

干旱胁迫下青稞全基因组可变剪切分析

曾兴权^{1,2}, 原红军^{1,2*}, 杨春葆^{1,2}, 于明寨^{1,2}, 徐齐君^{1,2},
王玉林^{1,2}, 扎桑^{1,2}, 尼玛扎西^{1,3}

(1. 省部共建青稞和牦牛种质资源与遗传改良国家重点实验室, 西藏 拉萨 850002; 2. 西藏自治区农牧科学院农业研究所, 西藏 拉萨 850002; 3. 西藏自治区农牧科学院, 西藏 拉萨 850002)

摘要:为阐明青稞在早胁迫下基因发生可变剪切的规律, 本研究以抗旱的“喜拉 16 号”和早敏感的“迪庆黑元桂”为试验材料, 分别对对照和 21 % PEG 浓度下处理的样本进行多时间点的全转录组测序。结果表明, 利用 LeafCutter 软件在所有的样品中共检测到了 22 181 个可变剪切事件; 通过 PCA 分析发现, 可变剪切具有明显的品种间特异性; 另外, 通过与抗旱基因数据库比对分析发现, 在干旱胁迫处理下 2 个类型品种间有多个抗旱相关的基因发生了显著的差异可变剪切, 分别是 *HVUL1H16429 (AtrbohF)*、*HVUL1H12808 (HAB1)*、*HVUL4H58649 (ABF4)*、*HVUL4H35985 (AtrbohD)*、*HVUL7H07799 (LOS5)*、*HVUL3H43774 (CLCc)*、*HVUL3H23631 (ABCG40)*、*HVUL1H16840 (MYB60)* 和 *HVUL6H08236 (AREB1)*; 对 2 个品种各自在对照与早胁迫条件下的差异可变剪切基因进行 pathway 分析, 发现了抗旱品种的差异可变剪切基因参与了更多的 pathway, 并且鉴别出 Fatty acid degradation、Inositol phosphate metabolism、alpha-Linolenic acid metabolism 和 Fatty acid metabolism 这 4 条通路显著富集。为解析青稞抗旱分子机理与培育青稞抗旱新品种奠定了理论基础。

关键词:青稞; 干旱胁迫; 全基因组; 可变剪切分析

中图分类号:S512.3 **文献标识码:**A

Genome-wide Variable Shear Analysis of Highland Barley under Drought Stress

ZENG Xing-quan^{1,2}, YUAN Hong-jun^{1,2*}, YANG Chun-bao^{1,2}, YU Ming-zhai^{1,2}, XU Qi-jun^{1,2},

WANG Yu-lin^{1,2}, Zhasang^{1,2}, Nimazhaxi^{1,3}

(1. State Key Laboratory of Hulless Barley and Yak Germplasm Resources and Genetic Improvement, Tibet Lhasa 850002, China; 2. Institute of Agricultural Research, Tibet Academy of Agricultural and Animal Husbandry Sciences, Tibet Lhasa 850002, China; 3. Tibet Academy of Agricultural and Animal Husbandry Sciences, Tibet Lhasa 850002, China)

Abstract: In order to study the pattern of variable shearing of genes in highland barley under drought stress, the drought-resistant ‘Xila 16’ and drought-sensitive ‘Diqing Heiyuanguai’ were used as experimental materials. The samples processed at control and 21 % PEG concentrations were separately sequenced at multiple time points. The results showed that a total of 22 181 variable shearing events were detected in all samples by Leaf Cutter software, and it was also found by PCA analysis that the variable shearing has obvious specificity among varieties. In addition, by comparing with drought resistance gene database, there were significant differences in drought resistance related genes between the two varieties under drought stress: *HVUL1H16429 (AtrbohF)*, *HVUL1H12808 (HAB1)*, *HVUL4H58649 (ABF4)*, *HVUL4H35985 (AtrbohD)*, *HVUL7H07799 (LOS5)*, *HVUL3H43774 (CLCc)*, *HVUL3H23631 (ABCG40)*, *HVUL1H16840 (MYB60)* and *HVUL6H08236 (AREB1)*. Comparing the variable shearing genes between two varieties by KEGG analysis, the drought-resistant varieties were involved in more pathways, and four significant enrichment pathways were identified, which were fatty acid degradation, inositol phosphate metabolism, alpha-Linolenic acid metabolism and fatty acid metabolism pathways. The research lays a theoretical foundation for analyzing the molecular mechanism of drought resistance of highland barley and breeding new varieties.

