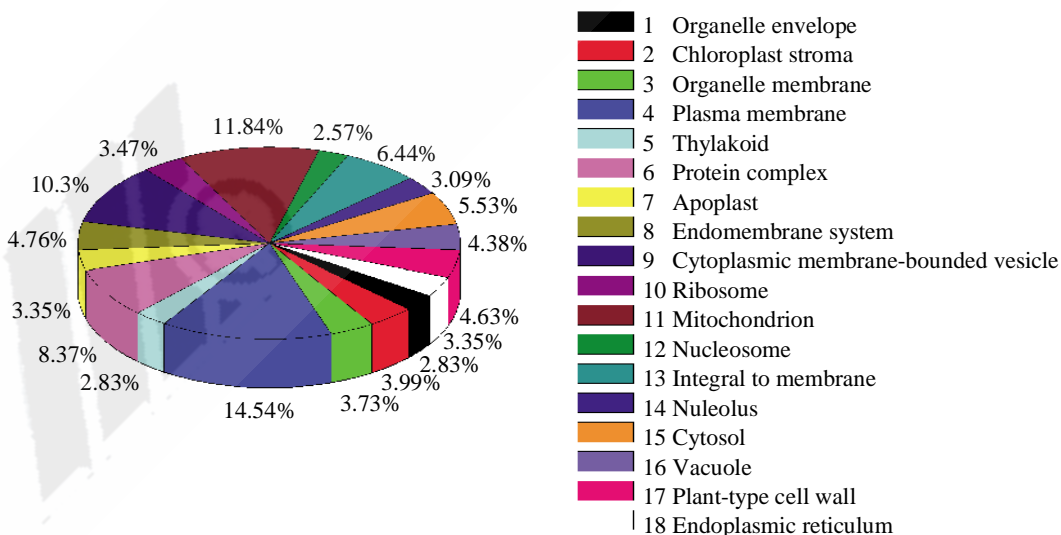
结合高代回交自交系田间数据,进而分析其基因芯片表达谱,从基因表达水平上对相关的差异表达基因进行研究,发现大量皮棉产量相关基因,为进一步从分子水平研究高产育种打下了良好的基础。



注:1. ATPase activity coupled:ATP 酶活性耦合;2. Active transmembrane transportor activity: 主动跨膜转运;3. Lipid binding:脂质结合;4. Lyase activity:裂合酶活性;5. Zinc ion binding: 锌离子结合;6. Vitamin binding:维生素结合;7. Endopeptidase activity: 肽链内切酶活性;8. Transcription factor activity: 转录因子活性;9. Coenzyme binding: 辅酶绑定;10. Electron carrier activity:电子载体活动;11. Transferase activity transferring hexcsyl groups:转录酶活性团体输送;12. Protein serine/threonine kinase activity:蛋白质的丝氨酸/苏氨酸激酶活性;13. Hydrolase activity acting on ester bonds:水解酶活性作用于酯键;14. Sequence-specific DNA binding: 序列特定 DNA 结合;15. Oxidoreductase activity: 氧化还原酶活性;16. Heme binding:血红素结合;17. Calmodulin binding:钙调蛋白绑定;18. Ligase activity forming carbon-nitrogen bonds:碳-氮键连接酶活性;19. Substrate-specific transmembrane transportor activity: 底物特定的跨膜运输活动;20. RNA binding:RNA 绑定;21. Structural constituent of ribosome:核糖体的结构组分。

