

杉木杂交组合内生长性状分离的机制探讨

王海蓉¹, 齐明^{2,3}, 叶金俊¹, 何贵平^{2,3}

(1. 浙江省遂昌县林业技术推广总站, 浙江 遂昌 323300; 2. 中国林业科学研究院亚热带林业研究所, 浙江 杭州 311400;
3. 浙江省林木育种技术研究重点实验室, 浙江 杭州 311400)

摘要: 为研究杉木 *Cunninghamia lanceolata* 杂交组合内子代生长性状分离的分子机制, 以杉木杂种和亲本为研究对象, 从基因表达水平揭示超亲杂种优于低亲杂种的形成分子机制, 采用 Illumina HiSeq 4000 高通量测序技术对不同生长势的杉木杂种(龙 15 × 1339)的超高亲子代(HF1)和超低亲子代(LF2)及其亲本(龙 15 和 1339)进行无参转录组测序和差异比较。结果表明, 所有样本经转录组测序共产生 Clean reads 5.8E+08 条, 总拼接长度为 49 803 726nt, 将 Clean reads 在 6 个数据库(Nr, Swiss-prot, KOG, KEGG, Pfam, GO)进行 BLASTX 分析, 比对结果产生 80 171 个基因;杉木同一组合内的 HF1 比 LF2 的生产力高,是因为 HF1 在不同的 GO 和 KEGG terms 内和 terms 间两个层面上的差异表达基因的分布处于不均匀、不平衡状态; HF1 的 GO terms 和 KEEG terms 的差异基因的分布不均匀、不平衡; LF2 GO terms 和 KEEG terms 的差异基因的分布趋于均匀、平衡状态。利用分子生物学技术对育种群体展开遗传多样性研究, 选择表达基因互补的亲本进行杂交组配, 这样可以降低杂种间生长性状的分离, 以达到收获期大体一致的目的, 是一条可行的技术路线。

关键词: 杉木; 杂交组合; 生长性状分离; 转录组测序; 分子机理

中图分类号: S718.46 文献标识码: A 文章编号: 1001-3776(2020)03-0001-08

Mechanism of Segregation of Growth Traits in *Cunninghamia lanceolata* Cross Combination

WANG Hai-rong¹, QI Ming^{2,3}, YE Jin-jun¹, HE Gui-ping^{2,3}

(1. Suichang Forestry Extension Station of Zhejiang, Sui chang 323300, China; 2. Research Institute of Subtropical Forestry of CAF, Hangzhou 311400, China; 3. Zhejiang Provincial Key Laboratory of Tree Breeding, Hangzhou 311400, China)

Abstract: Using the latest second-generation *Cunninghamia lanceolata* hybrids (Long 15 × 1339), HF1(3 super-parent hybrids), LF2(3 low-parent hybrids) and parents (P1 and P2), all sample groups were compared by transcriptome sequencing. The result demonstrated that in a total of 12 samples, 5.8E+08 Clean reads were generated by transcriptome sequencing, and the total length was 49 803 726nt. BLASTX analysis was performed on Clean reads in six databases (Nr, Swiss-prot, KOG, KEGG, Pfam, GO), through sequence alignment which resulted in 80 171 unigenes. Analysis indicated that the productivity of *C.lanceolata* hybrids associated to the following factors, growth of progenies in cross combination segregated, growth of HF1 faster than that of LF2, which was due to the uneven and unbalanced distribution of expression unigene at two-level: among GO and KEEG terms and within GO and KEEG terms. The differential genes of GO and KEEG terms in high parental generation were not evenly distributed and balanced; The distribution of the difference genes of GO terms and KEEG terms in the low parental generation tends to be uniform and balanced.

收稿日期: 2019-09-12; 修回日期: 2020-03-31

基金项目: 浙江省“十三五”林木新品种育种专项(2016C02056-5); 中国林科院基金项目(195005); 遂昌县牛头山林场省级林木良种基地建设项目(SL2019B02)

作者简介: 王海蓉, 工程师, 从事林木良种培育工作; E-mail:398843232@qq.com。通信作者: 齐明, 副研究员, 从事杉木多性状遗传改良; E-mail:youqingyi1962@163.com。

Key words: *Cunninghamia lanceolata*; cross combination; growth traits segregation; transcriptome sequencing; molecular mechanism

