Characterization of photoperiodic genes *Ghd8* and *Ghd7* on flowering time regulation in a mini-core collection of *Miscanthus sinensis*

（ミニコアコレクションを用いたススキにおける開花期制御に関する日長遺伝子 *Ghd8* と *Ghd7* の特徴づけ）

GUO Zhihui

Graduate School of Environmental Science

Hokkaido University

2021

Abstract

The genus *Miscanthus* is a rhizomatous, self-incompatible, C_4 perennial grass with a wide natural distribution, and is closely related to sugarcane (*Saccharum officinarum*) and sorghum (*Sorghum bicolor*). Owing to its environmental adaptability and high yields with low nutrient requirements, *Miscanthus* is regarded as a potential bioenergy crop. Optimization of flowering time is essential to obtain high biomass yield under different environments, and may also impact biomass quality for *Miscanthus*. Controlling flowering will facilitate the hybridization of genotypes from diverse geographical locations and assist the intergeneric crosses, such as between *Miscanthus* and *Saccharum*. Synchronizing flowering time will also be essential for the development of a seed-propagated crop. Flowering regulation in *Miscanthus sinensis,* one of the important species in *Miscanthus*, was so complicated, operated by degree days but also a photoperiod sensitivity mechanism. Nowadays, *M. sinensis* is identified as a facultative short-day (SD) plant and days to flower is strongly affected by photoperiod, but the genetic mechanism on controlling flowering in *M. sinensis* is poorly understood. The photoperiod regulation of flowering is well known in rice (*Oryza sativa*), and many significant flowering regulatory genes have been evolutionarily conserved in the Gramineae family. Two essential flowering genes in rice were selected for identification in *M. sinensis*. Therefore, the aim of the present study is 1) to identify allelic and deduced amino acid sequence diversity and geographic distribution of two flowering-related genes in a mini-core collection of *M. sinensis*, representing a wide range of flowering responses to photoperiod, genetic groups (population structure) and latitudes of origin, and 2) to analyze gene expression pattern by

quantitative real-time PCR (qRT-PCR) to characterize their response to photoperiod.

GRAIN YIELD, *PLANT HEIGHT AND HEADING DATE 8* (*Ghd8*), a major quantitative trait locus in rice, was isolated in *M. sinensis*. The deduced amino acid sequence of *Ghd8* in *M. sinensis* contained a HEME ACTIVATOR PROTEIN 3/NUCLEAR FACTOR-YB (HAP3/NF-YB) DNA-binding domain, which is critical for the transcription factor function of *Ghd8* gene products. Two homoeologous loci were identified, *MsiGhd8A* located on chromosome 13 and *MsiGhd8B* on chromosome 7, with one on each of this paleo-allotetraploid species' subgenomes of *M. sinensis*. A total of 46 alleles and 28 predicted protein sequence types were detected in 12 wild-collected accessions. Several variants of *MsiGhd8* showed a geographic and latitudinal distribution. Gene expression analysis revealed that *MsiGhd8* expressed under both long-day (LD) and SD conditions, and *MsiGhd8B* showed a significantly higher expression than *MsiGhd8A*.

GRAIN YIELD, *PLANT HEIGHT AND HEADING DATE 7* (*Ghd7*) was generated interest because of its genetic interaction with *Ghd8* and a monocot-specific flowering gene. *Ghd7* is evolutionarily conserved in *M. sinensis,* and the CONSTANS, CONSTANS-like AND TIMING OF CAB1 (CCT)- domain protein was preserved in *MsiGhd7*. One homoeologous locus, *MsiGhd7A* located at chromosome 11 in the A subgenome. And multiple *MsiGhd7B* loci with a repetitive region in the intron were found at chromosome 12 in the B subgenome. One putative loss-of-function allele, identified in *MsiGhd8B*, was characterized by an eight-base insertion in the first exon, resulting in a frameshift and eventual

premature termination of the protein, and entirely lack CCT domain. Both *MsiGhd7* homoeologous genes expressed higher in LD relative to SD, and the mRNA transcript level of *MsiGhd7* was abundant in the early morning under LD.