图 3 差异基因在分子功能上的聚类

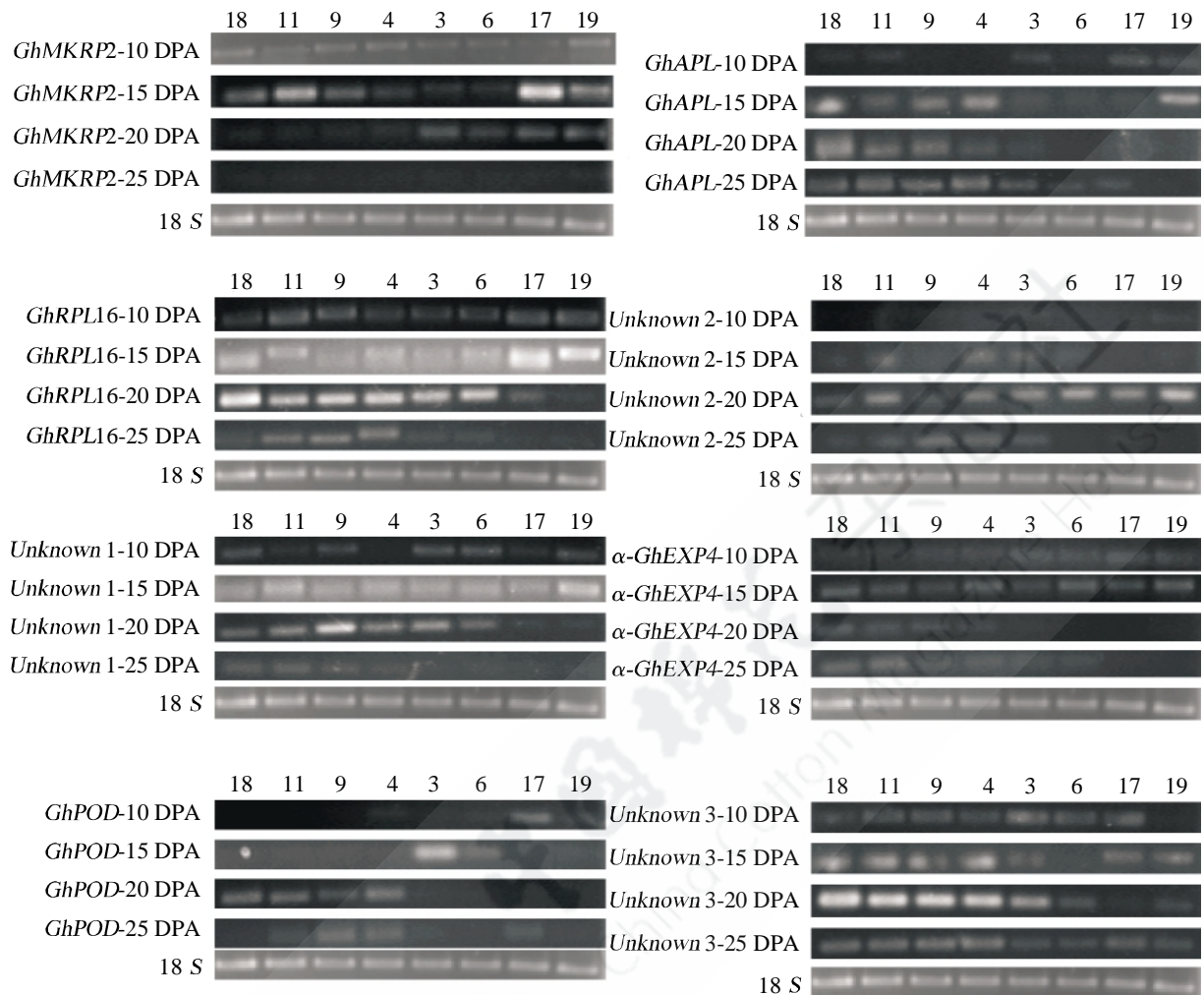
Fig. 3 Clustering of the differential expressed genes based on molecular functions



注:1. Organelle envelope:细胞器壁;2. Chloroplast stroma:叶绿体基质;3. Organelle membrane:细胞器膜;4. Plasma membrane:质膜;5. Thylakoid:类囊体;6. Protein complex:蛋白质复合体;7. Apoplast:质外体;8. Endomembrane system:内膜系统;9. Cytoplasmic membrane-bounded vesicle:细胞质膜泡;10. Ribosome:核糖体;11. Mitochondrion:线粒体;12. Nucleosome:核小体;13. Integral to membrane:整体膜分析;14. Nucleolus:细胞核;15. Cytosol:细胞溶质;16. Vacuole:液泡;17. Plant-type cell wall:株型细胞壁;18. Endoplasmic reticulum:内质网。

图 4 差异基因在细胞组分上的聚类

Fig. 4 Clustering of the differential expressed genes based on cellular components



19: recurrent parent upland cotton SG747; 18: male parent *Gossypium barbadense* Giza 75; 17: NMGA-026; 11: NMGA-066; 9: NMGA-095; 6: NMGA-017; 4: NMGA-023; 3: NMGA-144.

图5 8个基因在8个基因型不同发育时期纤维中表达的RT-PCR结果

Fig. 5 RT-PCR results of eight genes

参考文献:

- [1] XUE Wei-ya, Xing Yong-zhong, Weng Xiao-yu, et al. Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice[J]. *Nat Genet*, 2008, 40(6): 761-767.
- [2] HUANG Xian-zhong, Qian Qian, Liu Zheng-bin, et al. Natural variation at the *DEP1* locus enhances grain yield in rice[J]. *Nat Genet*, 2009, 41(4): 494-497.
- [3] JIAO Yong-qing, Wang Yong-hong, Xue Da-wei, et al. Regulation of *OsSPL14* by *OsmiR156* defines ideal plant architecture in rice[J]. *Nat Genet*, 2010, 42(6): 541-545.
- [4] PAULA M S. Branching out: the ramose pathway and the evolution of grass inflorescence morphology [J]. *Plant Cell*, 2006, 18: 518-522.
- [5] GALLAVOTTI A, Long J A, Stanfield S, et al. The control of axillary meristem fate in the maize *ramosa* pathway[J]. *Development*, 2010, 137: 2849-2856.
- [6] WU Xian-ting, Andrea S, Paula M S. Suppressor of sessile spikelets1 Functions in the ramose pathway controlling meristem determinacy in Maize[J]. *Plant Physiol*, 2009, 149: 205-219.
- [7] ULLAH S, Khan A S, Raza A et al. Gene Action analysis of yield and yield related traits in spring wheat (*Triticum aestivum*) [J]. *Intl J Agric Biol*, 2010, 12: 125-128.
- [8] DERE S, Yildirim M B. Inheritance of grain yield per plant, flag leaf width and length in an 8x8 diallel cross population of bread wheat (*T. aestivum* L.)[J]. *Turk J Agric*, 2006, 30: 339-345.
- [9] GORJANOVIC B, Balalic M K. Inheritance of plant height and spike length in wheat[J]. *Genetika*, 2005, 37: 25-31.
- [10] HASSAN S E, Khaliq I, Khan A S. Genetic mechanism of some physiological traits in spring wheat at two plant population regimes[J]. *J Agric Res*, 2008, 46: 395-401.
- [11] SHEN Xin-lian, Guo Wang-zhen, Lu Qiang-xian et al. Genetic mapping of quantitative trait loci for fiber quality and yield trait