杂交育种是杉木 *Cunninghamia lanceolata* 遗传改良的主要途径^[1]。在杉木杂交育种研究中,生长性状的杂种优势现象已有报道^[2],但对杂种的生长、材质等性状分离的分子机理研究的较少。杉木是我国重要的工业用材树种,按照短周期工业用材的要求,除了速生优质外,另外就是要收获期一致。杉木遗传改良进入了第三代,杉木育种群体是多世代与高强度选择的产物,杂合体居多。福建省已观察到杉木高世代种子园的后代生长明显分离,这明显不利于杉木工业用材林的收获。杉木无性系和亲本性状互补的亲本双系杂交种子园有着独有的优势,但是无性系对立地条件要求较高,且存在“C”效应(即年龄效应和位置效应)。杉木双系种子园有着独有的优势,但是这必须建立在优良杂交组合的子代分离研究的基础上。通过选择性状互补的优良亲本进行控制杂交,再经苗期评价,可以选育出生长节律大致相同的杂交新品种。

近年来,我国对一年生杂种实生苗分离状况的研究报道多在李^[3]*Prunus salicina*, 柚^[4]*Citrus maxima*, 枣^[5]*Ziziphus jujuba*, 苹果^[6-7]*Malus pumila* 和梨^[8-9]*Pyrus spp.* 等园艺植物上,研究多在1~3年生以上的生长势或亲本与童期的关系方面,结果发现,一年生植物枝干、叶片、针刺等性状可作为早期鉴定、早期选择的主要指标。在用材林研究上,邱有德等^[10]发现巨桉 *Eucalyptus grandis* 控制优质大径材的 *EgLBD* 基因及其作用机制,但杉木杂交组合的子代分离对收获期的影响尚未引起同行的注意。不少学者在杉木研究中借助 RNA-seq 技术来研究杉木纤维性状发育的形成机制和木材形成层活动的机理^[11-13]。

目前,尚未发现有人采用转录组技术研究杉木优良杂交组合内生长性状分离的分子机理。本研究在齐明等^[14-15]对杉木杂交子代试验林生长量前期试验调查的基础上,利用无参转录组测序技术,在基因差异研究的基础上,以基因差异表达分析为切入点,以杉木一个优良组合的超亲、低亲两组杂交子代及其亲本为研究材料,展开杂种子代和亲本两两间的比较分析,来探讨杉木杂种生长性状分离的分子机理,揭示杉木优良杂交组合高生产力的原因。

1 材料与方法

1.1 研究材料

研究材料取自浙江省遂昌县湖山乡大桥村的杉木杂交子代试验林,1996年重复制种,1997年育苗,1998年春营造试验林,试验设计为完全随机区组,19个处理,8次重复,纵向单行5株小区,每个家系40株参加试验,造林株行距为2m×2m。2005年11月进行生长量调查^[14-15]。2017年6月完成叶片取样,叶样来自样树顶部当年生嫩枝和针叶,在研究其遗传变异的基础上,选取表现最优的一个杂交组合(龙15×1339)中的超高亲子代(生长超过最优亲本的子代, HF1)和超低亲子代(生长低于中亲的子代, LF2),和同龄双亲(龙15和1339)。3个生物学重复,共12个样本,参与测序分析。

1.2 研究方法

1.2.1 遗传变异及生长性状的分离研究 本试验分析了研究材料的遗传变异、生长等性状的分离情况、以及优良组合龙15×1339的生产力,并将其与测序结果联系起来,进行研究。

1.2.2 cDNA 文库准备及 RNA-seq 测序 文库构建、无参转录组测序,以及随后的 unigene 功能注释、reads 富集与分类和基因表达等内容的分析方法参见文献^[16-20]。

1.2.3 序列比对及差异表达基因与杂种生产力间的关系分析 通过 Illumina Hiseq 4000 测序获得的转录组测序数据,需要经过生物信息学处理与分析^[16-20],方能获得有意义的结果。除了样本针叶的采集和研究方案的制定外,所有测序和初步分析(包括 GO 富集与分类、KEGG 的富集等分析)委托杭州联川生物技术股份有限公司完成。在此基础上,对注释的基因,GO 和 KEGG terms 富集结果,考查 terms 间和 terms 内基因的平衡状态,并追踪几个显著 KEGG terms 内,四个比较组间基因上调或下调规律,以揭示杂种生产力的高产和低产的原因。

unigene 的表达量 FPKM (Fragments per kb per Million fragments), 计算公式:

$$FPKM = 10^6 C / [(NL) / 10^3]$$

式中, FPKM 是某个基因 (A) 的表达量, C 是唯一比对到基因 A 的片段数, N 是唯一比对到所有 unigene 的总片段数, L 为 unigene A 的碱基数。

样本比较组的 GO-term 或 KEGG-term 间基因处于平衡或不平衡的状态指数 $k = \text{sum}(\text{第 } i \text{ 个下调 term 的基因数} / \text{第 } i \text{ 个上调 term 的基因数}) / n$, 上式中 i 取 $1 \sim n$, 标准差按常规公式计算。当 k 等于或接近于 1 时, 样本组间的差异表达基因处于平衡状态, 反之则处于不平衡状态。

