

刘锦宇, 吴士浩, 仲威宇, 等. 大豆矮化突变体 *df1* 的候选基因鉴定[J]. 土壤与作物, 2025, 14(2): 207–216.

LIU J Y, WU S H, ZHONG W Y, et al. Candidate gene identification for soybean dwarf mutant *df1* [J]. Soils and Crops, 2025, 14(2): 207–216.

大豆矮化突变体 *df1* 的候选基因鉴定

刘锦宇¹, 吴士浩², 仲威宇¹, 张文韬¹, 朱晓斌¹, 夏正俊¹

(1. 中国科学院东北地理与农业生态研究所, 黑龙江哈尔滨 150081;

2. 江苏省农业科学院经济作物研究所, 江苏南京 210014)

摘要: 植物在外界物理与化学因素的胁迫下会产生基因突变。这种染色体基因组的变化不但能创造新的遗传变异材料, 为植物遗传改良与育种服务, 同时也是研究基因功能的重要技术手段。本研究对我国东北地区的大豆主栽品种黑河 43 进行伽玛射线诱变, 获得极度矮化突变体 *df1*。遗传分析显示, *df1* 突变为隐性单基因遗传。对杂合型衍生的 *df1* 突变型与正常野生型进行混池测序后, 利用鉴定基因的 BVF-IGV 流程, 确定了位于 1 号染色体上包含 69 个基因的大片段缺失是引起 *df1* 突变的遗传因子。缺失区域序列开发的分子标记与表型间进行连锁分析及野生型与突变型混池的转录组分析结果, 佐证了上述结论。根据本研究中野生型与突变型的基因表达谱以及公共资源中 5 000 多份品种的转录组数据分析结果, 最终确定 *Glyma.01G066600* 等 26 个基因为候选基因。研究结果为进一步通过精细定位克隆 *df1* 的功能基因以及明确其调控大豆生长发育的分子机理奠定了基础。

关键词: 大豆; 矮化突变; 大片段缺失; BVF-IGA; 混池测序

中图分类号: S565.1

文献标识码: A

Candidate gene identification for soybean dwarf mutant *df1*

LIU Jinyu¹, WU Shihao², ZHONG Weiyu¹, ZHANG Wentao¹, ZHU Xiaobin¹, XIA Zhengjun¹

(1. Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Harbin 150081, China;

2. Institute of Industrial Crops, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China)

Abstract: Genetic mutations can be induced by external physical and chemical factors. This chromosomal genomic change not only creates new genetic variation materials for genetic improvement and molecular breeding of plants, but also serves as an important technical tool for studying gene function. In this study, a dwarf mutant, *df1*, was identified in the mutant library generated by gamma radiation of the widely cultivated cultivar Heihe 43 in Northeast China. Genetic analysis indicated that the *df1* mutation is controlled by a single recessive gene. By analyzing the resequencing data of *df1* and wild-type bulked pools sampled from a segregating line using the BVF-IGV pipeline, a large fragment deletion containing 69 genes on chromosome 1 is identified as the causal genetic factor for the *df1* mutation. Association analysis using molecular markers derived from the large fragment deletion and transcriptome analysis of RNA-seq data of bulked *df1* and wild-type supports the above conclusion. Twenty-six candidate genes including *Glyma.01G066600* are finally determined by combined analysis of gene expression profiles and over 5 000 cultivars from the public database. The results of this study could facilitate the cloning of the functional gene for *df1* mutation through fine-mapping and understanding of its molecular mechanism in regulating plant growth and development.

Keywords: soybean; dwarf; large-fragment deletion; BVF-IGV; bulk sequencing

收稿日期: 2024-08-07; 修回日期: 2024-10-18.

基金项目: 黑土地保护与利用科技创新工程专项资助 (XDA28070400); 国家重点研发计划政府间国际科技创新合作 (2022YFE0130300); 国家自然科学基金 (U21A20215, 32272094).

第一作者简介: 刘锦宇 (1997-), 硕士, 主要从事豆类分子生物学研究. E-mail: jy112190305@163.com.

通信作者: 夏正俊 (1962-), 博士, 研究员, 主要从事大豆遗传学及分子生物学研究. E-mail: xiazhj@iga.ac.cn.

0 引言

半矮秆品种的应用是第一次绿色革命的核心内容。植株矮化极大地提高了全球禾谷类作物的产量,缓解了世界许多发展中国家的粮食问题^[1]。与禾本科植物不同,大豆的生殖器官尤其是其结荚部位具有分散性,这使得其植株产量与株型的关系更为复杂。大豆理想株型概念正是为揭示这种关系而提出的^[2]。该概念与绿色革命的设想相结合,有助于深入研究大豆株型与产量及品质的关系^[3]。

大豆的株高是理想株型的核心组成部分。理论上讲,半矮秆大豆品种能够提高大豆的抗倒伏性,而增加半矮秆大豆的种植密度可以实现大豆的高产^[4]。大豆的株高是由多个基因控制的重要农艺性状,它不仅与植株的生长(结荚)习性密切相关,还与植株的生育期、种植密度、日照长度、光照强度、栽培管理等外部环境密切相关。大豆生长习性基因 *Dt1* 与 *Dt2* 位点间的相互作用决定了生长习性^[5-8]。*E1-E4*、*QNE1* 等生育期基因的成功克隆^[9-13],使得研究者通过解析光周期途径与生长习性基因之间的互作调控网络来定向对株高与株型进行分子调控成为可能。

在水稻绿色革命中,编码赤霉素合成的 GA20 氧化酶基因 *SD1* 降低了酶活性,从而减少了赤霉素的合成,促成了半矮秆水稻品种的育成。半矮秆水稻品种提高了水稻植株的抗倒伏能力,显著提升了水稻产量^[14]。在大豆中,已有不少研究者对控制株高与节间长度的 QTL 及相关基因进行了深入研究。在 QTL 定位方面,张秀彤^[15] 将一个矮秆基因定位于 19 号染色体中;Cheng 等^[16] 将一个矮化突变的候选基因定位于 1 号染色体 10.51~12.06 Mb 之间。在候选基因研究方面,张久坤等^[17] 鉴定了 18 个与株高相关的候选基因;Su 等^[18] 确定转录 *Aux/IAA* 蛋白的 *Glyma.09g193000* 基因是一个矮秆株型多分枝突变的候选基因。同时,研究表明赤霉素相关通路基因与株高及耐阴性相关联^[19-20]。转录组数据显示,大豆矮秆突变体 *Gmdwarf1* 中赤霉素通路及激素生物合成途径也受到明显调控^[21]。上述研究结果表明,大豆株高调控较为复杂,需要进一步深入研究。

