

DOI:10.11686/cyxb2024418

<http://cyxb.magtech.com.cn>

张志鹏, 蒋庆雪, 周昕越, 等. 转录组和蛋白组联合筛选饲用燕麦株高性状候选基因. 草业学报, 2025, 34(9): 147—161.

ZHANG Zhi-peng, JIANG Qing-xue, ZHOU Xin-yue, *et al.* Screening of candidate genes for plant height in forage oat (*Avena sativa*) through combined transcriptome and proteome analysis. Acta Prataculturae Sinica, 2025, 34(9): 147—161.

转录组和蛋白组联合筛选饲用燕麦株高性状候选基因

张志鹏, 蒋庆雪, 周昕越, 苗童, 唐俊, 仪登霞, 王学敏*, 马琳*

(中国农业科学院北京畜牧兽医研究所, 北京 100193)

摘要: 饲用燕麦是一种高产、优质且抗逆性强的饲草, 在我国饲草产业的发展中占据重要地位。本研究分别对饲用燕麦高株高(编号 972)与低株高(编号 1289)的极端材料取穗下茎节和节间组织进行高通量转录组测序(RNA-Seq)和蛋白组定量分析, 筛选差异表达基因(differentially expressed genes, DEGs)和差异表达蛋白(differentially expressed proteins, DEPs)。转录组分析筛选到 22762 个差异表达基因; 蛋白组分析获得 3934 个差异表达蛋白; 进一步的联合分析发现 1147 个差异表达基因(蛋白)重叠出现于转录组及蛋白组分析中。通过对重叠基因(蛋白)进行 GO 功能富集分析和 KEGG 信号通路分析, 发现很多基因被显著富集到与细胞生长、代谢和细胞壁形成的通路上。进一步结合转录因子分析, 筛选到 10 个饲用燕麦株高性状相关的候选基因。对候选基因进行 qRT-PCR 验证, 得到了测序结果的可靠性。此外, 对 10 个候选基因的组织表达特异性分析发现候选基因在燕麦茎、茎节中均有较高表达, 而在其他组织中表达量较低, 表明以上候选基因可能参与饲用燕麦的株高发育过程。综上, 通过转录组与蛋白组联合分析, 结合差异基因功能注释及转录因子分析, 筛选到 10 个饲用燕麦株高性状相关的候选基因。以上候选基因主要通过调控细胞生长、代谢和细胞壁发育参与饲用燕麦株高性状的形成。本研究为进一步探究饲用燕麦株高性状形成的分子机制奠定了基础, 并为今后饲用燕麦株高性状的生物育种提供了关键候选基因。

关键词: 饲用燕麦; 株高; 转录组; 蛋白组; 联合分析

Screening of candidate genes for plant height in forage oat (*Avena sativa*) through combined transcriptome and proteome analysis

ZHANG Zhi-peng, JIANG Qing-xue, ZHOU Xin-yue, MIAO Tong, TANG Jun, YI Deng-xia, WANG Xue-min*, MA Lin*

Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, Beijing 100193, China

Abstract: Forage oat (*Avena sativa*) is a high-yielding, high-quality, and stress-resistant forage that plays a significant role in China's forage industry. In this study, high-throughput transcriptome sequencing (RNA-Seq) and quantitative proteomic analysis were performed on stem nodes and internode tissues from high-stalk (No. 972) and low-stalk (No. 1289) forage oat varieties. The aim was to identify differentially expressed genes (DEGs) and differentially expressed proteins (DEPs) between these two varieties to discover plant height-related genes in *A. sativa*. We identified 22762 DEGs and 3934 DEPs between the two germplasm lines. Further integrated analyses revealed 1147 overlapping DEGs/DEPs in the transcriptome and proteome datasets. Gene Ontology (GO) enrichment and KEGG pathway analysis indicated that these 1147 overlapping genes/proteins were significantly

收稿日期: 2024-10-24; 改回日期: 2024-12-02

基金项目: 科技部、财政部国家科技资源共享服务平台项目-国家作物种质资源库(NCGRC-63)和财政部和农业农村部: 国家现代农业产业技术体系(CARS-34)资助。

作者简介: 张志鹏(1999—), 男, 内蒙古乌兰察布人, 在读硕士。E-mail: 224998720@qq.com

* 通信作者 Corresponding author. E-mail: wangxuemin@caas.cn; malin@caas.cn

enriched in pathways associated with cell growth, metabolism, and cell wall formation. Among these DEGs/DEPs, 10 candidate transcription factors were identified through transcription factor analysis, and their transcription profiles were validated using quantitative real-time polymerase chain reaction (qRT-PCR) analyses. The transcription patterns of the ten candidate genes were highly consistent with those predicted from the transcriptomic and proteomic data, confirming the reliability of the sequencing results. Tissue-specific analyses of their transcriptional profiles showed that these ten candidate genes exhibited higher transcript levels in stems and nodes but lower levels in other tissues, suggesting that they play roles in regulating plant height in forage oat. In summary, through integrated transcriptomic and proteomic analyses, along with differential gene function annotation and transcription factor analyses, we identified ten candidate genes related to plant height in forage oat. These candidate genes primarily regulate processes such as cell growth, metabolism, and cell wall development, contributing to plant height formation in forage oat. These findings provide a foundation for further exploration of the molecular mechanisms underlying plant height and offer key candidate genes for forage oat breeding programs.

Key words: forage oat; plant height; transcriptome; proteome; combined analysis

随着我国畜牧产业的快速发展,养殖户对饲草的需求量日益增大。作为一种具有优异营养品质和良好抗逆性的高产量饲草,饲用燕麦(*Avena sativa*)成为解决我国北方冬春牧场饲草供应不足的关键饲草之一。近年来,随着我国“粮改饲”等政策实施,饲用燕麦的种植面积不断扩大,对优良饲用燕麦品种的需求也日益迫切,进而对饲用燕麦的育种工作也提出更高的要求^[1-3]。株高是决定饲用燕麦产量的重要农艺性状之一,它与多个产量性状密切相关,与种植密度共同作用影响饲用燕麦的产量。因此,筛选饲用燕麦株高性状的候选基因,探究饲用燕麦株高性状的调控机制,对提高饲用燕麦产量具有重要意义。

株高作为直接影响植株产量的重要性状,既受到外部环境的影响,又受到内部遗传因素的控制。目前在燕麦中被发掘的株高相关位点共8个,分别为 *Dw1*~*Dw8*^[4],但仅有 *Dw6*、*Dw7* 和 *Dw8* 通过不同的试验方法被验证。Brown等^[5]首先在半矮株系中发现 *Dw6*。徐颖红^[6]又利用简单重复序列(simple sequence repeats, SSR)标记和转录组数据开发(kompetitive allele-specific PCR, KASP)标记对 *Dw6* 进行精细定位,进一步研究发现 *Dw6* 主要通过减少前3节茎的长度来降低株高。Milach等^[7]利用限制性片段长度多态性(restriction fragment length polymorphism, RELP)标记对 *Dw7* 进行定位,发现 *Dw7* 减少燕麦所有节间长度和节间数,降低植株株高。Morikawa^[8]发现 *Dw8* 通过减少燕麦所有节间长度而不影响植株的节间数来降低燕麦株高。然而,由于受到可用标记数量的限制,控制以上8个株高相关位点的基因除了 *Dw6* 研究较深入,其他均未被克隆。大量研究发现,植物节间细胞大小的变化是导致株高变化的主要因素。张孝波^[9]利用石蜡切片对野生型水稻(*Oryza sativa*)和半矮化突变体 *sd38* 抽穗期的倒二节间进行纵切比较发现,与野生型相比,突变体 *sd38* 倒二节间中部薄壁细胞长度及第一、二节间面积显著降低,表明细胞长度的降低可能是突变体 *sd38* 半矮化的原因。徐颖红^[6]对燕麦成熟期 *Dw6* 的近等位基因系的茎节纵向切片观察发现,矮秆株系穗下茎节细胞长度显著低于高秆株系,推测 *Dw6* 是通过缩短细胞长度来降低燕麦株高。Yamamuro等^[10]通过显微镜观察水稻野生型和突变体植株成熟期的节间切片,发现油菜素内酯(brassinolide, BR)信号受体 *OsBRI1* 的突变导致细胞变化,造成水稻茎秆伸长受阻,降低水稻的株高。Zhang等^[11]通过对水稻 P450 96B 家族基因 *sd17* 进行研究发现, *sd17* 突变体中与细胞分裂相关的基因表达受到干扰,导致茎秆中细胞数目减少进而使植株矮化。

目前,饲用燕麦株高性状相关基因的发掘、研究和利用相比其他作物少之又少,具体的分子调控机理也有待研究。随着后基因组学的快速发展,转录组学和蛋白质组学测序的研究方法和技术已经日趋成熟,并在植物领域得到了广泛的运用。

因此,本研究以饲用燕麦株高表型极端的品种为试验材料,运用 Illumina 高通量测序平台的转录组测序技术和 TMT(Tandem Mass Tags)标记蛋白技术,对饲用燕麦株高性状形成的关键部位进行测序,联合分析筛选出差异表达基因(differentially expressed genes, DEGs)和差异表达蛋白(differentially expressed proteins, DEPs),进一

步对差异共表达基因(蛋白)进行 GO、KEGG 分析和注释,结合转录因子及组织表达特异性分析,挖掘与饲用燕麦株高性状相关的候选基因,为揭示饲用燕麦调控株高性状的分子机制奠定基础。