2 结果与分析

2.1 研究材料的生长变异性和生长分离情况

对 9 年生时试验林生长性状的遗传变异进行调查分析^[14-15], 结果见表 1。

表 1 研究群体和优良杂交组合 (龙 15×1339) 9 年生的生长性状大小和分离信息^[14-15]
Table 1 Growth traits and segregation of *C. lanceolata* in a 9-year cross hybrid stand and cross combination 15×1339 in Suichang

项目	材积/dm ³	树高/m	胸径/cm	
杂交组合	龙 15×1339 平均值	19.69	5.75	8.35
	变异范围 (绝对值)	0 ~ 44.86	0 ~ 7.0	0 ~ 12.00
	表型 PCV/%	58.38	36.76	38.93
试验林	群体平均值	17.54	5.23	8.03
	变异范围 (绝对值)	0 ~ 68.11	0 ~ 9.1	0 ~ 13.30
	表型 PCV/%	62.69	20.20	24.43

注: PCV 表示变异系数。

由表 1 可知, 优良杂交组合的平均生长性状水平比群体的平均水平要高; 优良杂交组合的生长性状的变异幅度略小于全林试验结果, 这与参试子代样本数的多少有关 (组合样本数/全林子代样本数=36/602); 除材积的表型变异性, 群体水平略高于杂交组合外, 在树高、胸径上均为优良组合龙 15×1339 超过群体平均变异水平, 说明本试验中的优良组合龙 15×1339 中子代的分离变异还是较大的。

另外, HF1 的平均胸径是 LF2 胸径的 134.81%~138.32%; HF1 平均树高是 LF2 的 134.60%~150%。采样的子代植株、亲本龙 15 和 1339 生长发育均正常。

2.2 测序数据质量和基因表达概况

杉木 4 个处理 12 个样品, 测序测得的原始序列 (Raw reads) 介于 (4.3E+07)~(5.2E+07) nt; 12 个样本的 Clean reads 分别介于 (4.3E+07)~(5.2E+07) nt。Clean reads 有效数据占原始 Raw reads 的比例在 98%以上。二代测序中, 每测一个碱基会给出一个相应的质量值, 这个质量值是衡量测序准确度的。Q20 代表错误率为 1%; Q30 代表错误率为 0.1%。测序中的 Q20 与 Q30 则表示质量值≥20 或 30 的碱基所占百分比。12 个样本的 Phred 数值大于 Q20 和 Q30 的碱基占总体碱基的百分比, Q20 和 Q30 的百分比分别介于 98.20%~98.84%和 95.07%~96.4%。原始测序序列中, 碱基 G 和碱基 C 的数量总和占总碱基数的百分比 (GC%) 介于 44.20%~45.08%之间。

综合以上几个测序质量评价指标, 说明 12 份样品的测序质量较高, 能保证后续研究和满足后续数据分析的要求。

2.3 测序数据的拼接和基因注释结果

测序数据的拼接结果是, 平均 GC%达 40.78%; 当读长数达 50%时, 该读长的长度为 1 214nt。综合其它项的结果, 可以得出测序数据的组装拼接的结果很成功。

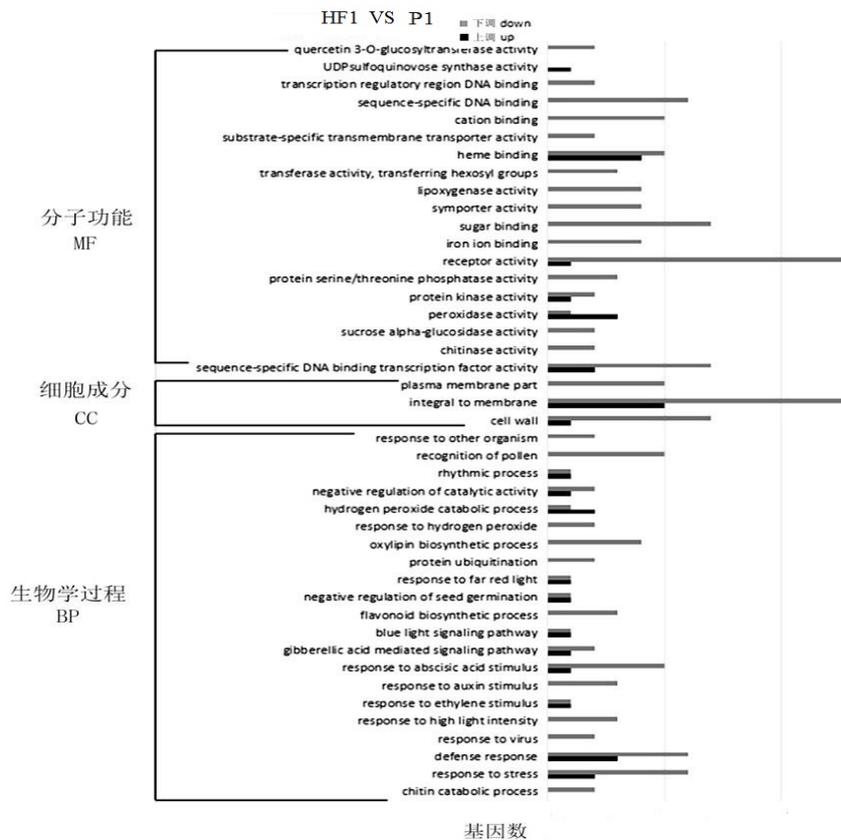
对 Clean reads 在六个数据库进行 BLASTX 分析, 将获得的基因分别在 Swiss-prot, Nr, Pfam, KEGG, KOG 和 GO 中注释, 结果杉木亲代和子代共注释有 80 171 个基因。