- by RIL approach in upland cotton [J]. *Euphytica*, 2007, 155: 371-380.
- [12] 王志伟, 王清连, 张永山. QTL 遗传效应正反交差异研究 [J]. *棉花学报*, 2007, 19(3): 194-198.
WANG Zhi-wei, Wang Qing-lian, Zhang Yong-shan. Difference of genetic effects due to QTL between reciprocal crosses in transgenic cotton [J]. *Cotton Science*, 2007, 19(3): 194-198.
- [13] WANG Bao-hua, Gao Wang-zhen, Zhu Xie-fei et al. QTL mapping of yield and yield components for elite hybrid derived-RILs in upland cotton [J]. *J Genet Genom*, 2007, 34: 35-45.
- [14] CHUANFU A, Johnie N J, Wu J X, et al. Use of fiber and fuzz mutants to detect QTL for yield components, seed, and fiber traits of upland cotton [J]. *Euphytica*, 2010, 172: 21-34.
- [15] 秦永生, 刘任重, 梅鸿猷, 等. 陆地棉产量相关性状的 QTL 定位 [J]. *作物学报*, 2009, 35(10): 1812-1821.
QIN Yong-sheng, Liu Ren-zhong, Mei Hong-xian, et al. QTL mapping for yield traits in upland cotton (*Gossypium hirsutum* L.) [J]. *Acta Agron Sin*, 2009, 35(10): 1812-1821.
- [16] 朱一超, 张天真, 贺亚军, 等. 棉花纤维伸长发育期的基因表达分析 [J]. *作物学报*, 2006, 32(11): 1656-1662.
ZHU Yi-chao, Zhang Tian-zhen, He Ya-jun, et al. Gene expression analysis during the fiber elongation period in cotton (*Gossypium hirsutum* L.) [J]. *Acta Agron Sin*, 2006, 32(11): 1656-1662.
- [17] SHI Yong-hui, Zhu Sheng-wei, Mao Xi-zeng et al. Transcriptome profiling, molecular biological and physiological studies reveal a major role for ethylene in cotton fiber cell elongation [J]. *Plant Cell*, 2006, 18: 651-664.
- [18] CHAUCHARY B, Hovav R, Rapp R, et al. Global analysis of gene expression in cotton fibers from wild and domesticated *Gossypium barbadense* [J]. *Evol Dev*, 2008, 10: 567-582.
- [19] 蒋建雄, 张天真. 利用 CTAB/ 酚法提取棉花组织总 RNA [J]. *棉花学报*, 2003, 15(3): 166-167.
JIANG Jian-xiong, Zhang Tian-zhen. Extraction of total RNA in cotton tissues with CTAB-acidic phenolic method [J]. *Cotton Science*, 2003, 15(3): 166-167.
- [20] 李龙云, 于雯雯, 翟红红, 等. 棉花纤维发育相关基因表达谱的比较分析 [J]. *分子植物育种*, 2010, 8(3): 488-496.
LI Long-yun, Yu Ji-wen, Zhai Hong-hong et al. Comparative analysis of cotton fiber development related gene expression profiling [J]. *Mol Plant Breed*, 2010, 8(3): 488-496.
- [21] 李龙云, 于雯雯, 翟红红, 等. 利用基因芯片技术筛选棉纤维伸长相关基因 [J]. *作物学报*, 2011, 37(1): 95-104.
LI Long-yun, Yu Ji-wen, Zhai Hong-hong, et al. Identification of fiber length-related genes using cotton oligonucleotide microarrays [J]. *Acta Agron Sin*, 2011, 37(1): 95-104.
- [22] MEI Wen-qing, Qin Yong-mei, Song Wen-qiang, et al. Cotton *GhPOX1* encoding plant class III peroxidase may be responsible for the high level of reactive oxygen species production that is related to cotton fiber elongation [J]. *J Genet Genom*, 2009, 36: 141-150.
- [23] WILKINS T A, Arpat A B. The cotton fiber transcriptome [J]. *Physiol Plant*, 2005, 124: 295-300.
- [24] HARMER S E, Orford S J, Timmis J N. Characterisation of six alpha-expansin genes in *Gossypium hirsutum* (upland cotton) [J]. *Mol Genet Genom*, 2002, 268(1): 1-9. ●