1 材料与方法

1.1 植物材料与表型调查

田间试验于 2023 年 3—6 月进行,试验地位于北京市昌平区,当地属暖温带半湿润半干旱季风气候,年平均气温 10~12℃,降水量 644 mm,无霜期 180~200 d,土壤 pH 约为 7.64。不同饲用燕麦种质资源于 2023 年 3 月 8 日播种,每份种质资源播种一行,单粒点播,株行距 30 cm,播深 3 cm 左右,灌溉方式为滴灌。

待植株进入乳熟期,对不同饲用燕麦种质资源的株高进行测定,株高为茎基部到主穗顶端的距离。根据调查结果选取两份极端材料,即株高最高(编号 972)与株高最低(编号 1289)的饲用燕麦种质资源(表 1),测定其主茎穗长、节数及节长等农艺性状,并对极端材料的穗下茎与茎节进行取样,设置 3 个生物学重复,委托武汉贝纳科技有限公司进行转录组测序和蛋白组测序以及多组学联合分析。

表 1 材料具体信息

Table 1 Materials details

编号 Code	名称 Name	材料来源 Source of material
972	Dookie 10	澳大利亚,维多利亚州 Victoria, Australia
1289	X61-7	美国,佛罗里达州 Florida, United States

1.2 RNA 提取与转录组测序

由武汉贝纳科技有限公司进行总 RNA 提取,质检合格后,进行 cDNA 文库构建,并在 Illumina Novaseq 6000 平台进行双端测序获得原始数据(raw reads);经过 fastp 校正,去除低质量的读数,利用 Star 将过滤后的转录组序列与饲用燕麦参考基因比对,获得高质量的序列数据(clean reads)。

1.3 蛋白质提取与蛋白组检测

使用总蛋白提取试剂盒(ZD412,北京庄盟国际生物基因科技有限公司)和 Bradford 蛋白浓度测定试剂盒(ZD301K,北京庄盟国际生物基因科技有限公司)对样本进行蛋白的提取和浓度定量。取等量的样品利用 TMT 稳定同位素技术标记肽段,在质谱仪(Orbitrap™ Ascend™ MultiOmics Tribrid,美国)中进行 LC-MS 分析。采用 DDA 模式采集图谱,获得质谱原始数据。

1.4 DEGs、DEPs 筛选及功能富集

使用 DESeq 2 对基因表达量数据进行标准化分析和差异分析,差异表达基因(DEGs)筛选的阈值为 $\text{padj} < 0.05$ 且 $|\log_2\text{Fold change}| > 1$ 。利用 t 检验对两组样本间蛋白表达量进行检验,差异表达蛋白(DEPs)筛选的阈值为 $\text{padj} < 0.05$ 且 $|\log_2\text{Fold change}| > 1$ 。合并 DEGs 和 DEPs,并分析转录组数据和蛋白质组数据之间的相关性,采用 clusterProfiler 软件进行 GO 功能富集分析和 KEGG 信号通路富集分析, P 值的取值为 $[0,1]$ 。

1.5 转录因子(transcription factors, TF)分析

将差异基因与转录因子数据库(<http://planttfdb.cbi.pku.edu.cn/index.php?sp=Vvi>)进行比对,获得差异表达的转录因子并对转录因子家族进行归类与分析。

1.6 候选基因组织特异性表达分析

根据 DEGs、DEPs 富集以及转录因子分析结果,筛选在株高极端材料间的差异表达候选基因,并采用实时荧光定量 PCR(qRT-PCR)分析候选基因在饲用燕麦苗期地下和地上组织;拔节期根、茎、叶、茎节;孕穗期根、茎、叶、茎节、穗;抽穗期根、茎、叶、茎节、穗等组织中的表达量,以 *AsActin* 为内参基因,利用 AlleleID 6.0 设计引物(表 2),引物由北京生工生物工程股份有限公司合成,根据北京庄盟国际生物基因科技有限公司的反转录试剂盒(HiScript IV 1st Strand cDNA Synthesis Kit)说明书进行 cDNA 第一条链的合成,采用北京庄盟国际生物基因科技有限公司荧光定量 PCR 试剂盒(2×HQ SYBR qPCR Mix)进行实时荧光定量 PCR(qRT-PCR)试验,反应体系为 20 μL:cDNA 模板 1 μL,2×HQ SYBR qPCR Mix (Low ROX) 10 μL,正、反向引物各 0.5 μL,加 ddH₂O 补至 20 μL。PCR 反应程序为:95℃预变性 5 min,95℃变性 10 s,60℃退火 30 s,40 个循环。 $2^{-\Delta\Delta CT}$ 法^[12]计算差异基因的表达量。每组样品设置 3 次重复。

表2 实时荧光定量 PCR 引物设计序列

Table 2 Primer sequences of real-time fluorescence quantitative PCR

基因 ID Gene ID	正向引物 Forward primer (5'—3')	反向引物 Reverse primer (5'—3')
AVESA. 00010b. r1. 2CG3384710	CATCACCGCCAACATCACC	GCAGCCTCCACGATCTCC
AVESA. 00010b. r1. 3CG0669440	AAGAAGTTGGTGGTTATTGG	CTGCTTCCTTACCTCTCC
AVESA. 00010b. r1. 4AG2537620	CGGACCTACAACCAGAACC	CAGCCCAAAGTGCCTCTC
AVESA. 00010b. r1. 5DG0437330	CGGTCTTCTTGCGAATGG	CGATGTACTCCACGAAACC
AVESA. 00010b. r1. 3CG2847600	ATGCTCTTCACCGTCTCC	CTGGTTGATGTAGTCCTTGG
AVESA. 00010b. r1. 4AG1741330	GCATAGCGAGCAGTGAAGAG	ACACGGAGATCAGCAGAGC
AVESA. 00010b. r1. 2CG1087750	CTCCTCCTCTGCTCGTC	GTGCCCTTGAACCTGTGG
AVESA. 00010b. r1. 2DG0384520	TGTGCTACGGGAGAAACG	AACTTGATGCCGCTATTCTG
AVESA. 00010b. r1. 2CG0426050	ACCTTCACGCTCAACTTCTCC	GCTCTCCACGCAGTTCTCC
AVESA. 00010b. r1. 4AG0885780	AACTTCCCGTGCTCTGATCC	GTCGTCTCGTCGGCTTCC

1.7 数据处理

利用 Microsoft Office Excel 2014 进行数据统计与处理, GraphPad Prism 8.0 进行作图, 结果以平均值±标准差表示。

2 结果与分析

2.1 株高极端材料的表型

田间表型观测发现, 编号为 972 的种质株高最高, 为(125. 60±1. 94) cm; 而编号为 1289 的种质株高最矮, 为(54. 20±1. 25) cm(图 1A)。进一步比较两份株高极端种质的穗长, 发现编号为 972 的种质穗长为(32. 50±1. 86) cm; 编号为 1289 的种质穗长为(11. 30±0. 75) cm(图 1B)。编号为 972 的主茎节数为(6±1); 编号为 1289 的主茎节数为(5±1)(图 1C)。通过比较两份株高极端种质的节间长度, 发现 972 种质的每一节间长度都大于 1289 种质(图 1D)。

2.2 转录组数据检测

经 Illumina Novaseq 6000 平台进行转录组测序(表 3), 质控处理后每个样品均获得 6 G 以上的过滤后序列, 碱基质量>30(Q₃₀)的占比大于 93. 99%, GC 含量为 51. 16%~53. 21%。测序数据与燕麦 Sang 参考基因组比对后, 发现约有 83. 39%~85. 83% 的过滤后数据能够特异地匹配至参考基因组。以上结果表明转录组测序数据可信度高, 可以满足后续分析。

2.3 蛋白质数据检测

通过 TMT 标记技术对样品进行蛋白质定量分析, 共获得了 10065 个蛋白。经过蛋白质理论数据搜索, 发现 9619 个蛋白具有定量信息。通过谱图分析, 发现共有 65098 条肽段。对鉴定到的肽段序列长度进行统计分析, 发现 99% 的肽段长度为 7~27 aa, 其中, 肽段分布的高峰区为 8~13 aa(图 2A)。对所有鉴定到的蛋白质进行统计, 发现蛋白相对分子质量主要为 10~60 kDa, 占总蛋白的 72. 52%。分子质量为 0~10 kDa 和 60~100 kDa 的蛋白分别占总蛋白的 0. 82% 和 18. 17%, 大于 100 kDa 的蛋白约占总蛋白的 8. 49%(图 2B)。

2.4 DEGs 及 DEPs 筛选

比较编号 972 与 1289 种质的转录组数据, 共鉴定出 22762 个差异表达基因(DEGs), 其中 8550 个基因在种质 972 中的表达量高于种质 1289, 而 14212 个基因在种质 972 中的表达量低于种质 1289。对两组样本的蛋白表达量进行 *t* 检验, 共获得 3934 个差异表达蛋白(DEPs), 其中 2083 个蛋白在种质 972 中的表达丰度高于种质 1289, 1851 个蛋白在种质 972 中的表达丰度低于种质 1289(图 3)。