The relative expression of *MsiGhd8* was at day time and night time, while *HEADING DATE 1* (*MsiHd1*) peaked at night, indicating that MsiGhd8-Hd1 complex might form and accumulate at night, subsequently activate the transcription of *MsiGhd7* in the morning under LD condition. And this MsiGHD8-HD1 complex potentially induce expression of *FLOWERING LOCUS T* (*FT*)- *like* genes [*CENTRORADIALIS 8* (*CN8*), *CN12* and *CN15*] under SD condition. *EARLY HEADING DATE 1* (*Ehd1*) showed a relative higher expression in SD. *MsiGhd7* might work as one of the upstream genes of *Ehd1* and suppress its expression in LD. Moreover, mRNA transcriptional level of *CN8*, *CN12* and *CN15* in *M. sinensis* were greatly promoted under SD condition. Thus, *Ehd1* might be one of the upstream genes of these three florigens. The comparison between days to flower and gene expression for each accession indicated that *CN8*, *CN12* and *CN15* affected flowering time in response to day length for most *M. sinensis* accessions. Whereas, for *M. sinensis* from high latitude, the SD might also be a signal to induce a dormancy response, which is epistatic to flowering. Taken together, these gene expression patterns for multiple flowering candidate genes characterize possible pathways that modulates photoperiodic flowering-time in *M. sinensis* for most accessions under LD and SD conditions.

The present study is the first of this kind of report that screened the diversity and geographic distribution of allele and protein variants, but also investigated the

gene expression in response to photoperiod in *M. sinensis*. Identifying these two major genes provides a novel perspective on flowering in *M. sinensis* and will accelerate the process to elucidate the flowering regulatory network of *Miscanthus*. Furthermore, it may provide information for the breeder to improve *Miscanthus* varieties as a bioenergy crop.

Acknowledgments

Firstly, I sincerely thank my supervisor Prof. Dr. Toshihiko Yamada for his guidance and thoughtful perspectives throughout the duration of my doctoral program at the Hokkaido University, Japan. I also thank my committee members: Profs. Dr. Ryouichi Tanaka, Dr. Akira Kanazawa and Dr. Yoichiro Hoshino for their expert advice and helpful feedback. I am very grateful to Dr. Erik J. Sacks and Dr. Lindsay V. Clark of University of Illinois at Urbana-Champaign, USA for providing me professional advice and valuable suggestions. I also thank the staff at Energy Crop Farm in University of Illinois at Urbana-Champaign, USA for their collaboration in preparing *Miscanthus* plants. I am deeply thankful to Dr. Meilan Xu for her time, patience, and tremendous help. Thanks are in order to Ms. Fiorella D. B. Nuñez, Mr. Hironori Nagano, Dr. Suraj Kar, Dr. Rafael Alexandre Muchanga, Dr. Yufita Dwi Chinta, Ms. Tzu-Ya Weng for advice on the experiment. I also express my deep gratitude towards all staff and technicians of our lab for their invaluable help during field and laboratory work.

Finally, I am thankful to my family for their unconditional supports and love throughout my life. The most heartfelt thanks go to my beloved husband, Mr. Yu Hao, for his encouragement and love. I also would express my appreciation to my friends for their company and friendship throughout the years.

And last but not least, I would like to thank China Scholarship Council for the gracious financial support to pursue my doctoral study. I also heartily thank Hokkaido University for waiving all of my tuition fees and giving me this opportunity to study in Japan.

Table of contents

Chapter 1

General inroduction

1.1 Bioenergy crops and *Miscanthus*

Climate change, which is mainly caused by carbon emissions, has become a global concern (Frank *et al.*, 2010). Achieving a low-carbon future will be challenging and will require a comprehensive technology and policy measurement. Fossil fuels have been identified as one of the main source of carbon emissions, however, with the rapid economic development all over the world, the consumption of fossil fuels is increasing, leading to the negative environmental effects, such as global warming and air pollution, which are driving a search for renewable energy sources that could potentially mitigate climate change and enhance energy security (Chu and Majumdar, 2012; Robertson *et al.*, 2017). Bioenergy becomes a particularly important part of the energy economy to supplementary to fossil fuels, accounting for over 70% of all renewable energy production with low-cost and low-maintenance and making a contribution to 9.3% of world total energy supply by source (International Energy Agency, 2017, 2020). According to the composition of energy carrier materials, bioenergy crops can be divided into three main categories: (a) starch and sugar crops that can be used in the production of fuel ethanol; (b) oil crops that can be catalyzed into bio-diesel; (c) lignocellulosic crops, which are rich in cellulose, hemicelluloses, and lignin, can be converted to generate heat, electricity, biogas and ethanol (Wang *et al.*, 2021). To accelerate the growth of bioenergy, a mix of manure and commercial technologies (such as biomass