BSA (Bulk segregant analysis) 混池技术已在基因定位中得到广泛应用。例如, BULK 测序 (bulk segregant RNA-seq)^[22]、RNA 混池转录组测序 (BSR-seq)^[23]、MutMap (Multi-platform mapping)^[24]、BVF-IGV 流程 (bulk sequencing, variant calling, functional annotation by SnpEFF software, and Integrative Genomics Viewer pipeline)^[25] 均可快速确定候选基因位点。

本研究通过⁶⁰Co γ 射线诱变大豆品种黑河 43,从突变后代中筛选出矮化突变体 *dfl*。通过田间及室内分析明确了 *dfl* 的株高及产量等性状,进而利用杂合型 *dfl* 的分离株行,构建突变池 *dfl* 与野生型池,通过测序获得基因组与转录组数据。最后,利用 BVF-IGV 流程确定 1 号染色体上的大片段缺失区段是 *dfl* 变异的分子基础。这一结论得到该区段序列开发的分子标记及缺失区域基因表达水平的佐证。同时,利用公共数据库中该区域内注释基因的表达特征,进一步缩小了候选基因的数量。

1 材料与方法

1.1 矮化突变体 *dfl* 的获得及特征观察

采用 200 Gy 剂量的⁶⁰Co γ 射线,对大豆品种黑河 43 进行诱变处理^[26]。种植 M0 代处理后的种子并混合收获,随后扩繁至 M1 代并单株收获。在 M2 代株行中进行变异观察, T878 株行共有 10 株植株,其中 2 株表现为极度矮化突变,将其定名为 *dwarf1* (*dfl*)。次年对 M2 代正常单株后代 M3 代群体进行表型调查,记录正常株与 *dfl* 突变株的比例,同时对正常株与 *dfl* 突变株的株高、节间数及单株产量等进行表型观察与考种。进一步扩繁至 M5 至 M6 代,统计分离株行中正常株与分离株的比例,以确定基因的变异是否符合孟德尔隐性单基因遗传。同时,将其用于基因型与表型的关联分析,以检测混池测序的结果。

1.2 基因组与转录组测序用混池样品的获取

为鉴定导致 *dfl* 变异的候选基因,对构建的 BSA 混池进行基因组与转录组测序。测序样本来源于

M3 代分离株行的突变株与正常株, 在植株长到 V5 至 V6 期时, 选取 30 株 *df1* 突变株顶部完全展开的嫩叶各一片, 依次叠放, 切取或剪取相同大小的叶片作为突变池混合样品, 确保每株叶片的体积相等。采用相同方法, 获取正常野生表型植株的混样。样品立即置于液氮中, 冷冻保存。DNA 与 RNA 提取、建库与测序均在安诺优达(北京)公司进行。

1.3 利用 MutMap 软件与 BVF-IGV 鉴定候选基因

使用 NGS QC Toolkit (v2.3.3) 在默认参数下, 对混池测序数据进行过滤。首先, 针对过滤后的混池测序数据, 利用 MutMap 查找候选基因位点及区间, 除将个体数设为 30 外, 其余参数均按照推荐设置进行运算^[24]。大豆参考基因组为 Wm82.a2.v1 (V275), 从 phytozome-next.jgi.doe.gov 下载(下载日期为 2016 年 4 月 23 日)。同时, 利用本实验室自行开发的 BVF-IGV 程序鉴定候选基因^[25], 即通过混池测序、变异检测, 并利用 SnpEFF 进行功能注释, 相关结果利用 Integrative Genomics Viewer (IGV) 软件进行批量观察, 以确定初步的候选基因或候选区域。

1.4 BVF-IGV 鉴定结果的验证

通过 BVF-IGV 流程, 确定一个大片段缺失区域为候选基因所在区域。随后, 获取大片段缺失所在区域的序列信息, 对每 500 bp 序列利用本地化的 primer3 软件 (<https://github.com/primer3-org/primer3>) 进行引物设计, 再经过 standalone_Blast 软件 (<http://blast.ncbi.nlm.nih.gov>) 即本地化的 Blast 扫描 V275 版本的基因序列, 选择特异性引物用于验证大片段缺失。在 M5 至 M6 代分离株行衍生的分离群体中, 进行基因型与表型的关联分析。采用本实验已发表的快速提取大豆种子 DNA 方法^[27] 以及叶片 DNA 提取方法^[28] 进行样本 DNA 提取。

1.5 混池转录组分析及大片段缺失内基因的表达特征

对于获得的混池转录组数据, 通过 RNAseq-workflow 软件 (<https://github.com/twbattaglia/RNAseq-workflow>) 进行转录组分析^[25,28]。为进一步明确大片段缺失内基因的表达特征, 从公共数据库 (<http://soy-atlas.venanciogroup.uenf.br>) (下载日期为 2023 年 12 月 26 日) 下载超过 5 000 份材料的不同组织的基因表达数据, 并对不同组织的表达量进行加权平均, 利用 TBtools 中的 heatmap 程序制作相关热图^[29-30]。

1.6 统计分析

在 GraphPad Prism 5 (v5.01, San Diego, USA) 中进行突变体与正常株等表型数据的比较 *t* 检验以及分离群体中正常株与 *df1* 的分离比的卡方检测。