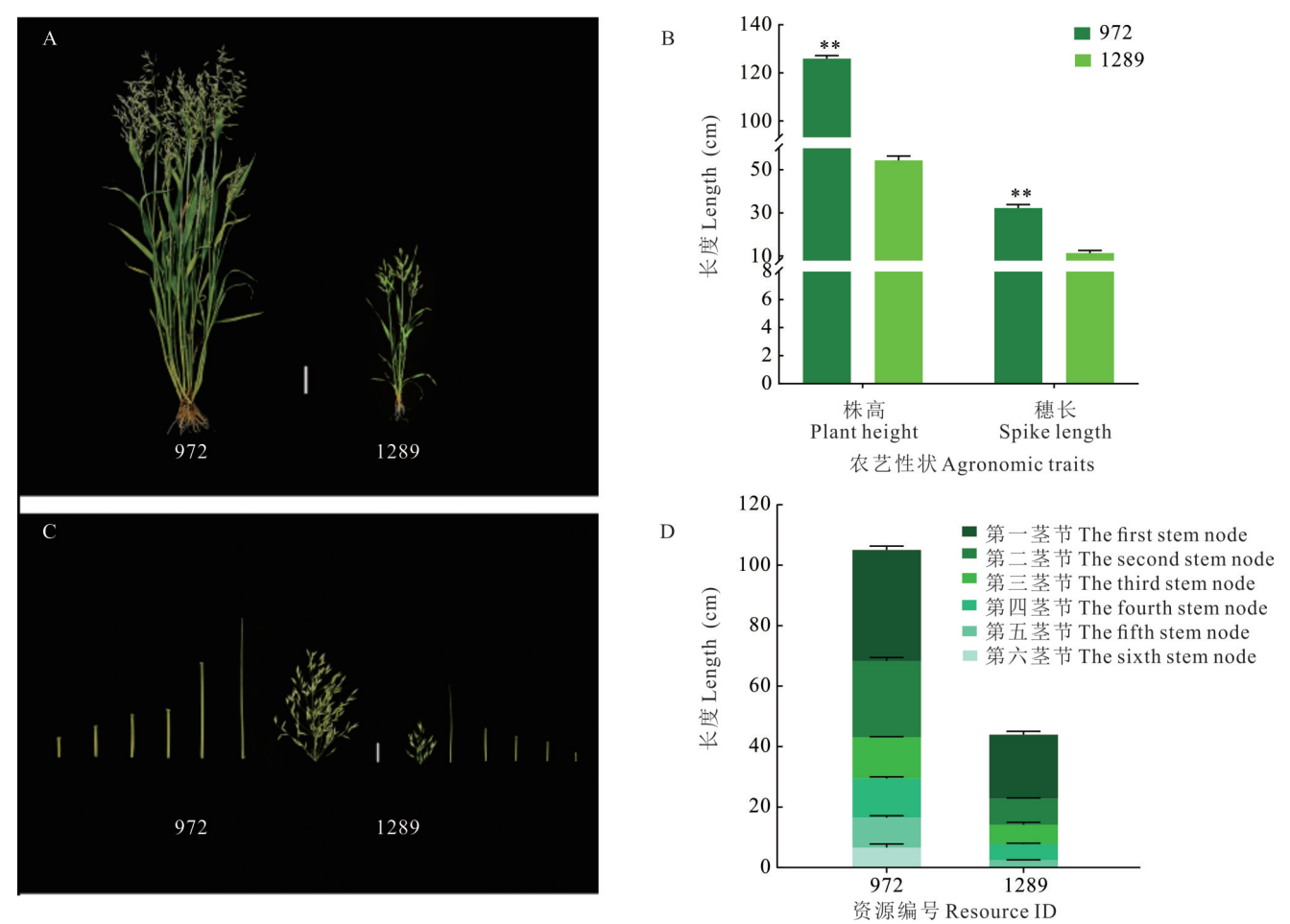


图 1 株高极端燕麦材料的表型分析

Fig. 1 Phenotypic observations of extreme plant height materials

972: 高株高种质 High plant height germplasm; 1289: 低株高种质 Low plant height germplasm. A: 成熟期的植株 (标尺=10 cm) Plants at maturity (scale bar=10 cm); C: 节数、各茎节长度和穗长表型 (标尺=5 cm) Number of nodes, length of each stem node and spike length phenotype (scale bar=5 cm). **: $P<0.01$.

表 3 转录组测序数据质控分析

Table 3 Quality test of transcriptome sequencing data

样品名称 Sample	过滤后数据 Clean reads (M)	过滤后序列 Clean bases (G)	碱基质量>30的 占比 Q ₃₀ (%)	GC 含量 GC content (%)	唯一比对 (对比率) Uniquely mapped (mapping ratio, %)	多方比对 (对比率) Multiple mapped (mapping ratio, %)
H-1	43.09	6.11	94.99	53.21	83.39	14.16
H-2	43.05	6.13	95.07	52.98	83.80	14.22
H-3	43.08	6.13	94.84	53.08	83.92	14.09
L-1	43.13	6.33	94.06	51.53	85.83	11.67
L-2	44.62	6.57	94.58	51.38	85.79	11.62
L-3	42.36	6.23	93.99	51.16	85.83	11.60

2.5 转录组与蛋白组联合分析

对转录组数据和蛋白组数据进行联合分析,共有 1147 个基因是重叠的(图 4A)。通过比较分析转录水平和对应蛋白质水平的变化趋势,可以帮助更好地理解生物体内的某些机制。转录组与蛋白组之间除了一对一的相互关系外,还存在其他更为复杂的调控关系,统计分析基因及其对应的蛋白表达量并绘制散点图(图 4B),结果显示转录组和蛋白质组之间呈正相关($R=0.14$),进一步比较转录组和蛋白质组中差异表达的转录本和蛋白个数

分布情况(图4C)。结果显示,饲用燕麦株高极端材料中484个基因在种质972中的表达量高于种质1289,而663个基因在种质972中的表达量低于种质1289。533个蛋白在种质972中的表达丰度高于种质1289,614个蛋白在种质972中的表达丰度低于种质1289。

为了比较分析各差异表达基因和蛋白的功能,通过GO功能富集分析将与DEGs表达趋势相同的DEPs进行分析,结果表明:当DEGs和DEPs共同上调时,其功能主要富集在细胞壁大分子分解代谢过程、脱落酸激活的信号通路、谷胱甘肽代谢过程、细胞质、谷胱甘肽转移酶活性、信号受体活性等(图5A);当DEGs和DEPs共同下调时(图5B),其功能主要富集在光合作用、水解酶活性、翻译负调控、谷胱甘肽转移酶活性等。为了关联分析DEGs-DEPs在各信号和代谢通路上的联系,对这些表达趋势相同基因的KEGG通路进行了比较。当DEGs和DEPs共同上调时,主要富集在谷胱甘肽代谢、氨基酸的生物合成、苯丙素的生物合成、氨基酸糖和核苷酸糖代谢、次级代谢物的生物合成和其他代谢途径中(图6A)。当DEGs和DEPs共同下调时(图6B),主要通路为碳代谢、苯丙素的生物合成、次级代谢物的生物合成和其他代谢途径中。

进一步比较分析各差异表达基因和蛋白的功能,通过GO功能富集分析方法对与DEGs表达趋势不同的DEPs进行分析,在324个转录本表达量下调,对应的蛋白表达量上调的相关因子中,其分子功能途径主要富集在过氧化物酶活性、葡聚糖内-1,3- β -D-葡萄糖苷酶活性和DNA聚合酶加工因子活性;在细胞组分途径中,这些相关因子主要富集在胞外体、细胞壁、质膜、MCM复合体等条目中;在生物学过程中,这些相关因子主要富集在DNA复制启动、脂质运输、木聚糖分解过程、过氧化氢分解代谢过程以及碳水化合物代谢过程条目上(图7A)。而在194个转录本表达量上调,对应的蛋白表达量下调的相关因子中,在分子功能途径中这些相关因子主要富集在与细胞壁成分相关的条目上,包括D-硫代阿糖1-脱氢酶活性、纤维素合成酶活性、mRNA结合等;在细胞组分途径中,这些基因主要富集在质膜和高尔基体条目;在生物学过程中,这些基因主要富集在植物次生细胞壁的生物生成、植物的细胞壁组织或生物生成、D-木糖代谢过程、纤维素生物合成过程等条目上(图7B)。这些通路多与细胞

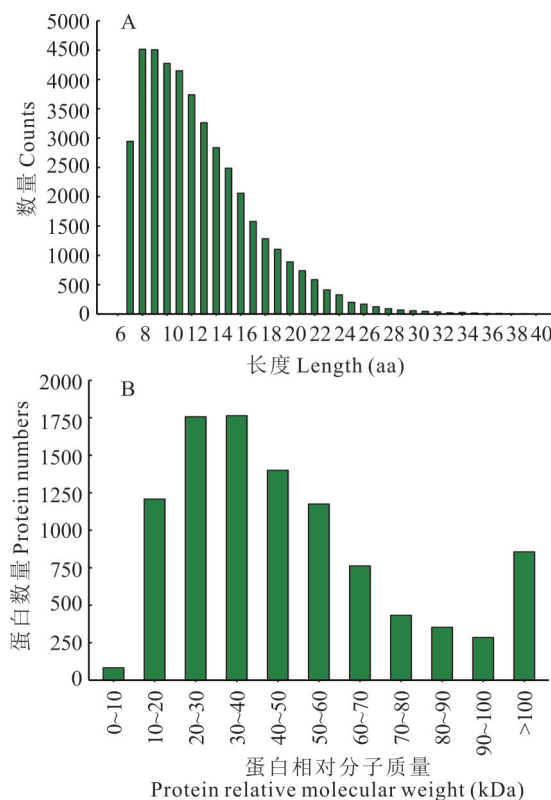


图2 蛋白质鉴定基本信息

Fig. 2 Basic information for protein identification

A: 肽段长度范围分布 Distribution of peptide length ranges; B: 蛋白质分子量的分布 Distribution of protein molecular weights.