gasification, pyrolysis and the production of ethanol from cellulosic feedstocks), appropriate policies of different country and regions, necessary investments and increased supplement of biomass feedstock for modern bioenergy uses are indispensable (International Energy Agency, 2017). Recently, many studies are conducting to improve candidate bioenergy crops and explore the dedicated biomass crops with policies set nationally to enhance production (Abraha *et al.*, 2020; Mandley *et al.*, 2020; Wu *et al.*, 2020). As a candidate biomass crop, the main objectives are to focus on (1) maximizing the productivity in a sustainable and cost-effective way, and (2) improving the conversion efficiency of biomass into biofuels such as ethanol. Comparing to C_3 plants, the group of C_4 grasses have photorespiration-suppressing modifications resulting in a relatively higher potential efficiency of converting solar energy to biomass (Ehleringer and Cerling, 2002; Zhu *et al.*, 2010), therefore, many of candidate biomass crops are C_4 grasses, such as maize (*Zea mays*), *Miscanthus* (*Miscanthus* spp.), sorghum (*Sorghum bicolor*), sugarcane (*Saccharum* spp.) and switchgrass (*Panicum virgatum*). Especially, perennial biomass energy crops (e.g. *Miscanthus* and switchgrass) are high resource-use efficient and low nutrient demanding, and can store nutrients over winter in underground roots at regrowth (Scordia *et al.*, 2014; Siri-Prieto *et al.*, 2020). Furthermore, perennial crops could also contribute to improving soil quality through the sequestration of carbon into the soil, preventing land degradation, and could provide an economically attractive means to begin to restore lands that would otherwise be extremely expensive to restore (Tilman *et al.*, 2006; Ye and Hall, 2020). Therefore, the development of dedicated bioenergy crops, especially breeding of new cultivars, play a role in the future supply chain of bioenergy and

thus benefit for the establishment of a viable cellulosic ethanol industry and finally favor to the global climates, which stimulates researches into the genomics, genetics, and breeding of the promising C_4 grasses.

Miscanthus is a perennial rhizomatous genus and belongs to the Gramineae family, tribe *Andropogoneae*, subtribe *Saccharinae* (Anzoua *et al.*, 2011). In the broad sense, *Miscanthus* spp. includes about 20 species (Clifford *et al.*, 1986), while in strictly, it contains about 12 species with a basic chromosome number of x = 19, named as *Miscanthus sensu stricto* (*s.s.*) (Hodkinson *et al.*, 2015). *Miscanthus s.s.* is classified into three sections: *Miscanthus* section, which contains *Miscanthus sinensis*, *Miscanthus floridulus* and *Miscanthus transmorrisonensis*, and *Triarrhetha* section, which includes *Miscanthus sacchariflorus* and *Miscanthus lutarioriparius*, and *Kariyasua* section, which includes *Miscanthus oligostachyus*, *Miscanthus tinctorius* and *Miscanthus intermedius* (Hodkinson *et al.*, 2015). Among them, *M. sinensis* and *M. sacchariflorus* are two dominating species and could from a species complex through hybridization (Adati and Shiotani, 1962; Hodkinson *et al.*, 2002), therefore, they are considered to be good source of high yielding plants suitable for biomass accumulation. To date, a natural interspecific hybrid between diploid *M. sinensis* (2n = 2x = 38) and allotetraploid *M. sacchariflorus* (2n = 4x = 76) is *Miscanthus* × *giganteus* (*M* × *g*), the only commercially dominant triploid (Greef and Deuter, 1993; Chramiec-Głabik *et al.*, 2012), thus *M* × *g* belongs to both *Miscanthus* and *Triarrhetha* sections. Additionally, *M. floridulus* generally grows at warm tropical climates and is clustered into a *M. sinensis* clade (Clark *et al.*, 2014). Taiwanese native *M. floridulus* achieved high biomass (Huang *et al.*, 2011) and showed great potential

for bioethanol production (Yeh *et al.*, 2016), it would be another suitable bioenergy crop, especially in southern region.