Key words: Highland barley; Drought stress; Whole genome; Variable shear analysis

青稞是我国藏区第一大粮食作物, 因其具有对

高原极端气候环境的适应性, 在距今约 3500 年前, 青稞就已作为我国藏区人民重要的粮食作物而得到广泛种植; 青稞也是青藏高原农区饲草主要来源, 青稞占西藏粮食总产的 80 % 以上, 占西藏农区饲草的 60 % 以上。

干旱是一种主要的非生物胁迫, 限制了农业生

收稿日期: 2019-02-05

基金项目: 西藏财政专项 (XZNKY-2018-C-021)

作者简介: 曾兴权 (1975-), 男, 博士, 研究员, 主要从事青稞遗传育种研究, E-mail: xingquanz_2@126.com, * 为通讯作者; 原红军 (1979-), 男, 硕士, 副研究员, 主要从事青稞遗传育种研究, E-mail: yuazhenbin@126.com。

产,严重威胁着全世界的粮食供应。青藏高原海拔落差大,气候多样,生态条件复杂,有季节性的干旱。与其它谷类作物相比,具有较高耐旱性的大麦是干旱相关研究的最合适的靶标,也是最有希望在现代大麦品种育种中应用的抗旱基因和基因区域的来源。然而,耐旱性是由许多不同基因控制的数量性状,是一种复杂的质量性状,由多个基因和基因家族共同控制,不易同时选择,需要在分子层次去剖析它的机理,为育种工作奠定基础。研究发现,在缓慢的干旱胁迫时,青稞抗旱品种的差异表达基因参与了内质网蛋白加工、TCA 循环、蜡质生物合成和剪切体;在急性干旱胁迫时,光合作用的碳固定、丙酮酸代谢、卟啉和叶绿素代谢都会发生改变^[1]。

为研究植物在逆境下的代谢规律和揭示调控机制,将传统的更具有描述性的生理学方法与功能基因组学技术相结合,即近几年来进行的高通量的转录测序、蛋白质组和代谢组研究方法被广泛研究是非常重要的。在植物中,可变剪切在植物的抗逆过程中起到十分重要的作用。目前有很多分析可变剪切的方法,可以识别不同的剪切模式,计算亚型的表达量。但是,由于亚型之间有大量的重复区域,每条二代的测序片段仅仅是一条转录本的一部分的原因,直接用 RNA-seq 的测序片段去计算亚型的表达量在准确性上面临很多挑战。Yang I. Li 等^[2]提出了一个新的更加灵敏和准确的基于外显子剪切的计算工具。

近年来,随着高通量测序技术的发展和运用,已有不少研究应用 RNA-seq 的手段研究青稞的性状^[3-4],但是,大部分研究仅仅局限于表达差异的分析和研究,但是表达的差异并不能完全对性状进行解释,需要从更多的角度去解析。已有少量的研究对青稞的 RNA 的可变剪切进行了研究,如 Yuan H 等^[5]通过转录组测序的方法对青稞的耐冻性进行了转录组调控分析,发现冷胁迫会改变 sering/arginine-rich (SR) 蛋白的可变剪切。但是对青稞这一物种的此类研究还不多,特别是还没有针对青稞的抗旱性状系统地转录组上寻找可变剪切的规律。

1 材料与方法

1.1 青稞不同品种和不同时期取样

供试材料为前期筛选的强抗旱性“喜拉 16 号”和干旱敏感材料“迪庆黑元桂”^[6]。青稞抗旱材料“喜拉 16 号”与青稞旱敏感材料“迪庆黑元桂”取样时间点设置为干旱处理后 0、1、4、24 和 48 h。