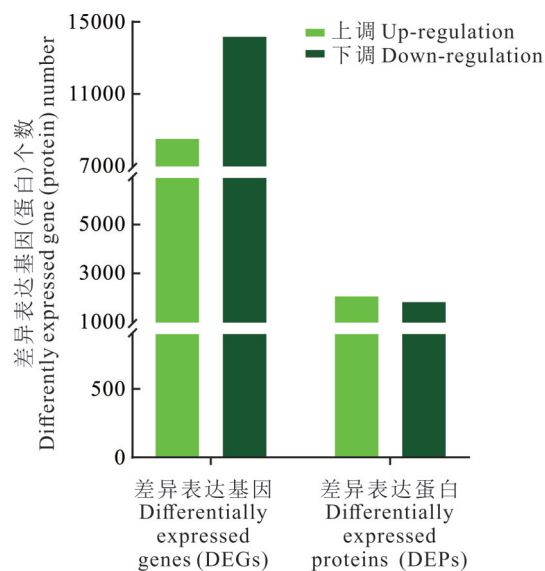


图3 极端材料差异表达基因和差异表达蛋白上调和下调基因数目

Fig. 3 Number of differentially expressed genes and differentially expressed proteins up- and down-regulated genes in extreme materials

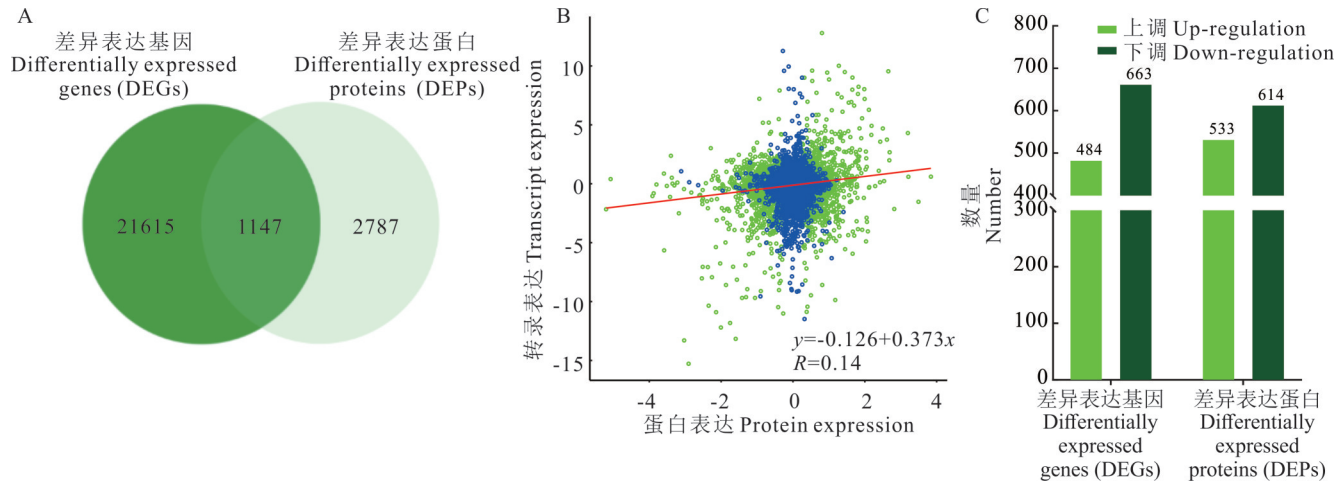


图 4 转录组和蛋白组关联分析

Fig. 4 Transcriptome and proteome association analysis

A: 差异表达转录组与蛋白组 Venn 图 Venn diagram of differentially expressed transcriptome and proteome; B: 转录组与蛋白组表达相关性分析 Transcriptome and proteome expression correlation analysis; C: 差异表达转录组与蛋白组数目统计 Statistics of the number of differentially expressed transcriptome and proteome.

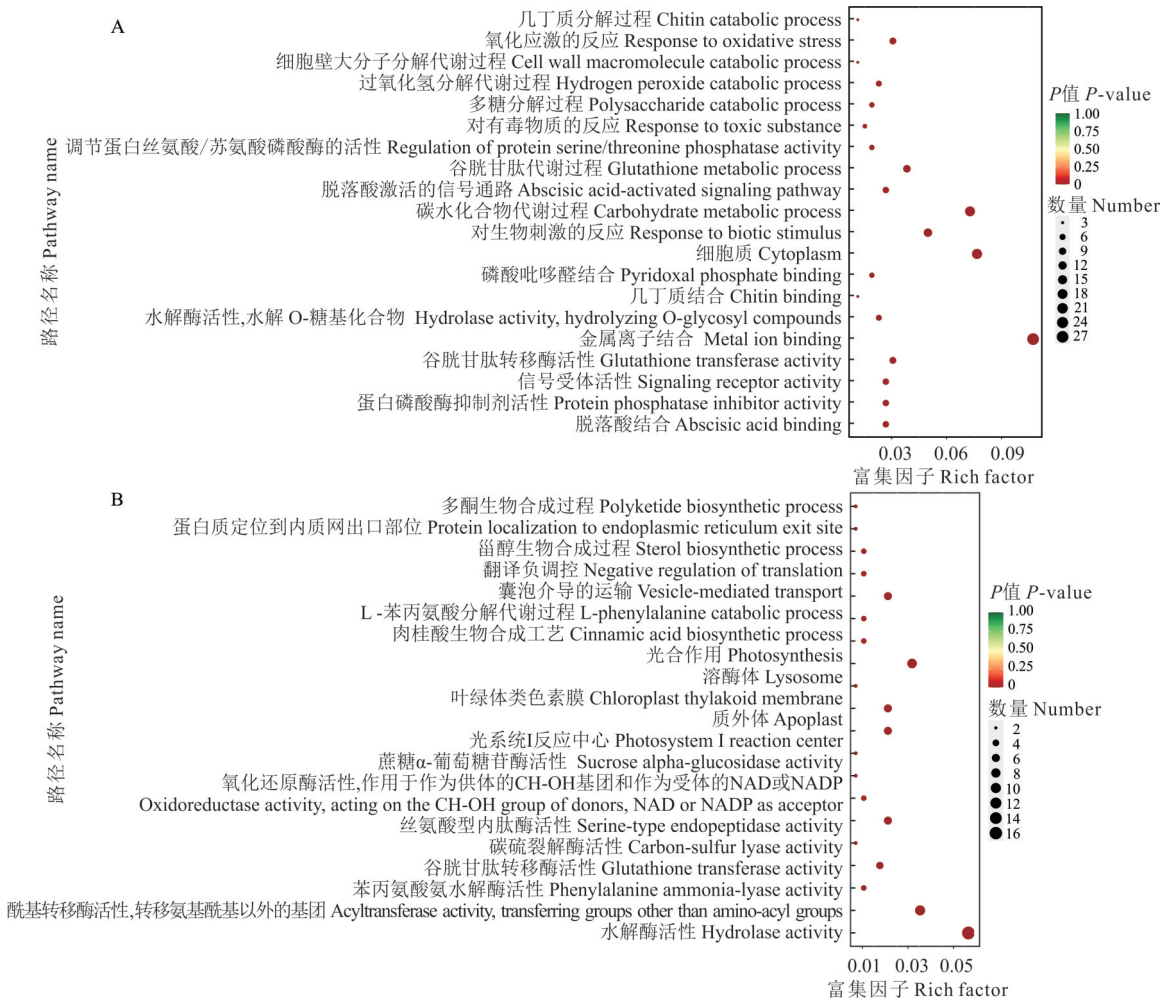


图 5 相同变化趋势相关因子的 GO 富集分析

Fig. 5 GO enrichment analysis of correlators with the same trend change

A: 转录本与蛋白共同上调的 GO 富集分析 GO enrichment analysis of transcripts and proteins jointly up-regulated; B: 转录本与蛋白共同下调的 GO 富集分析 GO enrichment analysis of transcripts and proteins jointly down-regulated.

的形成和发育相关。通过对不同变化趋势相关因子进行KEGG代谢通路富集分析,结果表明:324个转录本表达量下调,对应的蛋白表达量上调的这些相关因子富集通路为:萜类的生物合成、氨基酸糖和核苷酸糖代谢、苯丙素的生物合成、谷胱甘肽代谢、DNA复制和各种次级代谢物的生物合成中(图8A);194个转录本表达量上调,对应的蛋白表达量下调的相关因子中,碳代谢、氨基酸糖和核苷酸糖代谢、木质素、亚木质素和蜡的生物合成、半萜和三萜类化合物的生物合成、次级代谢物生物合成和其他代谢途径为主要通路(图8B)。