杉木为无参测序, 测得的数据通常采用 blastX 序列比对, 各数据库所注释总的基因比率大于 100%, 这表明有的基因同时在不同的数据库得到了重复注释。

测序数据质量、基因表达测序、数据的拼接和基因注释结果更详细的情况参见文献[15]。

2.4 差异表达基因的 GO 富集与分类分析和杂种子代的生产力

本研究选择 4 个样本组，HF1 VS P1（HF1 与龙 15 相比），HF1 VS P2（HF1 与 1339 相比），LF2 VS P1（LF2 与龙 15 相比），LF2 VS P2（LF2 与 1339 相比），对所有的 DEGs（differentially expressed genes）进行 GO 功能富集分析，从中挑选出基因表达量显著的 GO terms 进行 GO 分类分析作图。GO 功能分类体系中有参与生物过程（BP）、细胞组分（CC）以及分子功能（MF）3 个大类，43 或 45 个小类别。HF1 VS P1 中，GO 的功能富集，前 10 名的结果是：胁迫反应（ $k=3$ ）>糖结合（ $k=9/0$ ）>赤霉素介导的信号通路（ $k=2$ ）>受体活性（ $k=20$ ）>防御性反应（ $k=2$ ）>脱落酸活性（ $k=5$ ）>蛋白丝氨酸/苏氨酸磷酸酶活性（ $k=9/0$ ）>转录调控区域 DNA 结合（ $k=9/0$ ）>节律过程（ $k=1$ ）>过氧化物酶活性（ $k=0.33$ ）。这 10 个 GO terms 内，表达基因不均衡。



注：以上 GO terms 依次是 1-槲皮素 3-0-邻葡萄糖活性；2-UDP 磺基喹诺糖合酶活性；3-超越调控区 DNA 结合；4-特异系列 DNA 结合；5-阳离子结合；6-特异底物跨膜转运体活性；7-亚铁血红素结合；8-转移酶活性；9-脂氧合酶活性；10-同向转运活性；11-糖结合；12-铁离子结合；13-受体活性；14-丝氨酸蛋白磷酸酶活性；15-蛋白酶活性；16-过氧化物酶活性；17-蔗糖- α -葡萄糖苷活性；18-几丁质酶活性；19-序列特异性 DNA 结合转录因子活性；20-质膜部分；21-不可或缺的膜；22-细胞壁；23-对其它有机体的响应；24-识别标注；25-节律过程；26-负调节催化活性；27-过氧化氢酶催化过程；28-对过氧化氢酶的响应；29-氧的生物合成；30-蛋白质泛素化；31-对远红外光的响应；32-种子萌发负调控；33-类黄酮生物合成过程；34-蓝光信号传导途径；35-赤霉素介导的信号途径；36-对脱落酸刺激的响应；37-对生长素刺激的响应；38-对乙烯刺激的响应；39-对高强度光线的响应；40-对病毒的响应；41-防御响应；42-对胁迫的响应；43-几丁质分解过程。