2 结果与分析

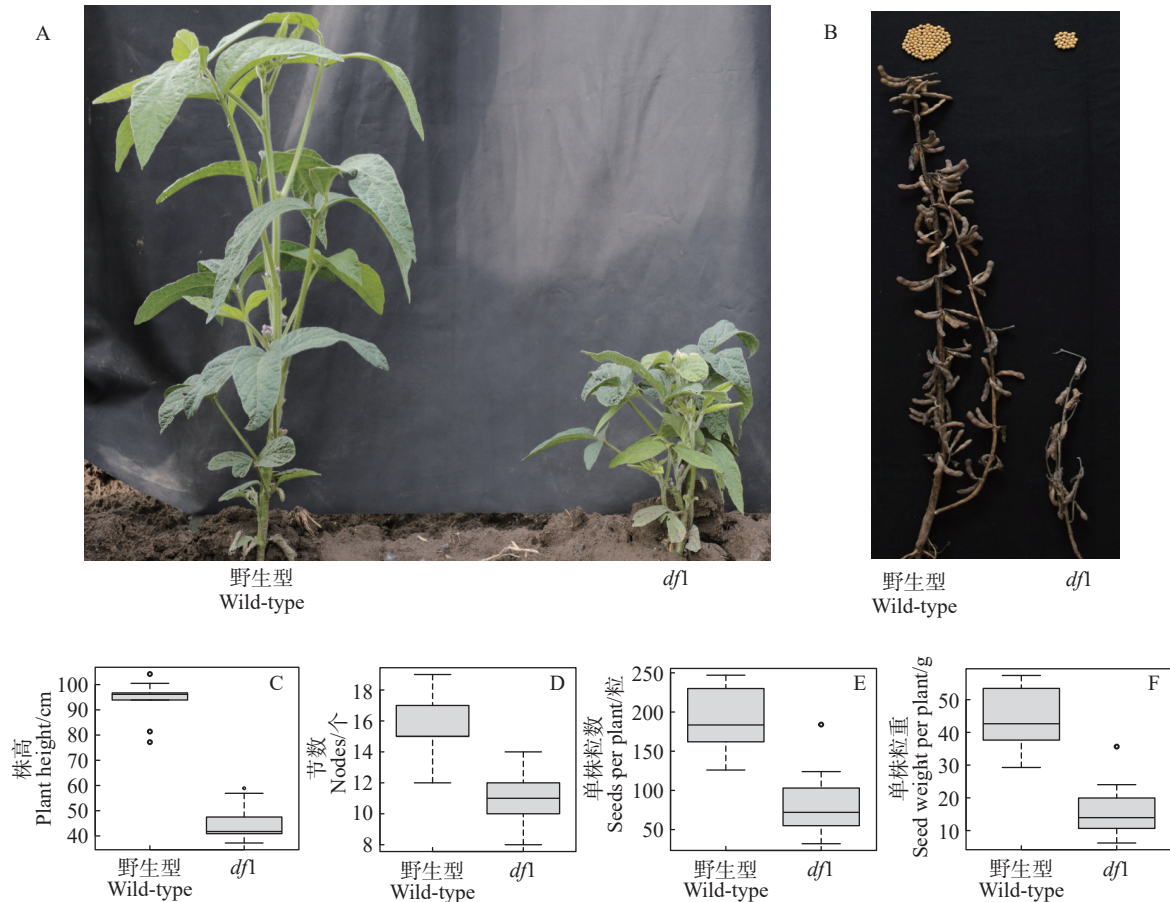
2.1 矮化突变 *df1* 的表型特征

在 M2 代的 T878 株行中, 共有 10 株植株, 其中 8 株呈现正常野生型表型, 2 株为矮化突变株, 将其定名为 *dwarf1* (*df1*) (图 1A)。对 M2 代正常株后代的一个分离株行进行调查, 发现野生型单株有 33 株, *df1* 表型株为 11 株, 符合 3 : 1 的分离比例。进一步对扩繁至 M4 代的分离株行进行统计, 正常株为 463 株, *df1* 表型株为 197 株, 经统计分析, 符合孟德尔隐性单基因 3 比 1 的分离比 ($P = 1.000$)。

与野生型植株相比, *df1* 具有更短的茎秆和较少的节间数。在 R8 期, 野生型株高为 92.1 cm ($n = 30$), 而 *df1* 突变株株高仅为 44.4 cm ($n = 30$), 为野生型的 48.2% (图 1C)。*df1* 突变株的节间数为 11.0 个 ($n = 30$), 比野生型少 5 个 (图 1D); 相应地, 平均节长也从野生型的 5.94 cm ($n = 30$) 下降至突变株的 3.97 cm ($n = 30$)。*df1* 突变株单株荚数为 41.0 个 ($n = 30$), 仅为野生型 71.0 个 ($n = 30$) 的 57.8% (图 1E); 单株粒数仅为野生型的 46.1%; 与野生型的单株粒重 50.1 g ($n = 30$) 相比, *df1* 突变株仅为 19.2 g ($n = 30$) (图 1F)。这表明矮化突变通过影响植株的基本生长发育, 改变了植株的产量。

2.2 MutMap 分析获得多个 SNP-index 峰值

对 BSA 分析的两个混池重测序数据进行质量控制分析, 结果显示: 野生型混池获得的 reads 数量为 168 615 124 个, 平均 reads 长度为 150 bp; 突变 *df1* 混池共获得的 reads 数为 166 790 766 个, 平均 reads



注: A: 黑河 43 由突变体 *df1* 分离株行中 V8 期的野生型 (左) 和 *df1* 突变体 (右); B: R8 野生型 (左) 和 *df1* 突变体 (右); C: V8 期株高; D: V8 期节间数; E: R8 期单株荚数; F: R8 期单株粒重。

Note: A: Appearance of *df1* and its wild type in the segregating line from a γ irradiation-induced mutant library of the soybean (*Glycine max*) cultivar "Heihe 43" at V8 stage; B: Appearance at R8 stage showing contrast difference in plant height and pod numbers; C: Plot for plant height at V8 stage; D: Plot for node numbers at V8 stage; E: Plot for pods per plant at R8 stage; F: Plot for weight per plant at R8 stage.