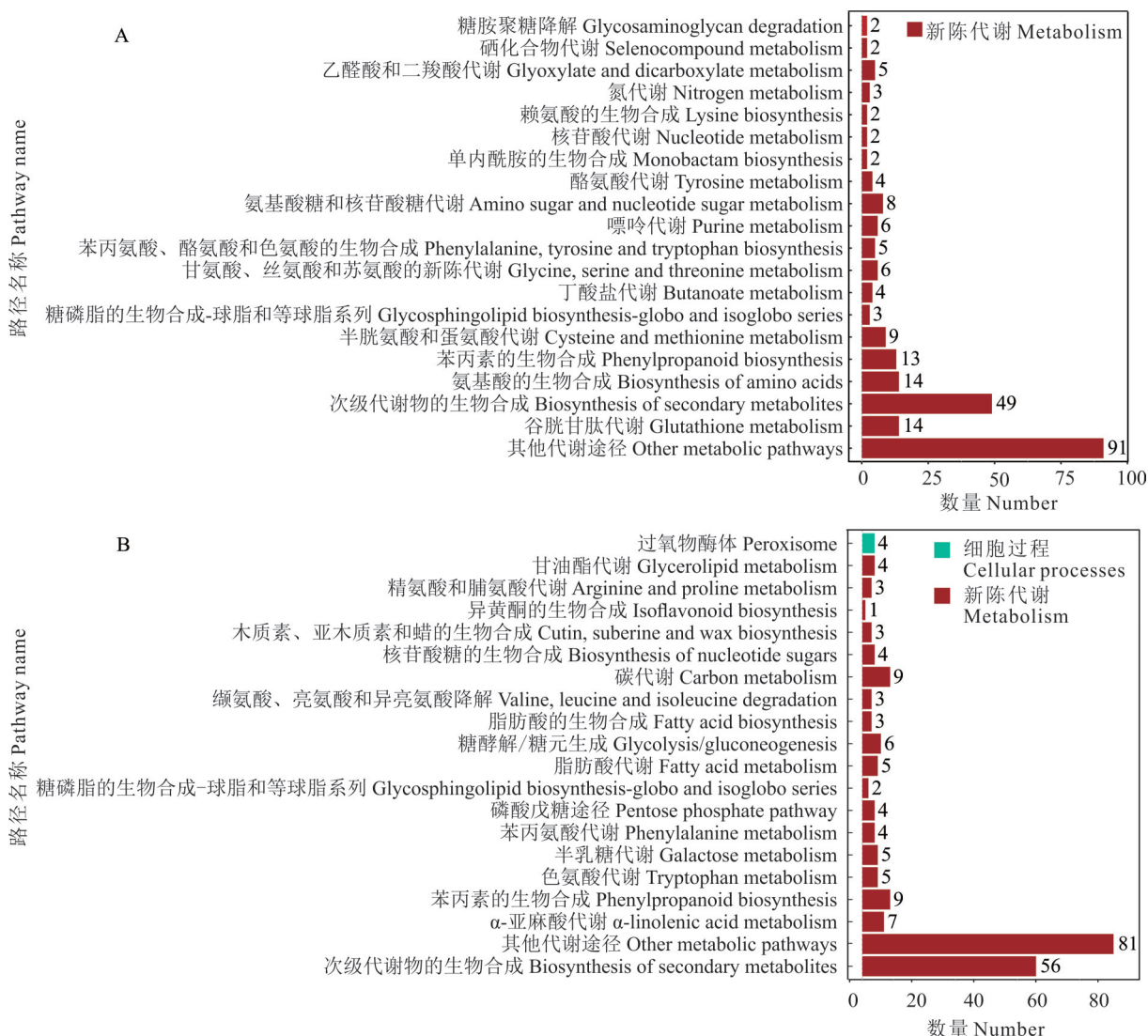


图6 相同变化趋势相关因子的KEGG通路分析

Fig. 6 KEGG pathway analysis of correlators with the same change trend

A: 转录本与蛋白共同上调的KEGG通路分析 Analysis of KEGG pathway jointly up-regulated by transcripts and proteins; B: 转录本与蛋白共同下调的KEGG通路分析 Analysis of KEGG pathway jointly down-regulated by transcripts and proteins.

2.6 转录因子筛选及分类

参考植物转录因子数据库 Plant TFDB(<http://planttfdb.cbi.pku.edu.cn/>),对1147个差异基因(蛋白)进行分析(表4),共鉴定到10个家族的12个转录因子,转录本与蛋白共同上调的差异转录因子1个,属于Nin-like家族转录因子;转录本与蛋白共同下调的差异转录因子5个,属于Trihelix、FAR1和ERF家族转录因子;转录本下调、蛋白上调的差异转录因子有3个,分别属于TCP、bHLH、NF家族转录因子;转录本上调、蛋白下调的差异转录因子有3个,分别属于Whirly、NAC、B3转录因子家族。

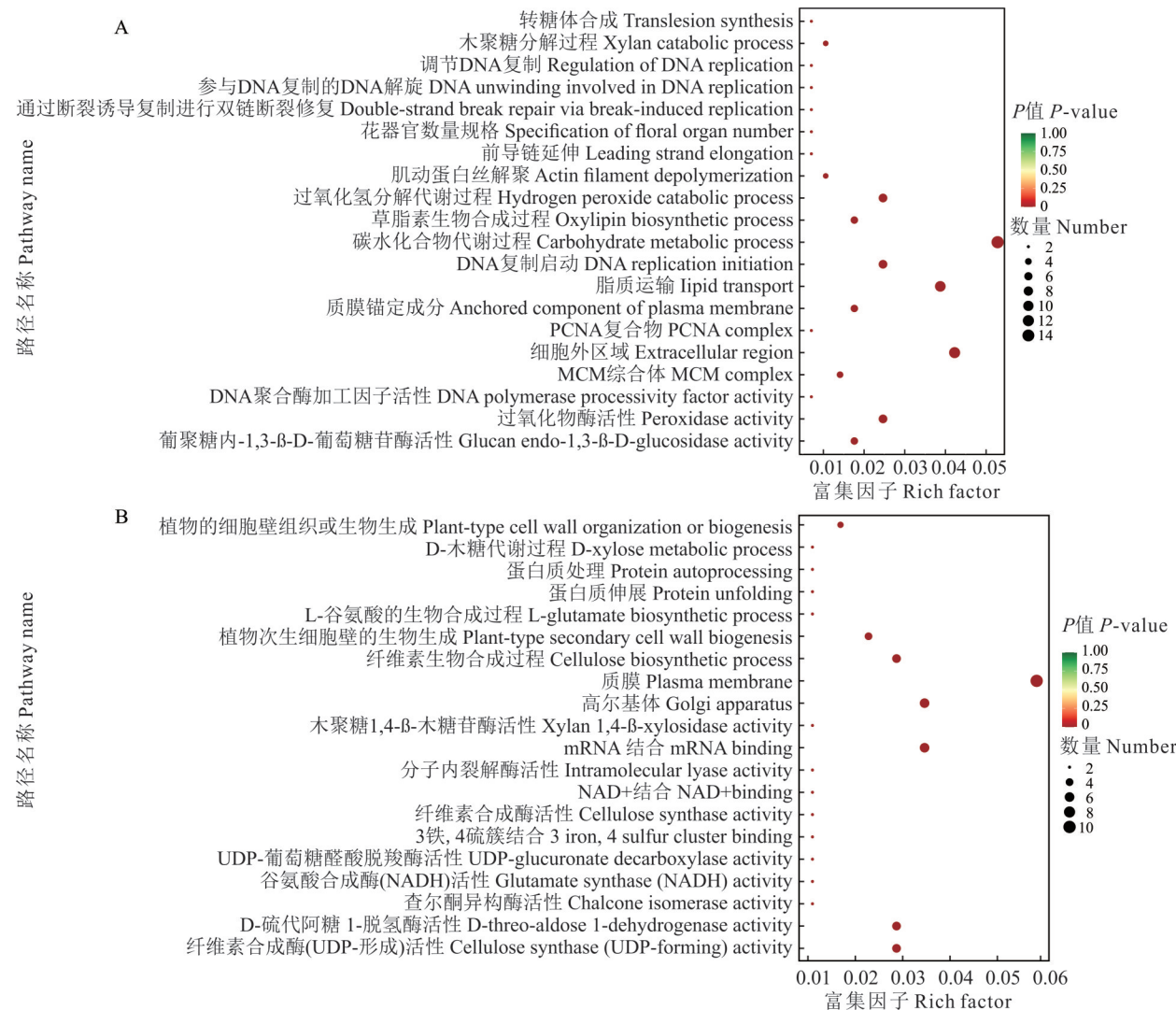


图 7 不同变化趋势相关因子的 GO 富集分析

Fig. 7 GO enrichment analysis of correlators with different trends

A: 转录本下调、蛋白上调的 GO 富集分析 GO enrichment analysis of transcript down-regulation and protein up-regulation; B: 转录本上调、蛋白下调的 GO 富集分析 GO enrichment analysis of transcript up-regulation and protein down-regulation.