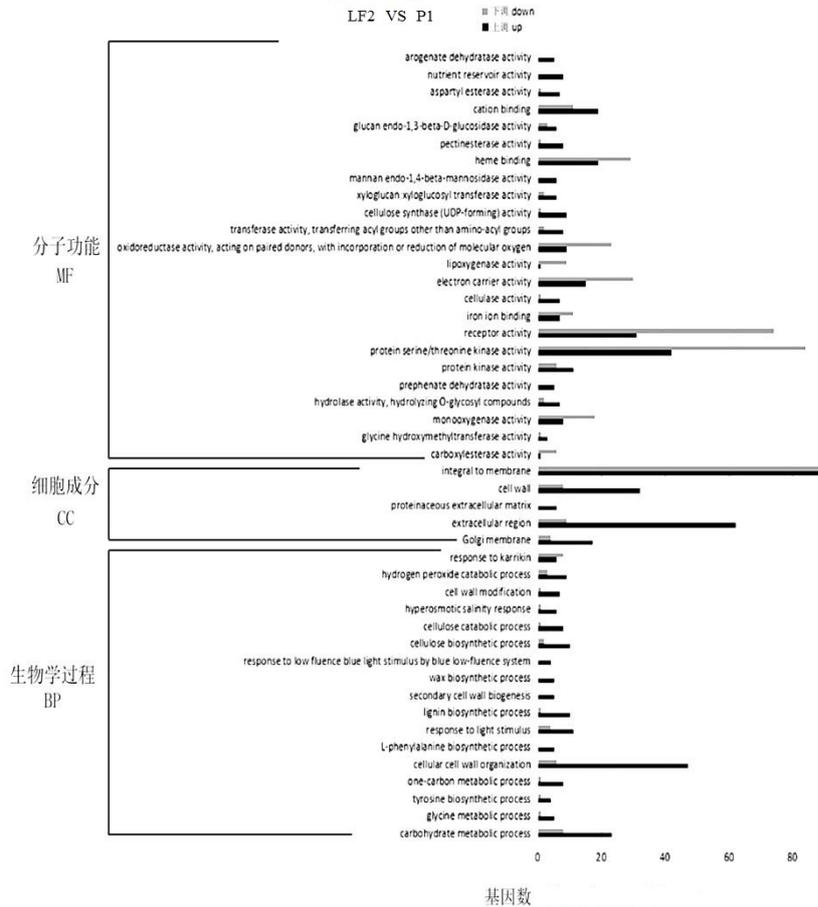
图 1 HF1 VS P1 差异表达基因的 GO 分类

Figure1 GO classification of differentially expressed genes in HF1 VS P1

比较组 LF2 VS P1 的 GO 的功能富集结果，前 10 名排列顺序为：细胞壁组织（ $k=0.1277$ ）>细胞外区域（ $k=0.1452$ ）>不可或缺的膜（ $k=1.2137$ ）>受体活性（ $k=2.3871$ ）>蛋白激酶活性（ $k=2$ ）>纤维素分解过程

($k=0.1429$) > 纤维素酶活性 ($k=0.1429$) > 碳水化合物代谢过程 ($k=0.3478$) > 转移酶活性、转移酰基以外的氨基酰基 ($k=0.4286$) > 纤维素生物合成的过 ($k=0.2$)。

图 2 是 LF2 VS P1 的 GO 富集与分类分析结果, 与图 1 一样, 仍以龙 15 (P1) 为参照物, LF2 VS P1 中不同的 GO terms 中参与新陈代谢的基因数目不同; 针对大多数 GO terms, 基因下调和上调有着相似的态势。同时, 在所有的 GO terms 上存在显著的上调基因, LF2 VS P1 中 GO terms 上基因分布趋于均匀、平衡。



注: 以上 GO terms 依次是 1-阿罗酸脱水酶活性; 2-营养库活性; 3-天门冬氨酸酯酶活性; 4-阳离子结合; 5-葡聚糖吡啶-1,3-β-D-葡糖苷酶活性; 6-果胶酯酶活性; 7-亚铁血红素结合; 8-甘露聚糖吡啶-1,4-β-甘露糖苷酶活性; 9-木葡聚糖转移酶活性; 10-纤维素合成酶活性; 11-转移酶活性; 12-氧化还原酶活性; 13-脂氧合酶活性; 14-电子载体活性; 15-纤维素酶活性; 16-铁离子结合; 17-受体活性; 18-蛋白质激酶活性; 19-预苯酸脱氢酶活性; 20-水解酶活性; 21-单氧酶活性; 22-甘氨酸羟甲基转移酶活性; 23-羧酸酯酶活性; 24-不可或缺的膜; 25-细胞壁; 26-非细胞质基的额外蛋白质; 27-细胞外区域; 28-高尔基膜; 29-对胰激肽的响应; 30-过氧化氢分解过程; 31-细胞壁的改变; 32-对高渗透盐度的响应; 33-纤维素分解代谢过程; 34-纤维素生物合成过程; 35-对低影响蓝光的响应; 36-蜡的生物合成过程; 37-次生壁的生物起源; 38-木质素的生物合成过程; 39-对光刺激的响应; 40-苯基丙氨酸生物合成过程; 41-细胞壁组织; 42-一碳代谢过程; 43-酪氨酸生物合成过程; 44-甘氨酸代谢过程; 45-碳水化合物代谢过程。