图 1 矮化突变 *df1* 及其野生型表型特征

Fig. 1 Phenotypes of *df1* mutant and wild-type

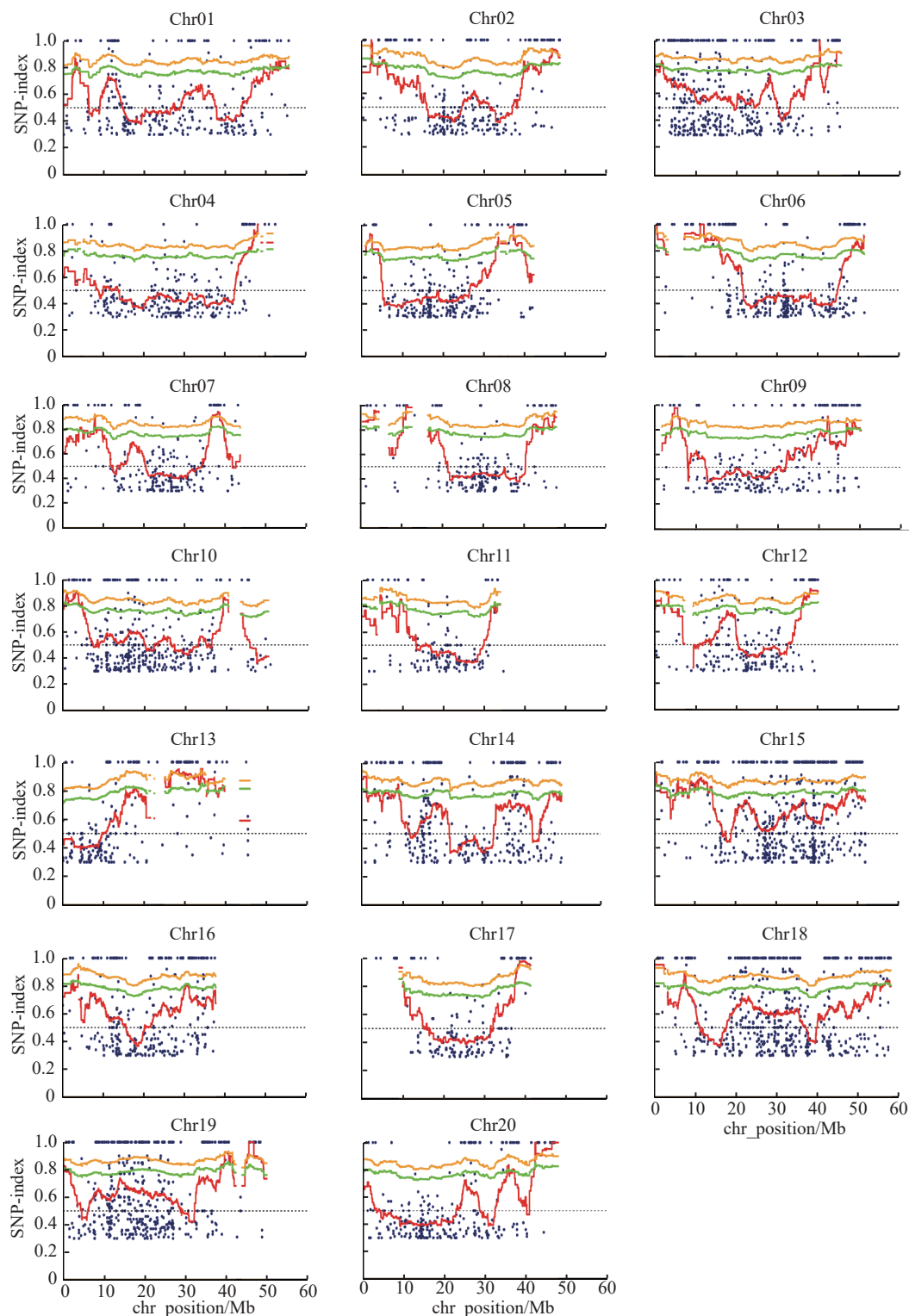
长度为 150 bp。利用 MutMap 软件对过滤后的测序数据进行分析^[24], 结果如图 2 所示: 几乎每条染色体上均存在平均 SNP-index 值超过 95% 置信区间的位点, 呈现出单峰或多峰型, 同时也有 SNP-index 值超过 99% 置信区间的单峰或多峰。

2.3 利用 BVF-IGV 流程确定 *df1* 位于 1 号染色体的大片段缺失中

利用本实验室建立的 BVF-IGV 流程对混池测序数据进行分析^[25], 结果显示: 与野生型相比, 突变体 *df1* 混池中特有的位点约有 1 227 个; 同样, 与突变池相比, 正常株混池中特有的位点有 1 109 个, 与这些位点相对应的基因数为 740 个。通过 BVF-IGV 流程的 batch 命令对候选基因进行初步筛选, 发现突变池中 *Glyma.01G071700* 基因呈现整个基因缺失型。进一步在 IGV 上确认该基因缺失范围是包含 69 个基因 (Chr01: 10396934 ~ 13719764/*Glyma.01G066300* ~ *Glyma.01G073100*) 的大片段缺失 (图 3)。同时, 利用 IGV, 我们观察了该区域左侧的 500 个基因与右侧的 500 个基因, 并未发现明确与表型相关的变异。

2.4 相关分子标记验证了 *df1* 突变体中的大片段缺失

将大片段缺失及附近序列依次截取成 500 bp 的片段, 利用本地化的 primer3 程序对相关序列逐一设计引物。接着, 利用 standalone-Blast 扫描 V275 大豆基因组上成功引物的分布情况, 选取均匀分布的 30 对特异性引物用于 PCR 验证 (表 1)。依据相关的分子标记, 可以明确判断 *df1* 突变体中存在大片段缺失 (图 3B)。同时, 对 M3 代分离株行的正常株自交衍生物进行了关联分析, 在分离群体的 660 株植株