2.7 株高性状候选基因筛选与 qRT-PCR 验证

通过转录组与蛋白组的联合分析,对可能调控株高性状的通路进行筛选,结合转录因子的分析,共筛选到 10 个与饲用燕麦株高性状相关的候选基因(表 5)。对筛选的 10 个饲用燕麦株高性状候选差异基因进行 qRT-PCR 验证(图 9),结果表明,基因相对表达量变化趋势与转录组测序结果一致,说明转录组结果准确可靠。

2.8 组织特异性表达分析

进一步分析筛选到的饲用燕麦株高性状候选基因的组织表达特异性(图 10),候选基因 AVESA. 00010b. r1. 2CG1087750、AVESA. 00010b. r1. 2DG0384520、AVESA. 00010b. r1. 4AG0885780 在苗期到抽穗期的茎与茎节均有表达,在拔节期时的茎与茎节表达量较高,其次是在抽穗期的茎组织与拔节期的叶片表达量较高,而在其他时期的组织表达量较低;候选基因 AVESA. 00010b. r1. 3CG0669440、AVESA. 00010b. r1. 4AG2537620 在拔节期的茎组织表达量最高,其次在拔节期和孕穗期的茎节表达量较高,在其他时期的组织中表达量较低;候选基因 AVESA. 00010b. r1. 5DG0437330 在抽穗期的茎和茎节组织表达量较高,其次在拔节期的茎与茎节组织表达量较高,在其他时期的组织均有表达但相对较低;候选基因 AVESA. 00010b. r1. 2CG0426050 和 AVESA. 00010b. r1. 3CG2847600 在孕穗期的茎与茎节表达量较高,其次在穗部组织表达量较高,其他时期的组

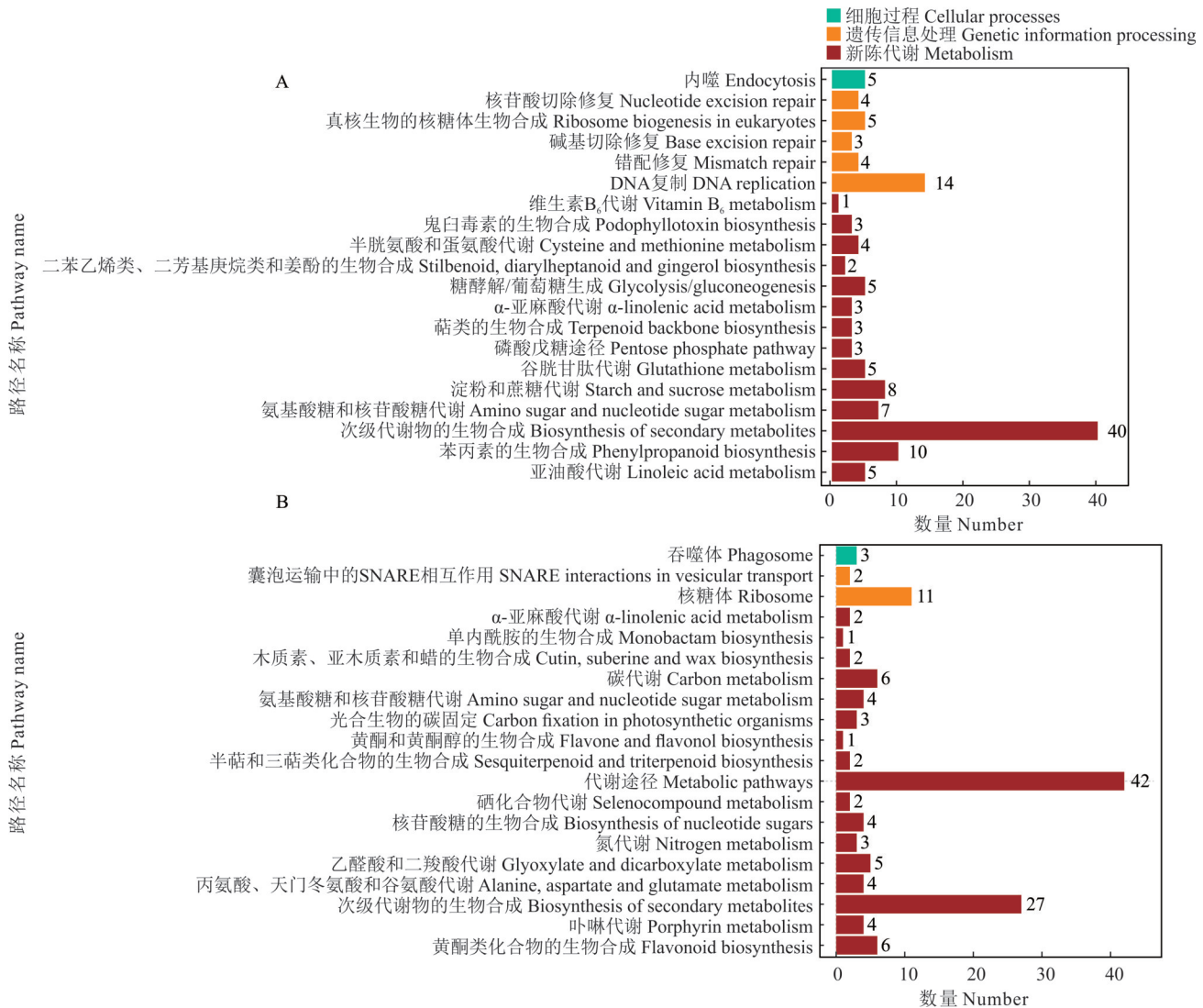


图8 不同变化趋势相关因子的KEGG通路分析

Fig. 8 KEGG pathway analysis of different trend correlators

A: 转录本下调、蛋白上调的KEGG通路分析 Analysis of KEGG pathway with transcript down-regulation and protein up-regulation; B: 转录本上调、蛋白下调的KEGG通路分析 Analysis of KEGG pathway with transcript up-regulation and protein down-regulation.

织均有表达但表达量较低;候选基因 AVESA. 00010b. r1. 2CG3384710在拔节期的茎与茎节表达量较高,在其他时期的组织表达量较低。候选基因 AVESA. 00010b. r1. 4AG1741330在各时期茎节组织表达量较高,然后是茎和穗部组织。上述结果表明,这些候选基因可能通过在茎、茎节中高表达,参与饲用燕麦的株高发育过程。

3 讨论

3.1 转录组与蛋白组联合分析

应用转录组学技术可以获取样本中转录本的基本信息,对所研究的植物在转录水平的调控规律有大致地了解。然而,通过转录组测序的数据仅能

表4 转录因子分类

变化趋势	转录因子家族	数量
Change trend	Transcription factor family	Number
转录本与蛋白共同上调 Transcripts and proteins jointly up-regulated	Nin-like	1
转录本与蛋白共同下调 Transcripts and proteins jointly down-regulated	Trihelix	3
	FAR1	1
	ERF	1
转录本下调、蛋白上调 Transcript down-regulation and protein up-regulation	TCP	1
	bHLH	1
	NF	1
转录本上调、蛋白下调 Transcript up-regulation and protein down-regulation	Whirly	1
	NAC	1
	B3	1

表 5 候选基因

Table 5 The candidate genes

基因 ID Gene ID	注释 Exegesis	基因差异倍数 Gene fold change	趋势 Tendencies	蛋白差异倍数 Protein fold change	趋势 Tenden- cies
AVESA.00010b.r1.2CG3384710	BHLH 结构域蛋白质 BHLH domain-containing protein	2.27	下调 Down	1.92	下调 Down
AVESA.00010b.r1.3CG0669440	GIDA 葡萄糖抑制分裂蛋白 GIDA glucose inhibited di- vision protein	4.41	上调 Up	1.31	上调 Up
AVESA.00010b.r1.4AG2537620	糖苷水解酶家族 17 Glycosyl hydrolases family 17	4.14	上调 Up	2.18	上调 Up
AVESA.00010b.r1.5DG0437330	未定性蛋白 Uncharacterized protein	4.85	下调 Down	1.20	上调 Up
AVESA.00010b.r1.3CG2847600	色氨酸氨基转移酶相关蛋白 Tryptophan aminotrans- ferase related protein	5.65	下调 Down	3.59	下调 Down
AVESA.00010b.r1.4AG1741330	过氧化物酶 Peroxidase	4.69	上调 Up	2.50	上调 Up
AVESA.00010b.r1.2CG1087750	γ -硫基家族 γ -thionin family	3.70	下调 Down	1.28	上调 Up
AVESA.00010b.r1.2DG0384520	糖基水解酶家族 17 Glycosyl hydrolases family 17	4.70	上调 Up	1.85	上调 Up
AVESA.00010b.r1.2CG0426050	谷氨酸脱羧酶 Glutamate decarboxylase	3.02	下调 Down	1.25	上调 Up
AVESA.00010b.r1.4AG0885780	乙烯反应性转录因子 ERF109 Ethylene-responsive transcription factor ERF109	9.04	下调 Down	1.89	上调 Up

从 RNA 水平对植物的分子调控有一个初步了解。想要更深层次地了解影响植物生长发育过程的分子机制相关信息,还需要蛋白组数据的结果来作为补充^[13]。目前已有许多研究采用转录和蛋白相结合的方法进行分析。罗京^[14]通过对矮秆油菜(*Brassica napus*)品系 DW871 和对照高秆品系 HW871 进行转录组和蛋白组联合分析,发现 24 个在转录和蛋白水平表达趋势相同的差异基因,进一步通过 GO 注释分析,结果表明,相关筛选基因被注释到分子调节功能等 12 个条目中。通过对激素处理后的玉米(*Zea mays*)幼苗进行转录组和蛋白质组联合分析,表明油菜素内酯信号转导对玉米光合作用产生影响^[15]。Lin 等^[16]通过对烟草(*Nicotiana tabacum*)幼苗的转录组和蛋白质组进行综合分析,以研究烟草幼苗在不同硼供应水平下基因/蛋白质的表达变化,揭示了烟草在不同硼供应水平的动态表达变化。

为筛选饲用燕麦株高性状的候选基因,本研究基于 Illumina NovaSeq 6000 测序平台及 TMT 标记分析的方法首次对极端株高性状的饲用燕麦进行转录组和蛋白质组数据关联分析。通过对两组学的测序结果进行关联分析,从 22762 个差异表达转录本和 3934 个差异表达蛋白中筛选出了 1147 个重叠的差异表达基因,其中 290 个转录本及其对应的蛋白表达量共同上调;339 个转录本及其对应的蛋白表达量共同下调;324 个转录本表达量下调,对应的蛋白表达量上调;194 个转录本表达量上调,对应的蛋白表达量下调。KEGG 注释分析表明,大部分相关因子分别注释到了谷胱甘肽代谢、苯丙素的生物合成、萜类生物合成、氨基酸糖和核苷酸糖代谢、木质素、亚木质素和蜡的生物合成以及次级代谢物生物合成和其他代谢途径通路中。究其原因,可能与饲用燕麦生长发育过程中的细胞增殖与代谢等有关。