图 2 LF2 VS P1 差异表达基因的 GO 分类

Figure2 GO classification of differentially expressed genes in LF2 VS P1

比较图 1 与图 2 发现, HF1 VS P1 和 LF2 VS P1 的分析结果不一致, 其一表现在 GO terms 层面, 富集分析前 10 名的 GO terms 中, 仅有活体受性一个 GO term 相同, 其它 terms 均不相同; 其二表现在 GO terms 内上调基因和下调基因数目的不均衡性上, HF1 VS P1 的不平衡结果, 比 LF2 VS P1 比较组的结果更甚。在 HF1 VS P1

中, 差异表达基因在不同 GO terms 上分布处于不均匀、不平衡, 而 LF2 VS P1 中, 整个系统间差异表达的基因数相差不大, 平衡系数接近于 1.0。由此得出, HF1 生长优于 LF2 正是 HF1 的基因系统处于非均匀、非平衡状态; 而 LF2 生长慢则是由于 LF2 的基因系统趋向于均匀、平衡状态。这是杉木杂种生长分离的内在遗传基础。

本研究还研究了 HF1 VS P2 和 LF2 VS P2 比较组 GO 的富集与分类研究结果。HF1 VS P2 和 LF2 VS P2 的分析结果与 HF1 VS P1 和 LF2 VS P1 基本一致, 尽管参照对象变了, 但 HF1 和 LF2 其系统内差异表达的基因数目所处的状态并没有改变。HF1 和 LF2 基因所处的状态、平衡参数的统计分析结果列于表 2。

表 2 杉木转录组内差异基因表达平衡与否的状态分析
Table 2 Differentially expressed genes in different group

比较组	差异基因经 GO 富集后的基因表达的情况 显著 GO terms 间平衡参数 k 值	差异基因经 KEGG 富集后的表达情况 255 个 KEGG terms 间平衡参数 k 值
HF1 VS P1	6.450 8±4.179 5	1.208 2±1.270 5
LF2 VS P1	0.832 6±1.636 6	1.011 9±1.401 1
HF1 VS P2	3.764 1±3.563 4	1.244 4±1.146 6
LF2 VS P2	2.422 8±4.460 8	0.930 5±1.369 3

从表 2 中可以发现, HF1 的差异表达基因的分布处于不均匀、不平衡状态, LF2 的差异表达基因的分布趋于均匀、平衡状态。就 KEGG terms 的富集结果, 我们使用了更多的 KEGG terms (上调/下调基因) 资料, 从系统的角度, 探讨了全部资料的 KEGG terms 的富集结果与杂种生长性状分离间的关系, 计算其平衡系数, 结果也列于表 2。由表 2 可知, 255 个 KEGG terms 中, 相对优良亲本, HF1 的平衡系数 $k > 1.2$; 而 LF2 的平衡系数 k 十分接近于 1.0。

2.5 差异表达基因的 Pathway 显著性富集分析与子代的生长优势

通过 KEGG 数据库中的 Pathway 富集分析, 确定 DEGs 参与的主要生化代谢途径和信号转导途径等, 结果显示, 子代/亲代 4 个比较组中, 差异表达的基因分布在 74 ~ 130 条 Pathway 中。表 3 为处理比较组 HF1 VS P1, HF1 VS P2, LF2 VS P1 和 LF2 VS P2 富集结果中最显著的前 10 条 Pathway。由表 3 可见, HF1 VS P1 和 HF1 VS P2 富集结果接近; LF2 VS P1 和 LF2 VS P2 的富集结果接近。但在相同的参照物下, HF1 与 LF2 的 KEGG 富集结果差异很大: LF2 VS P1 与 HF1 VS P1 仅有两个 KEGG terms (苯丙氨酸代谢和亚油酸的新陈代谢) 相同, 其它八个 KEGG terms 不相同; 同时相同的 KEGG term 内, 上调基因数与下调基因数不同。LF2 VS P2 与 HF1 VS P2 间的比较结果: 也有两个 KEGG terms (苯丙素的生物合成和苯丙氨酸代谢) 相同, 其它八个 KEGG terms 不相同, 同时在相同的 KEGG terms 内, 上调基因数与下调基因数相同或不同。

表 3 4 个比较组的 KEGG 富集结果中前 10 个代谢途径 (表中 $p < 0.01$)
Table 3 KEGG enrichment and top 10 metabolic pathways of four group samples of parents and cross seedlings

