Identification and analysis of flowering time candidate genes of *Dactylis glomerata* L.

Zhang, XQ[*1]; Feng, GY[1]; Huang, LK[1]; Zhao, XX[2].; Yang ZF[1]

*1 Department of Grassland Science, Animal Science and Technology College, Sichuan Agricultural University, Chengdu, 611130, China

2 Institute of Urban Agriculture, Chinese Academy of Agricultural Sciences, Chengdu, 610000, China

Key words: *Dactylis glomerata* L; flowering time; QTL; BSA

Abstract

Orchardgrass (*Dactylis glomerata* L.) is an important forage grass for cultivating livestock worldwide. Heading date is a surrogate measure of flowering time and is strongly correlated with the yield and quality of forage grasses. Here, we identified 210 flowering-related genes in orchardgrass by BLASTP. In addition, our integrated analysis of transcriptome, quantitative trait locus (QTL), and bulked segregant analysis (BSA) provided insights into the genetic network regulating flowering time. The results revealed that four main candidate genes may participate in flowering time control, including one *FT-like* gene and three MADS-box genes. Expression investigation between the early-flowering and late-flowering phenotype of orchardgrass indicated gene DG6G02970.1 was the only significantly differentially expressed gene, which encodes the MADS-box gene *AGL61*-like. Three nonsynonymous SNPs were identified in the *AGL61*-like gene, resulting in changes from alanine to valine, alanine to threonine and glycine to valine, most likely to cause differentially expression pattern in early-flowering and late-flowering phenotype of orchardgrass. Weighted gene co-expression network analysis (WGCNA) revealed that *AGL61*-like is associated with a set of 114 genes relevant to carbohydrate metabolic, oxidative stress response and glycolysis/gluconeogenesis process. These results provided useful clues for flowering time regulation in molecular breeding of forage grasses.

Introduction

Flowering is a crucial phase in the development of perennial grasses and most flowering plants, signalling the changeover from vegetative growth to reproductive growth (Simpson and Caroline, 2002). The applicable timing of anthesis is necessary for the plant's adaptation to environmental stresses, as well as the completion of the plant's life cycle (Curtis, 1968). In agriculture, normal flowering is a prerequisite for crop production whether the seeds or vegetative mass are harvested, and it is regulated by multiple genetic factors that interact with environmental conditions (Craufurd and Wheeler, 2009; Feng et al., 2017). Existing research on molecular biology has identified several critical pathways that play an important role in flowering regulation, such as the GA pathway, vernalization pathway, photoperiod pathway, and autonomous pathway (Mouradov et al., 2002). For most annual crops, the flowering time within a growing season is predominantly determined by the responses of temperature and photoperiod. Perennial crops, however, have an added vernalization requirement, and uncertain external environmental factors may lead to greater impact on the life cycle of perennial grasses. Hence an appropriate flowering time is strongly related to nutrient quality and forage yield (Bushman et al., 2012). For perennial forages, however, far too little attention has been paid to characterizing the flowering mechanisms.

Orchardgrass (*Dactylis glomerata* L.) is an important cool-season perennial forage grass that is cultivated and utilized worldwide. Indigenous to Eurasia and northern Africa, it has been naturalized on nearly every continent and utilized as a pasture or hay grass, which is important for the production of forage-based meat and dairy throughout the temperate regions of the world. As an excellent pasture and ecological improvement grass, orchardgrass has high biomass, abundant carbohydrates, shade tolerance, and broad adaptability (Volaire, 2010). Heading date is a surrogate measure for flowering time and is strongly correlated with the yield and quality of forage grasses. Interestingly, due to the widespread geographical distribution in orchardgrass, its substantial variation in flowering timing provided a valuable source of germplasm to tease apart the genetic mechanisms underlying flowering time regulation (Sheldrick et al., 2010). Therefore, in this study, we performed an integrated analysis to identify flowering genes based on transcriptome, QTL and BSA data, as well as detected potential regulators co-expressed with candidates. This research will contribute to our perception of the flowering adjustment of orchardgrass, further promoting selection in forage grass breeding.

Methods and Study Site

An F_1 population of 213 individuals derived from the cross between two orchardgrass cultivars-'Kaimo' (early flowering time) and '01436' (late flowering time)-was used for map construction and QTL. Plants were grown in the research field of the Grassland Science Department, Sichuan Agricultural University at Ya'an, Sichuan, China. Genomic DNA samples of the two parents and 213 F_1 progeny individuals were subjected to SLAF-seq with some modifications as described (Sun et al., 2013). Briefly, HighMap software was used for map construction and QTL analysis based on our previous report (Zhao et al., 2016). Significant loci associated with heading date and flowering time were identified based on LOD scores larger than the 5% cut-off value determined through 1,000 permutation tests.