1.2 RNA 提取

实验步骤参考 TIANGEN 说明书 (TIANGEN

BIOTECH (BEIJING) CO., LTD)。

1.3 文库构建和上机测序

每个样品取 3 μ g 总 RNA 作为起始量构建 lncRNA 文库。使用 Ribo-ZeroTM GoldKits 去除样品中的 rRNA,根据 NEB Next Ultra Directional RNA LibraryPrep Kit for Illumina (NEB, Ipswich, USA) 的操作说明分别选取不同的 index 标签建库。

文库构建的具体步骤为:首先使用试剂盒去除核糖体 rRNA,向反应体系中加入 Fragmentation Buffer 使 RNA 片段化成为短片段,再以片段化后的 RNA 为模板,用六碱基随机引物 (Random Hexamers) 合成 cDNA 第一链,并加入缓冲液、dNTPs、RNase H 和 DNA Polymerase I 合成 cDNA 第二链,经过 QiaQuick PCR 试剂盒纯化并加 EB 缓冲液洗脱对得到的 cDNA 第二链进行末端修复、加碱基 A,加测序接头,经琼脂糖凝胶电泳回收目的大小片段,加 UNG 酶消化 cDNA 二链,并进行 PCR 扩增,最后琼脂糖凝胶电泳回收目的大小片段,从而完成整个文库制备工作。最后用构建好的文库进行 Illumina 测序。

1.4 数据分析

原始数据通过质控后,通过 tophat^[7] 比对到青稞基因组。基于比对结果的 bam 文件,再通过 LeafCutter 系统鉴别样品中存在的可变剪切。KEGG 的富集分析先手计算每个途径的基因个数,然后应用超几何检验,找出与整个基因组背景相比,差异表达基因显著富集的途径。

2 结果与分析

2.1 可变剪切鉴别

mRNA 成熟过程中发生的可变剪切对真核生物的很多生物过程十分重要。现有的很多计算可变剪切的工具在计算亚型 (isoform) 是仍面临很多挑战,如剪切亚型与正常转录本之间的重合度很高,在估算丰度的时候会有一定的误差。Li Y I 等^[2]开发了 LeafCutter,基于 RNA-seq 短的 reads 鉴定可变的剪切事件,能够检测高度复杂的剪切事件。LeafCutter 不需要预先对转录本进行注释,并且克服了在复杂剪切事件中估算相对异构体使用的难题。

本研究使用 LeafCutter 对试验材料的 RNA-seq 数据通过 LeafCutter 方法计算可变剪切的 percentage spliced-in 值 (PSI),共检测到 22 181 个可变剪切事件。

2.2 广泛存在的品种与处理特有的剪切事件

为系统探究在不同的品种和处理间剪切的发生模式,对“喜拉 16 号”(XL)与“迪庆黑元桂”(DQ)

表 1 差异的可变剪切的数量

名称 Name	数量 Number
DQ-CK-S	39
DQ-XL-CK	480
DQ-XL-S	483
XL-CK-S	32

在处理(S)和对照(CK)间进行可变剪切差异分析。以 $Q\text{ value} < 0.05$ 为阈值,差异的可变剪切的数量如表所示。

由表 1 可知,品种间的差异剪切事件数量远远大于品种内不同处理间的差异剪切事件数量。

对这些差异剪切进行 PCA 分析,由图 1 可知,PCA 分析可以较为清晰地将不同的品种或者处理区分开来,说明在青稞中,品种、处理特异性的剪切事件是普遍存在的。从图 1 可见,DQ 与 XL 不管是在对照条件下,还是在处理条件下,PCA 结果都显示出比品种内的处理间更加明显的区分。进一步说

明,品种间特异性的剪切事件在青稞中是比较重要的现象。

由图 2 可知,在对照和处理下,DQ 与 XL 间有 238 个差异剪切事件是重合的,这些是品种间固有的差异,而 2 者所特有的 218 和 214 个剪切事件则分别是抗旱品种和旱敏感品种间在正常环境与旱胁迫下所特有的剪切事件。