LF2 VS P1	k 值	HF1 VS P1	k 值	LF2 VS P2	k 值	HF1 VS P2	k 值
淀粉和蔗糖的代谢	77/163	昼夜节律-植物	46/33	甲烷代谢	21/74	类黄酮生物合成	63/55
苯丙素的生物合成	74/137	抗原处理及存在	36/47	淀粉和蔗糖的代谢	74/166	昼夜节律-植物	46/33
甲烷代谢	25/70	剪接体	29/37	甘氨酸、丝氨酸和苏氨酸代谢	40/31	苯丙素的生物合成	82/129
甘氨酸、丝氨酸和苏氨酸代谢	28/43	MAPK 信号通路	57/51	苯丙素的生物合成	84/127	苯丙氨酸代谢	67/81
氰氨基酸代谢	24/54	类黄酮生物合成	66/52	类胡萝卜素生物合成	18/25	二萜生物合成	8/13
苯丙氨酸代谢	6/16	亚油酸的新陈代谢	4/0	利什曼病	13/7	黄酮和黄酮醇的生物合成	2/1
类黄酮生物合成	13/5	苯丙氨酸代谢	2/4	一个由叶酸组成的碳池	5/5	细胞色素代谢	1/2
亚油酸的新陈代谢	9/2	内吞作用	5/0	苯丙氨酸代谢	4/24	亚油酸的新陈代谢	2/1
类胡萝卜素生物合成	9/0	亚麻酸代谢	4/0	氰氨基酸代谢	2/15	香叶醇降解	0/2
戊糖和葡萄糖醛酸盐的相互转化	1/12	半乳糖代谢	3/0	氮代谢	8/9	亚麻酸代谢	2/2

综合以上 Pathway 富集分析结果可以发现, HF1 和 LF2 比较组的遗传差异, 表现在两个层面上: 其一是 KEGG-term 间的不平衡; 其二是 KEGG-term 内, 上调/下调的基因不均衡。这与 GO 的富集分析结果基本一致,

这正是 HF1 生长优于 LF2 的原因。但要揭示杉木杂交组合生长性状分离的内在机制, 最好在 HF1 和 LF2 比较组间, 选择相同的富集显著的 terms, 然后再比较 terms 内基因是上调表达, 还是下调表达。

按照这一思路, 我们比较了 LF2 VS P1 与 HF1 VS P1, HF1 VS P2 与 LF2 VS P2 几个相同的 KEGG terms 基因表达情况, 列于表 4。

表 4 相同 terms 下基因的差异表达情况
Table 4 Differential expression of genes within same terms

KEGG terms	HF1 VS P1	LF2 VS P1	KEGG terms	HF1 VS P1	LF2 VS P1	KEGG terms	HF1 VS P1	LF2 VS P1
亚油酸代谢基因	基因表达	基因表达	苯丙氨酸代谢	基因表达	基因表达	苯丙氨酸代谢	基因表达	基因表达
comp46528_c0	上调	下调	comp52987_c0	上调	上调	comp55452_c0	上调	上调
comp59067_c0	上调	下调	comp62983_c0	上调	上调	comp69809_c0	下调	基因沉默
comp59067_c1	上调	下调	comp61444_c0	上调	基因沉默			
comp52479_c0	上调	下调	comp71354_c2	下调	下调			
KEGG terms	HF1 VS P2	LF2 VS P2	KEGG terms	HF1 VS P2	LF2 VS P2	KEGG terms	HF1 VS P2	LF2 VS P2
苯丙素生物合成	基因表达	基因表达	苯丙素生物合成	基因表达	基因表达	苯丙素生物合成	基因表达	基因表达
comp71734_c2	上调	上调	comp60891_c0	下调	基因沉默	comp71175_c0	下调	基因沉默
comp61444_c0	上调	上调	comp73490_c0	下调	基因沉默	comp62983_c0	上调	上调
comp52987_c0	上调	上调	comp68751_c0	下调	下调	comp61921_c0	下调	基因沉默
comp62494_c0	上调	基因沉默	comp27611_c0	下调	基因沉默	comp60891_c1	下调	基因沉默

由表 4 可知, 在 LF2 中, 基因一般仍表现为下调表达或沉默, 而在 HF1 中上调表达, 只有亚油酸代谢中, 在 LF2 中下调表达, 在 HF1 中上调表达, 这是解释杂种生长分离的原因之一, 但它属于 terms 间的差异。

3 结论与讨论

杉木杂种高生产力的原因是其内在基因系统中基因分布不均匀、不平衡的原因。杉木优良组合中子代生长会出现分离: HF1 生长比 LF2 生长快, 是由于 HF1 GO terms 和 KEGG terms 层面不平衡, 以及 GO terms 和 KEGG terms 内的基因表达数量, 以及其内上调/下调的基因分布处于不均匀、不平衡状态; LF2 GO terms 和 KEGG terms 的差异基因的数量, 以及其内上调/下调的基因分布趋于均匀、平衡状态 (图 1 和图 2)。杉木杂种生产力的高低符合耗散结构理论: 不平衡体系, 能量大, 信息多, 生产力也最高。