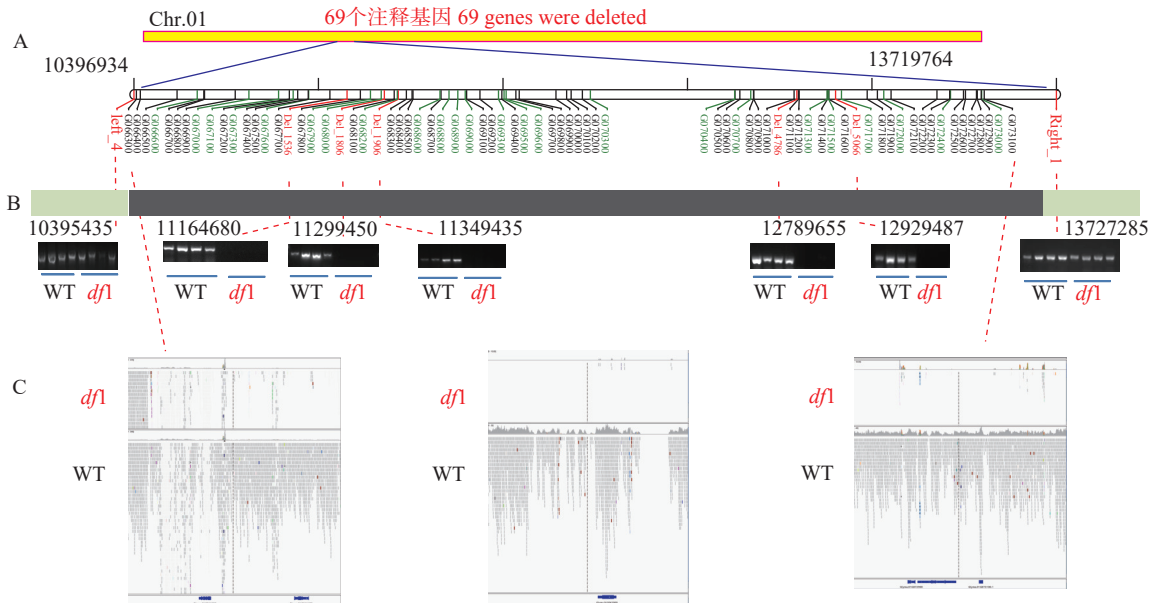


注: 矮化突变体 *dfl* 与野生型 WT 混池的株数均为 30, 两混池的基因组测序深度约为 25X。混池测序数据经 MutMap 计算而获得的每条染色体对应的 SNP-index 曲线; 混池蓝色点: SNP-index; 红色线为平均滑动窗口 4 Mb 及步长为 10 Kb 的 SNP-index; 绿色线为平均滑动窗口 95% 的置信区间; 橙色线为平均滑动窗口 99% 的置信区间。

Note: The number of plants in both the dwarf mutant *dfl* and the wild - type WT bulks is 30, and the genome sequencing depth of the two bulks is about 25X. The SNP - plot corresponding to each chromosome obtained by MutMap analysis of the resequencing data of bulks. Blue dot: SNP-index; red line: sliding window average for SNP-index with the sliding window average of 4 Mb interval with 10 Kb increment; green line: sliding window average for 95%-confidence interval; orange line: sliding window average for 99%-confidence interval.

图 2 M4 代杂合型突变体由来矮化突变体 *dfl* 与野生型 WT 叶片混池 SNP-index 曲线图

Fig. 2 SNP-index plots for mutant *dfl* and wild-type (WT) leaf bulks derived from a heterozygous plant in M4 generation



注：A：经对突变表型与野生表型的混池重测序数据经 BVF-IGV 鉴定位于 1 号染色体中约 3 322 830 bp 缺失（包含 69 个注释基因）为 *dfl* 突变的候选位点；B：利用大片断缺失序列开发的分子标记验证大片断缺失的真实性；C：BVF-IGV 分析过程中 IGV 截图示 *dfl* 突变混池中大片断缺失位置。

Note: A: Through the re-sequencing data of the mutant *dfl* and wild type bulks, a deletion of about 3,322,830 bp (including 69 annotated genes) located on chromosome 1 is identified by BVF - IGV as the candidate gene located region of the *dfl* mutation; B: Use molecular markers developed specifically from the sequences of the large fragment deletion to verify the authenticity of the large-fragment deletion; C: IGV snapshots showing the starting and ending position of the large fragment deletion in *dfl* during BVF-IGV analysis.

图 3 引致大豆矮化突变 *dfl* 的候选基因所在区域的鉴定及其分子验证

Fig. 3 Molecular verification of the large fragment deletion in *dfl* mutant

中，经 Del_1806 与 Del_4 786 的基因型鉴定所有的 197 株 *dfl* 突变表型个体均表现为缺失型，所有的 463 野生表型植株均呈现出条带，其表型与两分子标记间呈为共分离关系。鉴于 Del_1806 与 Del_4 786 分子标记不能区别杂合型与纯合的野生型，因而尚无法确定纯合野生型 *DF1/DF1* 与杂合型基因型 *DF1/dfl* 间的比例。关联分析的结果验证了 BVF-IGA 流程的鉴定结果。

表 1 鉴定大片断缺失的引物信息

Table 1 Primer pairs used for verification of the existence of the large deletion

标记物 Marker	正向引物 Forward primer (5'-3')	反向引物 Reverse primer (3'-5')
left4	TGGACATAACCTCATTGGCA	AGGCCCTGCTGCTAATATT
Del_1 536	TGTGCCTGATGTTTCACACA	AGAGCAGGTCTGTGTAGCCC
De_1806	AAGTGCCGAACTTCTCTCA	GTTGGCTGGTGGGAGAGTT
Del_1906	CCTGACTAATTCGGGCTAGG	TATCCAACGGTGCAAAACAA
del 4 786	TGACGATGGGTCTCTCGTAA	CGTGTAGCATTGCGAATCAC
del_5 066	GCTTGAAACATCGGAGGGTA	GGCTCGCCAAGGTGTATAAA
Right_1	ATTCATTCCCGAAAAAGGCT	GCTGGCCAAAAAGGGTAACT

2.5 大片段所在区域基因表达量分析及候选基因进一步筛选

为进一步验证候选基因鉴定的结果，利用 RNAseq-workflow 软件分析混池的转录组数据^[25,28]，获得了基因的相对表达量及校正后的表达量（图 4，表 2）。结果确认缺失区域内的所有基因均无表达，同时在大片段缺失区域外，缺失左侧与右侧基因在表达上未见明显差异，这进一步佐证了 *dfl* 突变体中大片段缺