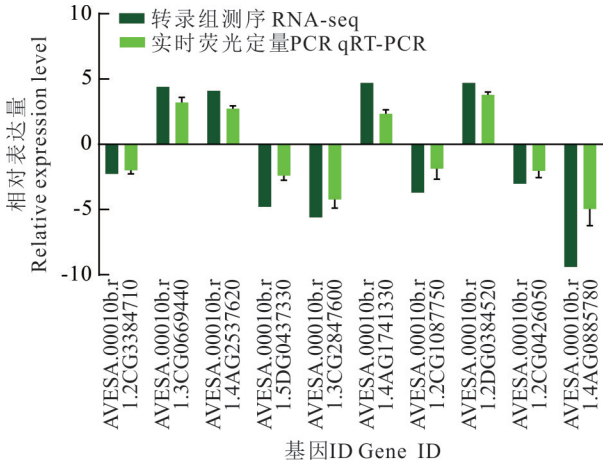


图 9 转录组数据的 qRT-PCR 验证

Fig. 9 qRT-PCR validation of transcriptome data

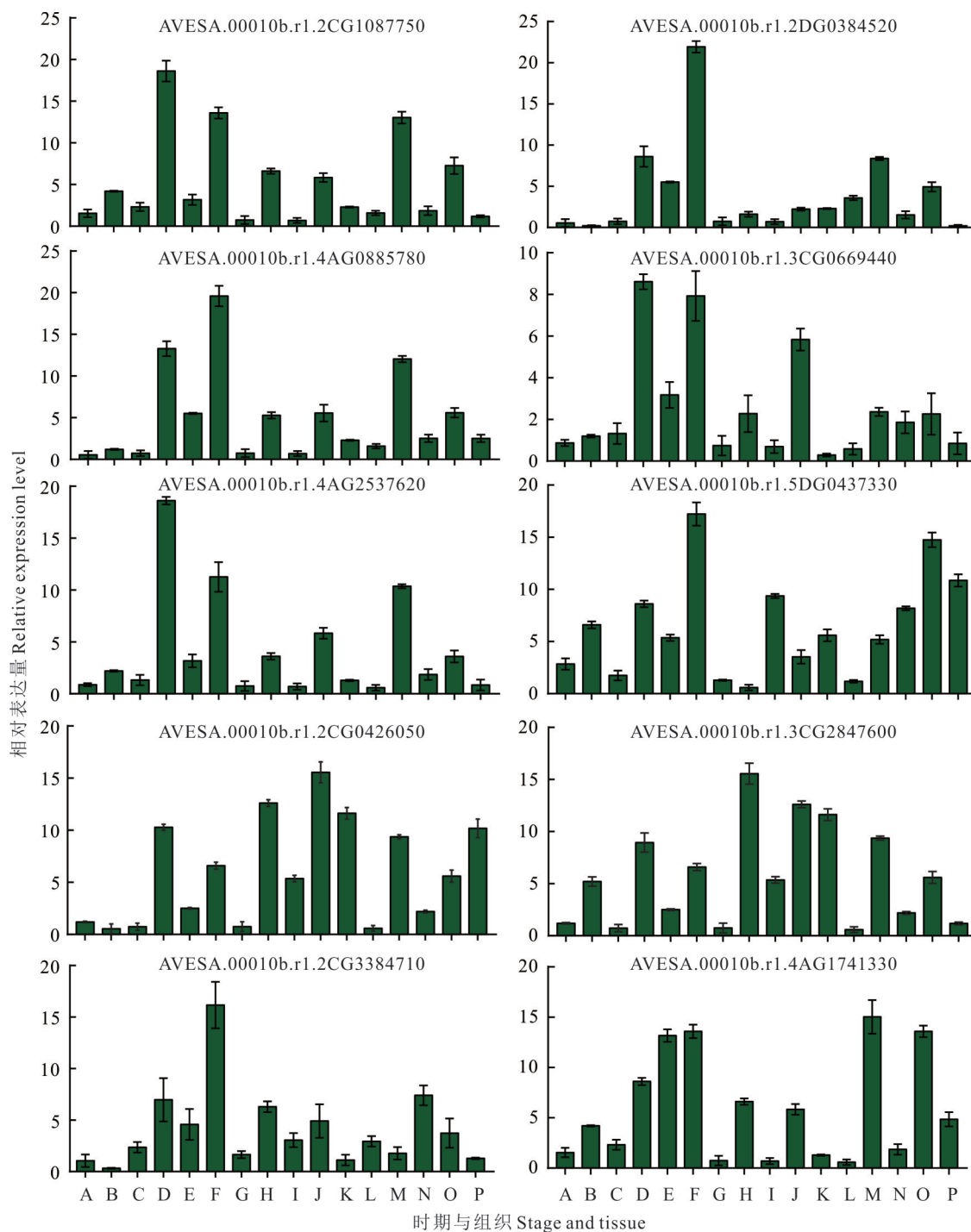


图 10 候选基因的组织特异性表达

Fig. 10 Tissue-specific expression of candidate genes

A: 苗期根 Seedling stage root; B: 苗期地上 Seedling stage aboveground; C: 拔节期根 Shooting stage root; D: 拔节期茎 Shooting stage stem; E: 拔节期叶 Shooting stage leaf; F: 拔节期茎节 Shooting stage stem nodes; G: 孕穗期根 Booting stage root; H: 孕穗期茎 Booting stage stem; I: 孕穗期叶 Booting stage leaf; J: 孕穗期茎节 Booting stage stem nodes; K: 孕穗期穗 Booting stage spike; L: 抽穗期根 Heading stage root; M: 抽穗期茎 Heading stage stem; N: 抽穗期叶 Heading stage leaf; O: 抽穗期茎节 Heading stage stem nodes; P: 抽穗期穗 Heading stage spike.

3.2 植物株高性状的调控机制

多项研究表明,植物株高基因通过调控茎的细胞长度或细胞数目来影响株高,如玉米 *D11*^[17]、水稻 *sdg*^[18] 和小麦 (*Triticum aestivum*) *Rht16*^[19] 都是通过缩短茎节细胞大小控制株高;而水稻 *sde*^[20]、玉米 *d2003*^[21] 和小麦 *Rht22*^[22] 是通过减少节间细胞数目控制株高;小麦 *Rht26*^[23] 同时影响茎节细胞数目和大小,进而调控小麦的株高。本研究

通过对饲用燕麦株高极端材料的穗下茎节与节间进行转录组和蛋白组联合分析,筛选出 1147 个差异表达基因(蛋白)。进一步对差异表达基因(蛋白)进行 GO 和 KEGG 富集分析发现,多数基因被显著富集到与细胞生长、代谢和细胞壁形成的通路上,这与前人的研究结果类似:Chen 等^[24]通过对小麦 *Rht12* 矮秆株系穗下茎组织进行转录组测序发现差异表达基因显著富集在细胞形态建成、钾离子转运和类胡萝卜素等与细胞伸长有关的通路上。事实上很多已报道的水稻矮秆基因,如 *sd1*、*d18*、*d61*、*OsBZR1*、*EUI1*、*GID2*、*D14*、*SRS5*、*OsCYP96B4*、*sd-t* 和 *GID1* 等均通过参与以上通路调控植株的高度。

本研究对饲用燕麦进行转录组与蛋白组测序的联合分析,通过对重叠基因进行 GO 富集分析和 KEGG 通路分析,在筛选到的 10 个饲用燕麦株高相关候选基因中,有 8 个基因的功能有明确注释:AVESA.00010b.r1.3CG0669440(*GID* 基因家族)、AVESA.00010b.r1.4AG2537620 和 AVESA.00010b.r1.2DG0384520(*GH* 基因家族)、AVESA.00010b.r1.2CG0426050(*GAD* 蛋白家族)、AVESA.00010b.r1.3CG2847600(色氨酸氨基转移酶)、AVESA.00010b.r1.2CG3384710(*bHLH* 转录因子)、AVESA.00010b.r1.4AG0885780(*ERF* 转录因子)、AVESA.00010b.r1.4AG1741330(过氧化物酶基因)。多项研究表明,以上基因家族成员在株高形成的过程中起关键作用:矮秆基因 *Rht* 通过与 *GID1*(*GID* 家族成员)的结合来影响植物的生长发育,显性矮秆基因 *Rht3* 对小麦的降秆强度为 36%,隐性矮秆基因 *Rht22* 通过减少细胞数目来降低小麦的株高^[25];王慧敏^[26]发现与野生型相比,超表达 *SlGH3.2*(*GH* 家族成员)的水稻株系明显矮化;在番茄(*Lycopersicon esculentum*)中过表达 *SlGH3.15*(*GH* 家族成员)表现出由节间距缩短所致的植物矮化、叶片卷曲及侧根变少^[27];Kazuhito 等^[28]通过对水稻成熟期的 *OsGAD2*(*GAD* 家族成员)敲除植株进行分析,发现突变体与野生型相比,株高呈极显著增加趋势;*bHLH* 转录因子成员 *AtbHLH146* 通过抑制细胞伸长相关基因的表达调控植物细胞伸长,过表达 *AtbHLH146* 的转基因拟南芥(*Arabidopsis thaliana*)植株表现出植株矮小的表型^[29];AP2/ERF 家族成员 *OsEATB* 通过下调赤霉素合成基因(贝壳杉烯合成酶 A)的表达,降低了成熟期水稻的株高^[30];另一个 AP2/ERF 家族成员 *OsRPH1*,通过调控赤霉素(gibberellin, GA)生物合成和 GA 失活相关基因的表达,减少水稻中的活性 GA 含量,进而调控水稻株高^[31]。因此,已筛选到的基因可能参与饲用燕麦株高的发育。本研究利用 qRT-PCR 技术分析候选基因在不同时期、不同组织的表达特性,发现这些基因均在燕麦不同生长时期的茎与茎节组织表达丰度较高,这些直接和间接的证据进一步证明本研究所筛选的候选基因很可能参与调控饲用燕麦株高性状的发育,但其具体生物学功能和调控方式还需作进一步的验证和研究。