2.3 重要的抗旱相关基因的差异剪切事件

为探究重要的抗旱基因在本次数据中的剪切模式,本研究使用 DroughtDB^[8] 数据库中的信息进行研究。首先,下载 DroughtDB 中抗旱基因的蛋白序列,使用 blastp 将青稞的基因序列翻译成蛋白序列与之比对。设定的阈值为 $e\text{ value} < 1E-5$,序列的覆盖比例为大于 50 %。

旱敏感品种 DQ 的处理与对照间未鉴别出抗旱相关基因的可变剪切事件,抗旱品种 XL 的处理与对照间鉴别出一个抗旱基因 *HVULAH35985* (*AtrbohD*, NADPH oxidase catalytical subunit)发生了差异

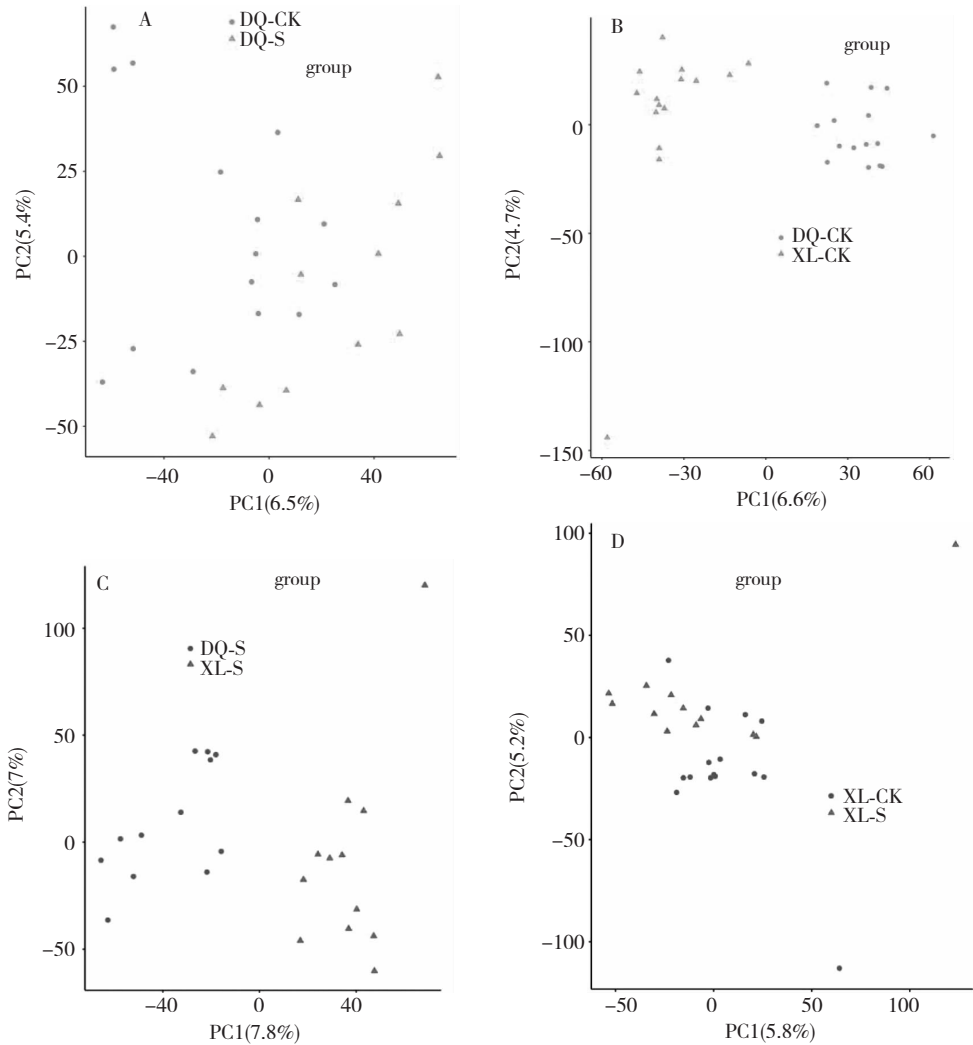


图 1 差异剪切 PCA 分析

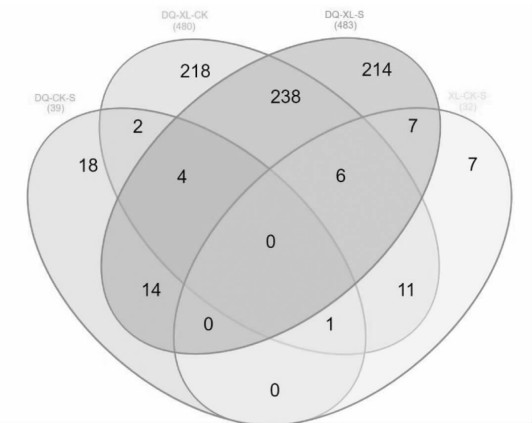


图 2 差异剪切韦恩图

可变剪切。据 Kwak et al. 等研究,ATABOHD/F 双突变体可导致 ABA 诱导气孔关闭受损。

值得关注的是,在干旱处理下,DQ 与 XL 间有多个发生差异剪切基因为抗旱相关基因,这些基因分别是 *HVUL1H16429* (*AtrbohF*)、*HVUL1H12808* (*HAB1*)、*HVUL4H58649* (*ABF4*)、*HVUL4H35985* (*AtrbohD*)、*HVUL7H07799* (*LOS5*)、*HVUL3H43774* (*CLCc*)、*HVUL3H23631* (*ABCG40*)、*HVUL1H16840* (*MYB60*)和 *HVUL6H08236* (*AREB1*)。