失的存在。同时, 我们从公共数据库 (<http://soyatlas.venanciogroup.uenf.br>) 下载了最新的 5 000 个样品的基因表达量数据, 通过对不同组织的样品进行加权平均处理, 利用 TBtools^[30] 的 heatmap 程序制作结果热图。结果显示, 各组织中目标基因的表达量与本研究 WT 混池中的大片段缺失所在区域的基因表达量趋势一致。在排除不同组织中不表达或表达量很低的基因后, 将大片段缺失内的候选基因数量从 69 个减少至 *Glyma.01G066600* (BES1-interacting Myc-like protein) 等 26 个基因 (表 2)。



注: 与野生型叶片 (WT) 与数据库来源的品种叶片的正常表达相比, 矮化突变混池 (*dfl*) 在大片断区间内的基因均为零表达, 佐证了大片断缺失的真实性。在图中显示的数值为从 <http://soyatlas.venanciogroup.uenf.br> 下载的 5 000 多个品种相关组织的平均转录丰度。Note: Compared with the normal expression in leaves of wild - type and varieties from databases, all genes in the large - fragment intervals of the *dfl* leaf bulk show no or near no expression, supporting the authenticity of large - fragment deletions. The data of average transcriptional abundances of relevant tissues of over 5000 cultivars are downloaded from <http://soyatlas.venanciogroup.uenf.br>.

图 4 大片断缺失区域内注释基因在矮化突变 *dfl*、野生型混池及公共数据库中 5 000 个品种不同组织中的表达谱

Fig. 4 Expression profiles of the annotated genes in the large - fragment deletion intervals in different tissues of the dwarf mutant *dfl*, the wild - type bulks, and 5,000 varieties deposited in the public database

3 讨 论

根据传统的正向遗传学, 利用图谱法已在大豆中成功克隆了一系列控制大豆生育期与株高的基因^[12,31]。随着基因组尤其是 T2T 基因组数据库^[32] 的发表, 以 BSA 技术为主体, 即以突变池与正常池的基因组与转录组测序数据为基础, 发展出了一系列寻找或鉴定突变候选基因的技术体系, 如 MutMap+^[33]、MutMap-gap^[34]、QTL-seq^[35]。如今, 已有多种鉴定程序与软件^[36] 可供研究者选择。本研究采用 MutMap 技术进行目

表 2 位于 1 号染色体的候选区间（大片断缺失）的基因信息
Table 2 Candidate genes within the deleted region on chromosome 1

基因 Gene name	功能注释 Functional annotation	表达量		表达量（更正）	
		Expression (reads)		Expression (corrected)	
		<i>dfl</i> 混池 <i>dfl</i> bulk	野生型混池 WT bulk	<i>dfl</i> 混池 <i>dfl</i> bulk	野生型混池 WT bulk
<i>Glyma.01G066600</i>	BES1-interacting Myc-like protein 2	0	766	0.00	286.94
<i>Glyma.01G067000</i>	Phosphoinositide 4-kinase gamma 7	0	2 738	0.00	1 025.65
<i>Glyma.01G067100</i>	UDP-Glycosyltransferase superfamily protein	0	1 424	0.00	533.43
<i>Glyma.01G067200</i>	GDSSL-like Lipase/Acylhydrolase superfamily protein	0	2	0.00	0.75
<i>Glyma.01G067300</i>	RNA polymerase II large subunit	0	96	0.00	35.96
<i>Glyma.01G067600</i>	Sugar transporter 1	2	681	0.75	255.10
<i>Glyma.01G067900</i>	Transducin/WD40 repeat-like superfamily protein	0	1 043	0.00	390.71
<i>Glyma.01G068000</i>	Chaperone protein htpG family protein	0	20 462	0.00	7 665.07
<i>Glyma.01G068200</i>	Senescence-related gene 1	0	88	0.00	32.96
<i>Glyma.01G068600</i>	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein	0	1 282	0.00	480.24
<i>Glyma.01G068800</i>	WPP domain-interacting protein 2	0	1 388	0.00	519.94
<i>Glyma.01G068900</i>		0	48	0.00	17.98
<i>Glyma.01G069000</i>		0	917	0.00	343.51
<i>Glyma.01G069300</i>	Basic leucine-zipper 42	0	10	0.00	3.75
<i>Glyma.01G069500</i>	TCP-1/cpn60 chaperonin family protein	0	7 816	0.00	2 927.87
<i>Glyma.01G069600</i>	0	0	271	0.00	101.52
<i>Glyma.01G070300</i>	Phosphoribulokinase / Uridine kinase family	0	216	0.00	80.91
<i>Glyma.01G070400</i>	Transferases; Folic acid binding	0	34	0.00	12.74
<i>Glyma.01G070700</i>	Chlorsulfuron/Imidazolinone resistant 1	4	669	1.50	250.61
<i>Glyma.01G071300</i>	Adenosylmethionine decarboxylase family protein	0	11 539	0.00	4 322.51
<i>Glyma.01G071500</i>		0	21	0.00	7.87
<i>Glyma.01G071700</i>	NSP-interacting kinase 2	0	302	0.00	113.13
<i>Glyma.01G071800</i>	Fatty acid desaturase 8	0	4	0.00	1.50
<i>Glyma.01G072000</i>	F-box/RNI-like superfamily protein	0	249	0.00	93.28
<i>Glyma.01G072400</i>	NB-ARC domain-containing disease resistance protein	0	12	0.00	4.50
<i>Glyma.01G073000</i>	FAD-linked oxidases family protein	0	678	0.00	253.98

标性状的 QTL 检测，然而所得结果呈现单峰或多峰型，这增加了后续进一步锚定与克隆基因的难度。

为继续定位和克隆目标基因，我们运用本实验室研发的 BVF-IGV 流程（BVF-IGV pipeline）^[25]，鉴定出目标基因位于一个大片段缺失中，并结合转录组数据进一步获取大片段缺失的信息。利用 BVF-IGV 流程，本实验室已成功鉴定了多种类型的大豆突变基因^[28-29]。成功地应用该流程不仅能准确地鉴定出 1 至数个 bp 缺失的变异，同时还可能用于发现和捕捉大片段缺失变异。此外，BVF-IGV 方法还可以帮助排除大豆基因组中由于杂合型位点而产生的假阳性候选位点。