杉木杂种生长性状存在分离现象, 对收获期不利。GO 和 KEGG 富集分析, 揭示有大量的基因参与了生长过程。标记辅助选择 MAS 研究工作量太大, 成本高, 技术复杂。利用分子生物学技术, 对育种群体展开遗传多样性研究, 选择表达基因互补的亲本, 进行杂交组配, 形成杂合体子代, 这样降低杂种生长性状的分离, 以达到收获期的大体一致, 是一条可行的技术路线^[21-22]。

参考文献:

- [1] 何贵平, 齐明, 程亚平, 等. 杉木杂交育种中亲本选配方法的研究[J]. 江西农业大学学报, 2016, 38 (4): 646-653.
- [2] 何贵平, 齐明. 杉木育种策略及应用[M]. 北京: 中国林业出版社, 2017: 33-77, P149-174.
- [3] 邓继光, 黄庆文, 刘凤君, 等. 一年生李杂种苗性状分离研究初报[J]. 沈阳农业大学学报, 1995, 26 (4): 363-368.
- [4] 王丹, 肖慈木, 李秀, 等. 柚杂交一年生苗期性状分离初步研究[J]. 中国南方果树, 2001, 31 (1): 3-7.
- [5] 鹿金颖, 毛永民, 申连英, 等. 枣实生后代性状分离研究[J]. 河北农业大学学报, 2003, 26 (5): 53-58.
- [6] 王冬梅, 伊凯. 苹果杂种叶片与果实的相关性研究[J]. 北方果树, 2004, (3): 12-14.
- [7] 张小猜, 赵政阳, 樊红科, 等. 苹果杂种 F1 代叶片性状分离及早期选择研究[J]. 西北农业学报, 2009, 18 (5): 228-231.
- [8] 崔艳波, 张绍铃, 吴华清, 等. 梨杂交后代童期和童程的研究[J]. 中国农学通报, 2011, 27 (2): 128-131.
- [9] 吴敏, 刘永杰, 李世强, 等. 库尔勒香梨杂种实生苗性状分离初步研究[J]. 新疆农业科学, 2011, 48 (5): 826-831.
- [10] LU Q, SHAO F, MACMILLAN C, et al. Genomewide analysis of the lateral organ boundaries domain gene family in *Eucalyptus grandis* reveals members that differentially impact secondary growth[J]. Plant Biotechnol J, 2018, 16 (1): 124-136.

- [11] ZHANG Y, HAN X, SANG J, *et al.* Transcriptome analysis of immature xylem in the Chinese fir at different developmental phases[J]. Peerj, 2016, 4 (17) : e2097; DOI 10.7717/peerj.2097.
- [12] HUANGHH, XULL, TONG ZK, *et al.* De novo characterization of the Chinese fir (*Cunninghamialanceolata*) transcriptome and analysis of candidate genes involved in cellulose and lignin biosynthesis[J]. BMC Genomics, 2012, 13 (1) : 648 – 658.
- [13] WANG ZJ, CHENJH, LIU WD, *et al.* Transcriptome Characteristics and Six Alternative Expressed Genes Positively Correlated with the Phase Transition of Annual Cambial Activities in Chinese Fir (*Cunninghamialanceolata* (Lamb.)Hook)[J]. PLOS one, 2013, 8 (8) : 1 – 14.
- [14] 齐明. 浙南地区杉木杂交组合再选择研究[J]. 林业科学研究, 2005, 18 (6) : 722 – 725.
- [15] 齐明, 何贵平, 周建革, 等. 杉木生长性状的杂种优势转录组分析[J]. 林业科学研究, 2019, 32 (3) : 113 – 120.
- [16] 李阳. 亚硝酸盐对水稻胚性愈伤组织的诱导作用及机制[D]. 湖北: 武汉大学, 2016.
- [17] 丁健. 沙棘果肉和种子油脂合成积累及转录表达差异研究[D]. 哈尔滨: 东北林业大学, 2016.
- [18] 冯延芝. 杜仲种仁转录组测序及 FAD3 基因鉴定与功能研究[D]. 北京: 中国林业科学研究院, 2016.
- [19] 翟荣荣. 超级稻协优 9308 根系杂种优势转录组分析[D]. 杭州: 中国农科院水稻所, 2013.
- [20] 蒋桂雄. 油桐种子转录组解析及油脂合成重要基因克隆[D]. 长沙: 中南林业科技大学, 2014.
- [21] 齐明, 王海蓉, 彭九生. 杉木育种园的遗传多样性和杂种优势群的划分[J]. 南方林业科学, 2018, 46 (6) : 17 – 21.
- [22] 齐明, 何贵平, 周建革, 等. 杉木无性系光合特性和呼吸速率的遗传变异性及杂交亲本筛选[J]. 江西农业大学学报, 2018, 40 (2) : 182 – 189.