2.4 抗旱品种响应旱胁迫的可变剪切基因参与的 pathway

应用超几何检验对差异可变剪切的基因参与的 KEGG pathway 进行富集分析,BH(Benjamini-Hochberg)法用于 *P* value 调整,得到 *Q* value 值,*Q* value <0.05 的 pathway 为显著富集的 pathway。

图 5A 为青稞早敏感材料“迪庆黑元桂”的差异可变剪切的基因参与的 pathway 的显著值,其中 Nitrogen metabolism 显著富集。由图 5B 可见,抗旱材料“喜拉 16 号”的差异可变剪切基因参与的 pathway

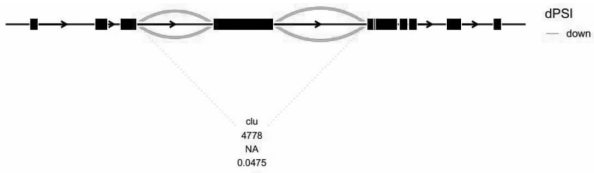


图 3 HVUL4H35985 的基因结构

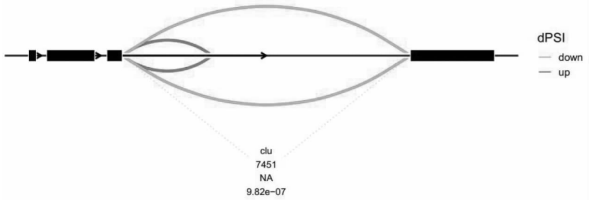


图 4 HVUL1H12808 结构

的数量较多,显著的 pathway 有 Fatty acid degradation、Inositol phosphate metabolism、alpha-Linolenic acid metabolism 和 Fatty acid metabolism。

3 讨论

青稞作为青藏高原藏区主要粮食作物,从基因调控层次上研究清楚抗旱机制,对于品种的改良具有重要的意义。可变剪切使得生物体内基因在转录后产生不同的转录本,进而翻译产生结构和功能不同的蛋白,是形成基因组和蛋白组多样性的一种重要的机制^[10-11]。目前的研究表明,可变剪切在动植物体内广泛存在,对生物体的调控具有重要的作用。