To identify SNPs of genes involved in flowering time, 29 full-sib individuals from an F_1 mapping population of 213 lines were used for QTL sequencing. SNPs that were homozygous in one parent and heterozygous in the other parent were prioritized and extracted from the 'vcf' output files. The homozygous genotype of the parent was used as the reference to calculate the number of reads of this parent's genotype in the individuals in the offspring pools. The ratio of reads harbouring the SNP that was different from the reference sequence was calculated as the SNP index of the base site. Sliding-window methods were used to present SNP indexes across the whole genome. The SNP index for each window was calculated as the average of all SNP indexes in the selected window of the genome. The window size was set as 1 Mb, and the step size was set as 1 Kb. The difference in the SNP index of the two pools, namely, one earlier flowering pool and one later flowering pool, was calculated as the transformed Δ (SNP index).

Using the *Arabidopsis. thaliana* homologues as queries, the putative flowering orthologous candidate genes in orchardgrass were identified by BLASTP with an E-value cut-off of 1e-5 (Bouché et al., 2016). If these genes were in common families in OrthoMCL, then their protein domains were predicted by Pfam (http://pfam.xfam.org/). Only genes that had the same protein domain as X were considered orthologous to the *A. thaliana* genes. To identify the differentially expressed genes and elaborated the expression pattern of candidate flowering genes, the expression data was extracted based on our previous research (Feng et al., 2017; Feng et al., 2018). The weighted gene co-expression network analysis was performed by using WGCNA package in R (v3.3.0) (Langfelder and Horvath, 2008).

Results

Identification of flowering time genes in orchardgrass

In this study, 603 orthologues and paralogues in the orchardgrass genome were identified, corresponding to 210 flowering-related genes in the *Arabidopsis thaliana* flowering-time gene dataset (Bouché et al., 2016). Of these, 85 orchardgrass orthologues and paralogues corresponding to 53 flowering-related genes were differentially expressed between early- and late-flowering phenotypes of orchardgrass. Meanwhile, 25 and 5 were detected in the vernalization and photoperiod pathways, respectively. Several key flowering regulators such as the photoperiod related gene *CO1*, vernalization related genes *VRN1* and *VRN2*, circadian clock related gene *LUX1* and flowering integrator *FT* paralogue were differentially expressed between early- and late-flowering phenotypes, potentially contributing to the difference in heading date. Additionally, five *FT* orthologues might have undergone expansion during orchardgrass evolution, suggesting their essential roles in flowering time.

Flowering time gene mapping and expression analysis

To identify candidate genetic regions and key regulators associated with flowering time, we integrated QTL analysis and BSA with transcriptome expression-profiling data. The peak value for the transformed Δ (SNP index) localized to two regions spanning from 154.344 Mb to 156.231 Mb and from 157.05 to 159.599 Mb on chromosome 6. Based on the QTL results, we also identified a major locus at 157.639 Mb (np6325) on chromosome 6 that overlapped with the BSA candidate regions (Figure 1a). Fine-mapping analysis identified a 4.426-Mb overlapped region on chromosome 6 that may harbour the major locus contributing to orchardgrass heading date. After removing genes that were not expressed among the pre-vernalization, vernalization, post-vernalization, pre-heading, and heading stages, 30 candidate genes were predicted within this region (Figure 1b). Polymorphism detection identified 6 nonsynonymous SNPs corresponding to 4 candidates, including one *FT*-like gene and three MADS-box genes, in the early- and late-flowering populations (Figure 1c). In previous reports, the MADS-box family was revealed to be a highly conserved gene family involved in flowering time, floral organ formation and inflorescence architecture (Schilling et al., 2018). In the orchardgrass reference sequence, we identified 94 MADS-box genes, including 58 type I and 36 type II genes (Gramzow and Theissen, 2010). The MADS-box gene family was markedly expanded in the orchardgrass genome compared with other grass genomes, which likely drive the extensive variation in

heading date and strong adaptability to environmental conditions of orchardgrass.

To investigate the gene expression of these four candidates, comparative transcriptome analysis was performed between the early-flowering and late-flowering orchardgrass lines. Gene DG6G02970.1 was the only significantly differentially expressed gene; this gene encodes the MADS-box gene *AGL61*-like, which plays an essential role in pollen tube guidance and the initiation of endosperm development (Steffen et al., 2008). Mutants of the A. thaliana homologue AT2G24840.1 (*AGAMOUS-LIKE 61*, *AGL61*) express female fertility reduction and defective central cells with abnormal morphology. *AGL61*-like showed higher expression among five critical flowering stages in the early-flowering line than in the late-flowering line. Three nonsynonymous SNPs were identified in the *AGL61*-like gene, resulting in changes from alanine to valine, alanine to threonine and glycine to valine (Figure 1c). Thus, DG6G02970.1 might participate in flowering regulation of orchardgrass.