Miscanthus s.s. originate from Eastern Asia, Southeastern Asia and the South Pacific (Vermerris, 2008). In nature, its latitudinal range extends from Northeastern Siberia at 50 °N to tropical Polynesia at 22 °S, and the native longitudinal distribution extends from Burma, and Andaman and Nicobar Islands at 92 °E to Fiji at 179 °W (Vermerris, 2008). Additionally, Clark *et al*. (2014) reported that South-eastern China was the origin of *M. sinensis* populations found in temperate eastern Asia, which is consistent with this area probably having been a refugium during the last glacial maximum (LGM) by restriction site-associated DNA sequencing (RAD-seq). After the LGM, *M. sinensis* migrated directly from south-eastern China to Japan before migrating to the same latitudes in China and South Korea, in consistent with natural ongoing changes in the climate global. Owing to human introduction, *Miscanthus* has also been imported into many regions of the world, including Eurasia, North and South America, and New Zealand, as an ornamental crop (Meyer *et al.*, 2010; Quinn *et al.*, 2010, 2012; Matlaga *et al.*, 2012; Clark *et al.*, 2014, 2015).

As a perennial grass, the lifetime of *Miscanthus* can extend to 15 years (Clifton-Brown *et al.*, 2019) or potentially up to 25 years. Once established and fertilized in the first two seasons, *Miscanthus* is a permanent crop requiring no field management except harvest and can maintain high production of lignocellulosic biomass (Dubis *et al.*, 2019). Moreover, even under temperate climate or low temperature, the photosynthetic efficiency of *Miscanthus* is still high and some

genotypes could compete with C_3 plants (Beale and Long, 1995; Tubeileh *et al.*, 2016; Jiao *et al.*, 2017; Pignon *et al.*, 2019). Meanwhile, the water (Zhuang *et al.*, 2013; Hamilton *et al.*, 2015; Kørup *et al.*, 2018) and nutrient use efficiency (Iqbal *et al.*, 2015; Oliveira *et al.*, 2017) of *Miscanthus* are at a high level. Iqbal *et al.* (2015) reported that *Miscanthus* were sensitive to fertilizer by comparing to switchgrass, and the nitrogen content in the harvested biomass increased along with increasing in fertilizer levels. In summary, *Miscanthus* exhibits prominent characteristics to drought tolerance (van der Weijde *et al.*, 2017*a*; Scordia *et al.*, 2020), frost and cold tolerance (Clifton-Brown and Jones, 1997; Głowacka *et al.*, 2014; Dong *et al.*, 2019*a*,*b*), salt and alkali stresses, and broad resistance to a variety of diseases and insects (Jørgensen, 2011). Therefore, these advantages allow it to adapt divergent ecological environments, and *Miscanthus* may be applied as an emerging bioenergy crop.

1.2 Flowering of *Miscanthus*

Floral initiation is a major physiological change that sets the switch from vegetative to reproductive development in most plant species. The transition from a vegetative (production of stem and leaves) to a reproductive stage (production of inflorescences and flowers) determines the time of flowering (or heading date in cereals) and is one of the most important developmental switches in the life cycle of plants. From a human perspective, flowering time has major consequences for crop production, both in terms of vegetative biomass in the case of vegetables, and inflorescence biomass primarily in the case of grains, fruits, and seeds. Flowering time was regulated by the integration of environmental inputs with endogenous

cues (Shrestha *et al.*, 2014). According to their requirement for daylength, plants can be classified into three categories. Long-day (LD) plants flower when the photoperiod exceeds a critical daylength, short-day (SD) plants flower when the photoperiod is shorter than a critical daylength and day-neutral (DN) plants flower at the same time regardless of the photoperiodic conditions (Jackson, 2009). Many flowering plants are potentially photoperiodic; typical LD plants include *Arabidopsis thaliana*, *Medicago truncatula*, pea (*Pisum sativum*), barley (*Hordeum vulgare*) and wheat (*Triticum aestivum*). SD plants include but are not limited to rice, soybean (*Glycine max*), maize (*Zea mays*) and sorghum. Commonly, tomato (*Solanum lycopersicum*) and some tobacco (*Nicotiana tabacum*) species are DN plants. The critical daylength for floral induction is specific to each species but often varies between accessions of the same species. Additionally, plants that respond to daylength can be further subdivided into obligate (or qualitative) types, where a particular daylength is essential for flowering, or facultative (or quantitative) types, where a particular daylength accelerates but is not essential for flowering. The previous studies pointed out that SD could induce dormancy response and reduce or prevent flowering in switchgrass and big bluestem (*Andropogon gerardii*) (especially for high-latitude populations), which are quantitative SD, perennial, C_4 grasses (Benedict, 1941; McMillan, 1959; Castro *et al.*, 2011). Besides, accumulated temperatures also showed a significant impact on flowering times of *M. sacchariflorus*, which is considered as a quantitative SD plant (Jensen *et al.*, 2013).