本研究利用青稞抗旱材料“喜拉 16 号”与早敏感材料“迪庆黑元桂”,通过 RNA 建库测序方法研究了青稞在受到干旱胁迫后可变剪切的变化,从分析结果来看,不管是在对照还是在干旱处理下,品种间的差异可变剪切都比品种内的多,表明可变剪切很可能是这两个品种对逆境敏感度不一样的原因之一。从分析结果来看,抗旱品种在旱处理与对照下

表 2 旱胁迫下 2 个品种间差异可变剪切中的抗旱相关基因信息

基因 Gene	表型 Phenotype	参考文献 Reference
<i>ABCG40</i>	knockout; upregulation of ABA responsive genes is delayed, stomata close more slowly	Kang et al. , 2010
<i>ABF4</i>	overexpression; enhanced drought tolerance	Kang et al. , 2002
<i>AREB1</i>	needs ABA-dependent phosphorylation for activation;phosphomimicking variant overexpressed; expression of ABA-inducible genes w/o ABA treatment	Furihata et al. , 2006
<i>AtrbohD</i>	atrbohD/F double mutant; ABA induced stomatal closing is impaired	Kwak et al. , 2003
<i>AtrbohF</i>	atrbohD/F double mutant; ABA induced stomatal closing is impaired	Kwak et al. , 2003
<i>CLCc</i>	expressed guard cells, root an pollen;tonoplast;mutant; ABA dependent stomata closure impaired	Jossier et al. , 2010
<i>HAB1</i>	mutant; ABA hypersensitive; hab1G246D mutant is unable to close their stomata to limit water loss in response to ABA synthesized during a drought stress	Robert et al. , 2006
<i>LOS5</i>	knockout; drought susceptible, ABA deficient, increased water loss	Xiong et al. , 2001
<i>MYB60</i>	expression downregulated under dehydration;guard cell-specifically expressed;knockout ; water loss under drought is limited;enhanced drought tolerance	Cominelli et al. , 2005

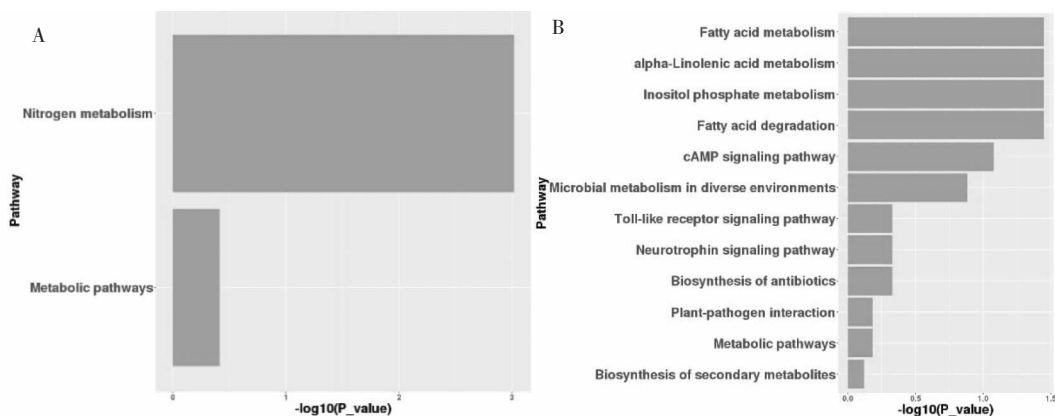


图 5 差异剪切基因参与的 pathway 及其显著性

的差异剪切基因参与的 pathway 中, Fatty acid metabolism 在前人研究中表明可能参与了植物的干旱适应过程^[9]。其他 pathway 中的基因也是候选的通过改变基因的剪切模式来调控植物的抗旱性的基因。