在鉴定到大片段缺失的基础上，我们对附近的 500 个基因进行了筛查，未发现其他与表型一致的变异。同时，本研究中的转录组数据及针对缺失区域设计的引物分析结果，进一步验证了大片段缺失是引致 *dfl* 变异的遗传因子。在转录组数据的支持下，我们将 69 个基因缩小至 26 个基因。本研究中所获得的大片段

缺失位点, 与已有的报道的缺失位点极为相近^[16]。在已明确的调控大豆株高的因素中, 除生长习性基因、生育期基因外, 赤霉素 (GAs) 和油菜素内酯 (BRs) 是调节植物高度的两种主要激素^[37]。赤霉素和油菜素内酯的生物合成途径受阻常会导致植株矮化^[8,14,38], 异源过表达 *AtDREB1A* 会影响 GA 生物合成途径, 引发植株矮化^[39]。Qin 等^[31] 克隆了大豆株高调控主效基因 *PHI3*, 该基因编码一个 WD40 蛋白, 与 GmCOP1s 互作降解 STF1/2 转录因子, 从而促进茎秆伸长。UV-B 依赖型 E3 连接酶 GmILPA1 通过 GA 途径控制大豆植株的高度与夹角^[40], 同时该基因还通过 GA 等通路控制大豆叶柄长度^[25]。此外, Cheng 等^[16] 通过对 *dfl* 类的突变体进行研究指出, 该类突变表型不能被施加外源 GA 激素得以恢复, 说明 *dfl* 突变引致的停滞与 GA 通路关系并不密切, 这意味着存在其他调控大豆株型的通路。

4 结 论

大豆主栽品种黑河 43 经伽玛射线诱变后, 获得了极度矮化突变体 *dfl*。通过 BVF-IGV 流程, 确定了位于 1 号染色体上包含 69 个基因的大片段缺失是引发 *dfl* 突变的分子基础, 这一结论得到了缺失区域分子标记关联分析以及混池转录组数据分析的有力佐证。同时, 依据缺失区段内的基因表达特征, 确定了 26 个基因为候选基因。本研究的结果为进一步图位克隆以及解析 *dfl* 调控大豆生长发育的机理奠定了坚实基础。

参考文献 (References):

- [1] KHUSH G S. Green revolution: the way forward[J]. *Nature Reviews Genetics*, 2001, 2(10): 815–822.
- [2] 杜维广, 盖钧镒. 大豆超高产育种研究进展的讨论 [J]. *土壤与作物*, 2014, 3(3): 81–92.
DU W G, GAI J Y. A discussion on advances in breeding for super high-yielding soybean cultivars[J]. *Soils and Crops*, 2014, 3(3): 81–92.
- [3] LIU S L, ZHANG M, FENG F, et al. Toward a “green revolution” for soybean[J]. *Molecular Plant*, 2020, 13(5): 688–697.
- [4] 朱金龙, 刘锦宇, 朱晓斌, 等. 大豆耐密植相关性状遗传调控机制研究进展 [J]. *大豆科学*, 2023, 42(1): 109–117.
ZHU J L, LIU J Y, ZHU X B, et al. Research progress on genetic regulatory mechanism of soybean traits related to the tolerance to high planting density[J]. *Soybean Science*, 2023, 42(1): 109–117.
- [5] TIAN Z X, WANG X B, LEE R A, et al. Artificial selection for determinate growth habit in soybean[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2010, 107(19): 8563–8568.
- [6] LIU B H, WATANABE S, UCHIYAMA T, et al. The soybean stem growth habit gene Dt1 is an ortholog of *Arabidopsis* TERMINAL FLOWER1[J]. *Plant Physiology*, 2010, 153(1): 198–210.
- [7] PING J Q, LIU Y F, SUN L J, et al. Dt2 is a gain-of-function MADS-domain factor gene that specifies semideterminacy in soybean[J]. *The Plant Cell*, 2014, 26(7): 2831–2842.
- [8] SUN T P. The molecular mechanism and evolution of the GA-GID1-DELLA signaling module in plants[J]. *Current Biology*, 2011, 21(9): R338–R345.
- [9] LIU B H, KANAZAWA A, MATSUMURA H, et al. Genetic redundancy in soybean photoresponses associated with duplication of the phytochrome A gene[J]. *Genetics*, 2008, 180(2): 995–1007.
- [10] WATANABE S, HIDESHIMA R, XIA Z J, et al. Map-based cloning of the gene associated with the soybean maturity locus E3[J]. *Genetics*, 2009, 182(4): 1251–1262.
- [11] WATANABE S, XIA Z J, HIDESHIMA R, et al. A map-based cloning strategy employing a residual heterozygous line reveals that the *GIGANTEA* gene is involved in soybean maturity and flowering[J]. *Genetics*, 2011, 188(2): 395–407.
- [12] XIA Z J, WATANABE S, YAMADA T, et al. Positional cloning and characterization reveal the molecular basis for soybean maturity locus E1 that regulates photoperiodic flowering[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2012, 109(32): E2155–E2164.
- [13] XIA Z J, ZHAI H, ZHANG Y F, et al. QNE1 is a key flowering regulator determining the length of the vegetative period in soybean cultivars[J]. *Science China Life Sciences*, 2022, 65(12): 2472–2490.
- [14] WANG Y H, LI J Y. Molecular basis of plant architecture[J]. *Annual Review of Plant Biology*, 2008, 59: 253–279.
- [15] 张秀彤. 大豆隐性矮秆基因 19-DW 的精细定位 [M]. 长春: 吉林大学, 2019: 30–51.
ZHANG S T. Mapping of soybean recessive Dwarf gene 19-DW[M]. Changchun: Jinling University, 2019, 30–51.
- [16] CHENG W, GAO J S, FENG X X, et al. Characterization of dwarf mutants and molecular mapping of a dwarf locus in soybean[J]. *Journal of Integrative Agriculture*, 2016, 15(10): 2228–2236.