WGCNA was used to search for candidate genes that were associated with flowering regulators. WGCNA revealed that DG6G02970.1 (*AGL61*-like) is associated with a set of 114 genes in the early-flowering line. GO term enrichment indicated that carbohydrate metabolic process genes were particularly enriched, and glycolysis/gluconeogenesis pathway genes were enriched in the KEGG analysis. Among the biological processes, four terms related to carbohydrate metabolic process and two terms related to response to oxidative stress were highly enriched. The need for a high level of carbohydrates for enhanced flowering has been demonstrated. Carbohydrate accumulation is related to the transition from vegetative growth to flowering (Kozłowska et al., 2007). Assuming a conserved function of *AGL61*-like in flowering regulation, we annotated genes that were differentially expressed in pre-vernalization stage versus post-vernalization stage or pre-heading stage versus heading stage comparisons in the early-flowering line. This analysis identified a potential relationship between *AGL61*-like and the carbohydrate metabolic process. However, transgenic evidence needs to be provided to further confirm that the difference in heading date is caused by *AGL61*-like alone or the cooperation of *AGL61*-like and other co-expressed genes.

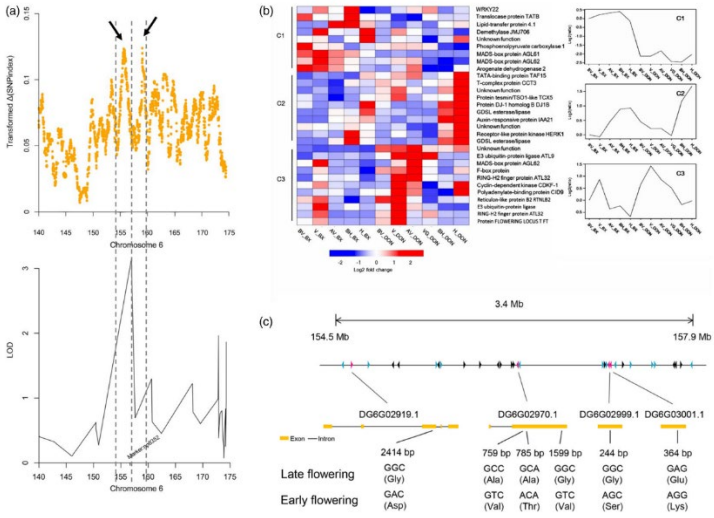

Figure 1. Genetic mapping of the orchardgrass flowering time candidate genes. (a) Mapping the flowering-time genes by BSA and QTL analysis. The transformed Δ (SNP index) is the product of the Δ (SNP index) and normalized SNP density in each 1-Mb sliding window (10-kb steps). The dark arrow and dashed line indicate the positions of the 1.89-Mb and 2.55-Mb peaks, respectively. (b) The clusters and expression patterns of 30 candidate genes. The heatmap on the left side shows the expression of 30 candidate genes, the line chart on the right side show the expression pattern of clusters. (c) Exon-intron structure and nonsynonymous SNPs of four candidates in two phenotypes.

Discussion

Flowering regulation is a convoluted process for higher plants. A network consisting of more than 300 genes in Arabidopsis as only the tip of an iceberg in entire network of flowering-regulation (Bouché et al., 2016). In this research, 210 flowering-related genes were identified in orchardgrass, and further performed the comparison of expression patterns of these flowering genes. In addition, we found 50 potential flowering-related genes were differentially expressed between early- and late-flowering phenotype of orchardgrass, and most of these DEGs were enriched in photoperiod pathway and vernalization pathway. These results indicated there may be diversity in flowering regulatory networks among different species, but

the key regulators that have significant impacted on flowering time are relatively conservative. Due to the complex genome organization, one or a limited number of differentially expressed flowering-genes may not have a decisive influence on flowering time of orchardgrass. Whereas the cumulative effect of many differentially expressed verified flowering genes or novel genes may be an important reason for the flowering time diversity of orchardgrass. Furthermore, four main candidates were identified as key regulator via the combination of transcriptome, QTL and BSA data. Of which, three nonsynonymous SNPs may result in distinct expression patterns of *AGL61*-like in early- and late-flowering phenotype of orchardgrass. The functional studies of *AGL61* suggest a role in female gametophyte development, embryo sac development and early seed development (Bemer et al., 2008).Our results indicated that *AGL61*-like may play a similar function in development regulation of orchardgrass. In future studies, more flowering time regulating genes will be identified and validated by using known flowering regulators and hub genes involved in the gene network constructed by WGCNA.