4 结论

通过转录组与蛋白组数据联合分析发现,细胞生长、代谢和细胞壁形成的过程均在调节饲用燕麦株高能力中具有十分重要的作用,并且筛选到 10 个与饲用燕麦株高发育关系密切的基因,可以作为候选基因,进一步探究饲用燕麦株高性状的分子调控机制。

参考文献 References:

- [1] Liu W H, Jia Z F, Liang G L. The current situation, problems and suggestions for the development of China's oat feed industry. Qinghai Science and Technology, 2020, 27(3): 82—85.
刘文辉,贾志锋,梁国玲.我国饲用燕麦产业发展现状及存在的问题和建议.青海科技,2020,27(3):82—85.
- [2] Ye X L, Gan Z, Wan Y, et al. Advances and perspectives in forage oat breeding. Acta Prataculturae Sinica, 2023, 32(2): 160—177.
叶雪玲,甘圳,万燕,等.饲用燕麦育种研究进展与展望.草业学报,2023,32(2):160—177.
- [3] Chen X, Wu B, Zhang Z W. Evaluation of adaptability and stability for important agronomic traits of oat (*Avena* spp.) germplasm resources. Journal of Plant Genetic Resources, 2016, 17(4): 577—585.
陈新,吴斌,张宗文.燕麦种质资源重要农艺性状适应性和稳定性评价.植物遗传资源学报,2016,17(4):577—585.
- [4] Milach S, Rines H W, Phillips R L, et al. Inheritance of a new dwarfing gene in oat. Crop Science, 1998, 38(2): 356—360.
- [5] Brown P D, Mckenzie R, Mikaelsen K. Agronomic, genetic, and cytologic evaluation of a vigorous new semidwarf oat. Crop

- Science, 1980, 20(3): 303—306.
- [6] Xu Y H. Fine mapping of oat dwarfing gene *Dw6*. Chengdu: Sichuan Agricultural University, 2023.
徐颖红. 燕麦矮秆基因 *Dw6* 的精细定位. 成都: 四川农业大学, 2023.
- [7] Milach S C K, Rines H W, Phillips R L. Molecular genetic mapping of dwarfing genes in oat. *Theoretical & Applied Genetics*, 1997, 95(5/6): 783—790.
- [8] Morikawa T. Genetic analysis on dwarfness of wild oat, *Avena fatua*. *The Japanese Journal of Genetics*, 1989, 64: 363—371.
- [9] Zhang X B. Map-based cloning and functional analysis of leaf shape gene *CFL2* and plant height gene *SD38* in rice. Chongqing: Southwest University, 2021.
张孝波. 水稻叶形基因 *CFL2* 和株高基因 *SD38* 的图位克隆与功能分析. 重庆: 西南大学, 2021.
- [10] Yamamuro C, Ihara Y, Wu X, *et al.* Loss of function of a rice *brassinosteroid insensitive1* homolog prevents internode elongation and bending of the lamina joint. *The Plant Cell*, 2000, 12(9): 1591—1605.
- [11] Zhang J, Liu X, Li S, *et al.* The rice semi-dwarf mutant *sd37*, caused by a mutation in CYP96B4, plays an important role in the fine-tuning of plant growth. *PLoS One*, 2014, 9(2): e88068.
- [12] Schefe J H, Lehmann K E, Buschmann I R, *et al.* Quantitative real-time RT-PCR data analysis: Current concepts and the novel “gene expression’s CT difference” formula. *Journal of Molecular Medicine*, 2006, 84(11): 901—910.
- [13] Yang F, Ye R, Ma C, *et al.* Toxicity evaluation, toxin screening and its intervention of the jellyfish *Phacellophora camtschatica* based on a combined transcriptome-proteome analysis. *Ecotoxicology and Environmental Safety*, 2022, 46(6): 1—12.
- [14] Luo J. *Brassica napus* L. dwarf traits proteome and transcriptome joint analysis. Guiyang: Guizhou Normal University, 2021.
罗京. 甘蓝型油菜矮化性状的蛋白质组和转录组联合分析. 贵阳: 贵州师范大学, 2021.
- [15] Li H, He X W, Gao Y F, *et al.* Integrative analysis of transcriptome, proteome, and phosphoproteome reveals potential roles of photosynthesis antenna proteins in response to brassinosteroids signaling in maize. *Plants*, 2023, 12(6): 1290—1307.
- [16] Lin J, Zheng X, Xia J, *et al.* Integrative analysis of the transcriptome and proteome reveals the molecular responses of tobacco to boron deficiency. *BMC Plant Biology*, 2024, 24(1): 689—707.
- [17] Wang Y J, Lu W J, Zhao J, *et al.* Transcriptome dynamics of dominant maize dwarf *Dwarf11* (*D11*) revealed by RNA-seq and co-expression analysis. *Plant Molecular Biology Reporter*, 2017, 35(3): 355—365.
- [18] Sui J M. Fine mapping of one semidwarf gene *sdt3* and candidate-gene screening and functional analysis of the other semidwarf gene *sdg* in rice (*Oryza sativa* L.). Yangzhou: Yangzhou University, 2006.
隋炯明. 水稻半矮秆基因 *sdt3* 的精细定位和 *sdg* 的克隆与功能分析. 扬州: 扬州大学, 2006.
- [19] Gong Y S, Wei S H, Peng Z S, *et al.* Genetic study on plant height and its components, partial yield traits in durum wheat ‘ANW16F’. *Southwest China Journal of Agricultural Sciences*, 2021, 34(2): 229—235.
龚胤书, 魏淑红, 彭正松, 等. 硬粒小麦 ANW16F 株高及构成因子与部分产量性状遗传研究. *西南农业学报*, 2021, 34(2): 229—235.
- [20] Wu M Y. Genetic analysis and gene mapping of three (semi-) dwarf genes in rice. Beijing: Chinese Academy of Agricultural Sciences, 2020.
吴明月. 三个(半)矮秆水稻基因的遗传分析和基因定位. 北京: 中国农业科学院, 2020.
- [21] Lv H K, Zhang J, Wang T Y, *et al.* The maize *d2003*, a novel allele of *VP8*, is required for maize internode elongation. *Plant Molecular Biology*, 2014, 84(3): 243—257.
- [22] Peng Z, Li X, Yang Z, *et al.* A new reduced height gene found in the tetraploid semi-dwarf wheat landrace Aiganfanmai. *Genetics and Molecular Research*, 2011, 10(4): 2349—2357.
- [23] Song J, Li L, Liu B Y, *et al.* Fine mapping of reduced height locus *RHT26* in common wheat. *Theoretical and Applied Genetics*, 2023, 136(3): 62—72.
- [24] Chen L, Yang Y, Mishina K, *et al.* RNA-seq analysis of the peduncle development of *Rht12* dwarf plants and primary mapping of *Rht12* in common wheat. *Cereal Research Communications*, 2020, 48(2): 1—9.
- [25] Shan Q Q. Analysis of allelic variations of wheat plant height regulation genes *Rht-1* and *GID1* and their interaction mechanism. Zhengzhou: Henan Agricultural University, 2018.
单强强. 小麦株高调控基因 *Rht-1* 和 *GID1* 的等位变异分析及其编码蛋白的互作机制. 郑州: 河南农业大学, 2018.
- [26] Wang H M. Characterization of the expression patterns of tomato *SIGH3.2* and its potential functions in rice plants. Nanjing: Nanjing Agricultural University, 2015.
王慧敏. 番茄 *SIGH3.2* 的表达特征及在水稻中的功能研究分析. 南京: 南京农业大学, 2015.

- [27] Ai G. Functional dissection of *SlGH3-15* in tomato. Wuhan: Huazhong Agricultural University, 2017.
艾国. 番茄 *SlGH3-15* 基因的功能解析. 武汉: 华中农业大学, 2017.
- [28] Kazuhito A, Fumio T. C-terminal extension of rice glutamate decarboxylase (*OsGAD2*) functions as an autoinhibitory domain and overexpression of a truncated mutant results in the accumulation of extremely high levels of GABA in plant cells. *Journal of Experimental Botany*, 2007, 58(10): 2699—2707.
- [29] Jia Y T. Function analysis of transcription factor *bHLH146* in *Arabidopsis thaliana*. Changchun: Jilin University, 2022.
贾雨彤. 拟南芥转录因子 *bHLH146* 的功能研究. 长春: 吉林大学, 2022.
- [30] Qi W W, Sun F, Wang Q J, *et al.* Rice ethylene-response AP2/ERF factor *OsEATB* restricts internode elongation by down-regulating a gibberellin biosynthetic gene. *Plant Physiology*, 2011, 157(1): 216—228.
- [31] Ma Z M, Wu T, Huang K, *et al.* A novel AP2/ERF transcription factor, *OsRPH1*, negatively regulates plant height in rice. *Frontiers in Plant Science*, 2020, 11(13): 709—724.