- [17] 张久坤, 齐阳阳, 李立竹, 等. 利用BSA法定位大豆全基因组株高QTL及候选基因分析[J]. 华北农学报, 2020, 35(S1): 1–10.
ZHANG J K, QI Y Y, LI L Z, et al. Genome-wide mapping of QTLs and candidate genes underlying plant height in soybean using BSA method[J]. *Acta Agriculturae Boreali-Sinica*, 2020, 35(S1): 1–10.
- [18] SU B H, WU H T, GUO Y, et al. GmIAA27 encodes an AUX/IAA protein involved in dwarfing and multi-branching in soybean[J]. *International Journal of Molecular Sciences*, 2022, 23(15): 8643.
- [19] LIU X, WICKLAND D P, LIN Z C, et al. Promoter deletion in the soybean Compact mutant leads to overexpression of a gene with homology to the C20-gibberellin 2-oxidase family[J]. *Journal of Experimental Botany*, 2023, 74(17): 5153–5165.
- [20] SHAN F X, SUN K X, GONG S D, et al. Effects of shading on the internode critical for soybean (*Glycine max*) lodging[J]. *Agronomy*, 2022, 12(2): 492.
- [21] ZHANG F, SHEN Y T, SUN S, et al. Genome-wide expression analysis in a dwarf soybean mutant[J]. *Plant Genetic Resources*, 2014, 12(S1): S70–S73.
- [22] LIU S Z, YE H C T, TANG H M, et al. Gene mapping via bulked segregant RNA-Seq (BSR-Seq)[J]. *PLoS One*, 2012, 7(5): e36406.
- [23] LI Y H, SHI X H, HU J H, et al. Identification of a recessive gene PmQ conferring resistance to powdery mildew in wheat Landrace qingxin-mai using BSR-seq analysis[J]. *Plant Disease*, 2020, 104(3): 743–751.
- [24] ABE A, KOSUGI S, YOSHIDA K, et al. Genome sequencing reveals agronomically important loci in rice using MutMap[J]. *Nature Biotechnology*, 2012, 30(2): 174–178.
- [25] GAO Y, ZHU J L, ZHAI H, et al. Dysfunction of an anaphase-promoting complex subunit 8 homolog leads to super-short petioles and enlarged petiole angles in soybean[J]. *International Journal of Molecular Sciences*, 2023, 24(13): 11024.
- [26] 郭宁, 郑佳佳, 胡博, 等. 菜豆品种黄金勾伽玛射线突变体库的建立及突变表型观察[J]. 土壤与作物, 2018, 7(2): 168–176.
GUO N, ZHENG J J, HU B, et al. Construction and phenotypic observation of a gamma radiation mutant library of common bean cultivar Golden Hook[J]. *Soils and Crops*, 2018, 7(2): 168–176.
- [27] 夏正俊, 李玉卓, 朱金龙, 等. 快速、无损大豆种子连续取样技术及其DNA制备[J]. 植物学报, 2021, 56(1): 56–61.
XIA Z J, LI Y Z, ZHU J L, et al. A rapid, non-destructive and continuous sampling technique and DNA extraction for soybean seed[J]. *Chinese Bulletin of Botany*, 2021, 56(1): 56–61.
- [28] XU K, ZHU J L, GUO N, et al. A novel 7-base pair deletion at a splice site in MS-2 impairs male fertility via premature tapetum degradation in common bean (*Phaseolus vulgaris* L.)[J]. *TAG Theoretical and Applied Genetics Theoretische Und Angewandte Genetik*, 2023, 136(3): 56.
- [29] ZHU X B, WANG H Y, LI Y Z, et al. A novel 10-base pair deletion in the first exon of GmHY2a promotes hypocotyl elongation, induces early maturation, and impairs photosynthetic performance in soybean[J]. *International Journal of Molecular Sciences*, 2024, 25(12): 6483.
- [30] CHEN C J, WU Y, LI J W, et al. TBtools-II: a "one for all, all for one" bioinformatics platform for biological big-data mining[J]. *Molecular Plant*, 2023, 16(11): 1733–1742.
- [31] QIN C, LI Y H, LI D L, et al. PH13 improves soybean shade traits and enhances yield for high-density planting at high latitudes[J]. *Nature Communications*, 2023, 14(1): 6813.
- [32] ZHANG C, XIE L, YU H, et al. The T2T genome assembly of soybean cultivar ZH13 and its epigenetic landscapes[J]. *Molecular Plant*, 2023, 16(11): 1715–1718.
- [33] FEKIH R, TAKAGI H, TAMIRU M, et al. MutMap+: genetic mapping and mutant identification without crossing in rice[J]. *PLoS One*, 2013, 8(7): e68529.
- [34] TAKAGI H, UEMURA A, YATEGASHI H, et al. MutMap-Gap: whole-genome resequencing of mutant F₂ progeny bulk combined with *de novo* assembly of gap regions identifies the rice blast resistance gene Pii[J]. *The New Phytologist*, 2013, 200(1): 276–283.
- [35] TAKAGI H, ABE A, YOSHIDA K, et al. QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations[J]. *The Plant Journal*, 2013, 74(1): 174–183.
- [36] ZHOU H K, TANG K Q, LI G, et al. A robust and rapid candidate gene mapping pipeline based on M2 populations[J]. *Frontiers in Plant Science*, 2021, 12: 681816.
- [37] YANG C J, ZHANG C, LU Y N, et al. The mechanisms of brassinosteroids' action: from signal transduction to plant development[J]. *Molecular Plant*, 2011, 4(4): 588–600.
- [38] LI R Q, XIA J X, XU Y W, et al. Characterization and genetic mapping of a Photoperiod-sensitive dwarf 1 locus in rice (*Oryza sativa* L.)[J]. *TAG Theoretical and Applied Genetics Theoretische Und Angewandte Genetik*, 2014, 127(1): 241–250.
- [39] SUO H C, MA Q B, YE K X, et al. Overexpression of AtDREB1A causes a severe dwarf phenotype by decreasing endogenous gibberellin levels in soybean [*Glycine max* (L.) Merr][J]. *PLoS One*, 2012, 7(9): e45568.
- [40] SUN J Q, HUANG S Y, LU Q, et al. UV-B irradiation-activated E3 ligase GmLPA1 modulates gibberellin catabolism to increase plant height in soybean[J]. *Nature Communications*, 2023, 14(1): 6262.