Acknowledgements

This research work was funded by the National Basic Research Program (973 Program) in China (No. 2014CB138705), the National Natural Science Foundation of China (NSFC 31872997), the Earmarked Fund for Modern Agro-industry Technology Research System (No. CARS-34) and the National Project on SciTec Foundation Resources Survey (2017FY100602).

References

Bemer, M., Wolters-Arts, M.U., and Angenent, G.C. 2008. The MADS Domain Protein DIANA Acts Together with AGAMOUS-LIKE80 to Specify the Central Cell in Arabidopsis Ovules. *Plant Cell*., 20:2088-2101.

Bouché, F., Lobet, G., Tocquin, P., and Périlleux, C. 2016. FLOR-ID: an interactive database of flowering-time gene networks in *Arabidopsis thaliana*. *Nucleic Acids Research*., 44:D1167-D1171.

Bushman, B.S., Robins, J.G., and Jensen, K.B. 2012. Dry Matter Yield, Heading Date, and Plant Mortality of Orchardgrass Subspecies in a Semiarid Environment. *Crop Science*., 52:745-751.

Craufurd, P.Q., and Wheeler, T.R. 2009. Climate change and the flowering time of annual crops. *Journal of Experimental Botany*., 60:2529-2539.

Curtis, D.L. 1968. The Relation Between the Date of Heading of Nigerian Sorghums and the Duration of the Growing Season. *Journal of Applied Ecology*., 5:215.

Feng, G., Huang, L., Li, J., Wang, J., Xu, L., Pan, L., Zhao, X., Wang, X., Huang, T., and Zhang, X. 2017. Comprehensive transcriptome analysis reveals distinct regulatory programs during vernalization and floral bud development of orchardgrass (*Dactylis glomerata* L.). *BMC Plant Biology*., 17:216.

Feng, G., Xu, L., Wang, J., Nie, G., Bushman, B.S., Xie, W., Yan, H., Yang, Z., Guan, H., Huang, L., et al. (2018). Integration of small RNAs and transcriptome sequencing uncovers a complex regulatory network during vernalization and heading stages of orchardgrass (*Dactylis glomerata* L.). *BMC Genomics*., doi: 10.1186/s12864-018-5104-0.

Gramzow, L., and Theissen, G. 2010. A hitchhiker's guide to the MADS world of plants. *Genome Biology*., 11:214.

Kozłowska, M., Rybus-Zając, M., Stachowiak, J., and Janowska, B. 2007. Changes in carbohydrate contents of Zantedeschia leaves under gibberellin-stimulated flowering. *Acta Physiologiae Plantarum*., 29:27-32.

Langfelder, P., and Horvath, S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*., 9(1):559.

Mouradov, A., Cermer, F., and Coupland, G. 2002. Control of Flowering Time: Interacting Pathways as a Basis for Diversity. *Plant Cell*., 14:S111-130.

Schilling, S., Pan, S., Kennedy, A., and Melzer, R. 2018. *MADS*-box genes and crop domestication: the jack of all traits. *Journal of Experimental Botany*., 69(7).

Sheldrick, R.D., Lavender, R.H., and Tewson, V.J. 2010. The effects of frequency of defoliation, date of first cut and heading date of a perennial ryegrass companion on the yield, quality and persistence of diploid and tetraploid broad red clover. *Grass & Forage Science*., 41(2):137-149.

Simpson, G.G., and Caroline, D. 2002. *Arabidopsis*, the Rosetta stone of flowering time? *Science*., 296(5566):285-289.

Steffen, J., Kang, I., Mf, Lloyd, A., and Drews, G. 2008. *AGL61* interacts with *AGL80* and is required for central cell development in *Arabidopsis*. *Plant Physiology*., 148(1):259-268.

Sun, X., Liu, D., Zhang, X., Li, W., Liu, H., Hong, W., Jiang, C., Guan, N., Ma, C., and Zeng, H. 2013. SLAF-seq: An Efficient Method of Large-Scale De Novo SNP Discovery and Genotyping Using High-Throughput Sequencing. *Plos One*., 8:e58700.

Volaire, F. 2010. Seedling survival under drought differs between an annual (*Hordeum vulgare*) and a perennial grass (*Dactylis glomerata*). *New Phytologist*., 160:501-510.

Zhao, X., Huang, L., Zhang, X., Wang, J., Yan, D., Ji, L., Lu, T., Li, X., and Shi, T. 2016. Construction of high-density genetic linkage map and identification of flowering-time QTLs in orchardgrass using SSRs and SLAF-seq. *Scientific Reports*., 6:29345.