研究发现 2 个品种在对照与处理下的差异可变剪切基因富集的通路有较大差异, 在喜拉中, HVUL3H17775 与 HVUL1H10694 分别参与了富集的 4 个通路中的 3 个通路。HVUL3H17775 序列比对到 GenBank 的编号为 BAJ86175. 1, 是一个预测的蛋白, 其主要的功能是 3-羟基酰基-CoA 脱氢酶。HVUL1H10694 为酰基辅酶 A 氧化酶 4, 过氧化物酶体, 其主要的作用是酰基辅酶 A 氧化酶。HVUL0H14419 (肌环己六氢醇 2-脱氢酶)、HVUL2H19484 (肌醇-五磷酸盐 2-激酶) 也参与了富集的通路。在以表达量为标准的差异分析模式下, 未发生差异表达的基因之间存在的很多的剪切差异是被忽略的, 而这些剪切可引起后续转录本及蛋白的不一致从而对植物的表型产生影响。从差异可变剪切的角度剖析不同品种在不同环境下的响应, 为后续探究翻译这些酶蛋白的基因由于剪切而造成功能的变化奠定了基础。

4 结 论

本研究在所有的样品中共检测到了 22 181 个可变剪切事件; 发现了在旱胁迫处理下 2 类型品种间有均多个抗旱相关的基因发生了显著的差异可变剪切, 且具有一定的物种差异性; 研究也发现了抗旱品种的差异可变剪切基因参与了更多的 pathway, 并且鉴别出 Fatty acid degradation、Inositol phosphate metabolism、alpha-Linolenic acid metabolism 和 Fatty acid metabolism 这 4 条通路显著富集, 研究为青稞抗旱新品种培育奠定了理论基础。

参考文献:

- [1] LIANG J, CHEN X, DENG G, et al. Dehydration induced transcriptomic responses in two Tibetan hulless barley (*Hordeum vulgare* var. *nudum*) accessions distinguished by drought tolerance [J]. BMC genomics, 2017, 18(1): 775.
- [2] LI Y, KNOWLES D, HUMPHREY J, et al. Annotation-free quantification of RNA splicing using LeafCutter [J]. Nature Genetics, 2018, 50(1): 151.
- [3] TANG Y, ZENG X, WANG Y, et al. Transcriptomics analysis of hulless barley during grain development with a focus on starch biosynthesis [J]. Functional & Integrative Genomics, 2017, 17(1): 107-117.
- [4] WEI Z, ZENG X, QIN C, et al. Comparative transcriptome analysis revealed genes commonly responsive to varied nitrate stress in leaves of Tibetan hulless barley [J]. Frontiers in Plant Science, 2016, 7: 1067.
- [5] YUAN H, ZENG X, LING Z, et al. Transcriptome profiles reveal cold acclimation and freezing tolerance of susceptible and tolerant hulless barley genotypes [J]. Acta Physiologiae Plantarum, 2017, 39(12): 275.
- [6] 王玉林, 徐齐君, 原红军, 等. PEG 模拟干旱胁迫处理对青稞幼苗生长和生理特性的影响 [J]. 大麦与谷类科学, 2018, 35(1): 6-12.
- [7] TRAPNELL C, PACTHER L, SALZBERG S L. TopHat: discovering splice junctions with RNA-Seq [J]. Bioinformatics, 2009, 25(9): 1105-1111.
- [8] ALTER S, BADER K, SPANNAGL M, et al. DroughtDB: an expert-curated compilation of plant drought stress genes and their homologs in nine species [J]. Database, 2015: 046.
- [9] ZHONG D, DU H, WANG Z, et al. Genotypic variation in fatty acid composition and unsaturation levels in bermudagrass associated with leaf dehydration tolerance [J]. Journal of the American Society for Horticultural Science, 2011, 136(1): 35-40.
- [10] EXCOFFON K, BOWERS J, SHARMA P. Alternative splicing of viral receptors: A review of the diverse morphologies and physiologies of adenoviral receptors [J]. Recent Research Developments in Virology, 2014, 9: 1-24.
- [11] HAMID F, MAKEYEV E. Emerging functions of alternative splicing coupled with nonsense-mediated decay [J]. Biochemical Society Transactions, 2014, 42(4): 1168-1173.