Miscanthus has a long life span while typically flowers annually (Clifton-Brown *et al.*, 2019). Within *Miscanthus*, *M. sacchariflorus* is a quantitative SD

plant that would not flower when daylength exceeded a certain critical threshold (estimated at ~12.5h) (Jensen *et al.*, 2013). Flowering regulation in *M. sinensis* was more convoluted. Initially, flowering induction of *M. sinensis* was regarded as a DN plant (Deuter, 2000) and flowering induction was much affected by accumulated heat units degree days (Jorgensen and Muhs, 2001), whereas Jensen *et al*. (2011) demonstrated that flowering of *M. sinensis* was manipulated by multiple factors, including degree days, temperature, growing season rain fall and photoperiod. In the latest study, Dong *et al*. (2021) reported that *M. sinensis* was a facultative SD plant, and photoperiod strongly affected *Miscanthus* flowering.

As with most grasses, flowering terminates the production of leaves at the stem apex, potentially reducing the length of the growing season, thereby lessening the time over which radiation is intercepted and hence decreasing potential biomass accumulation. In the previous work, Clifton-Brown *et al*. (2001) described earlier flowering *Miscanthus* genotypes produced lower yields than those that flower late, or that never flower. As a perennial grass, flowering time of *Miscanthus* were also linked to the impact on nutrient remobilization to the underground rhizome, thereby promoting crop sustainability and improving combustion quality by minimizing moisture, ash, potassium (K), chloride (Cl), nitrogen (N), and sulfur (S) from the harvested biomass (Lewandowski *et al.*, 2003). The commercial cultivar, *M*×*g*, a sterile triploid, is propagated depending on the clone (rhizome) (Lewandowski, 1998). Compared to the seed-based crops, the cropping of $M \times g$ is more expensive, which limits the development of the *Miscanthus* industry that requires novel hybrid production by the intra- and interspecific hybridization between *Miscanthus* and *Saccharum* (Xue *et al.*, 2015). Therefore, the optimization or the uniformity of

flowering time in *Miscanthus* could be essential for biomass yield and quality as well as *Miscanthus* commercialization. These information also suggest that manipulation of genes that regulate the process of flowering induction in *Miscanthus* may be candidates for the genetic engineering. However, the information on signaling pathways and key genes associated with *Miscanthus* flowering are poorly understood.

1.3 Flowering genetic regulatory pathways in Arabidopsis and Gramineae family

Molecular mechanisms for flowering regulation have been extensively studied in the model plant *A. thaliana* (Fornara *et al.*, 2010; Kim, 2020). A complex interplay of environmental cues and endogenous signals determines flowering at the appropriate time (Amasino and Michaels, 2010; Preston and Fjellheim, 2020). To date, at least six flowering pathways have been characterized in *A. thaliana* including the vernalization pathway, photoperiod pathway, gibberellin (GA) pathway, ambient temperature pathway, aging pathway and autonomous pathway (Fornara *et al.*, 2010; Kim, 2020). These pathways converge on floral integrators and floral meristem which integrate signals from a set of flowering pathways and then activate downstream genes to induce or repress the floral transition. Photoperiodism, or day-length sensing, has been considered a highly reliable cue of the changing seasons, the magnitude of annual change being determined by degrees north or south of the equator (Andrés and Coupland, 2012; Preston and Fjellheim, 2020). Besides, plants have evolved the ability to measure environmental photoperiod to ascertain time of year (Borchert *et al.*, 2005; Itoh *et al.*, 2010). The

photoperiodic response is generated by the integration of external signals (e.g. light) into internal circadian clock (Song *et al.*, 2010, 2015; Johansson and Staiger, 2015).

Currently, molecular mechanisms of photoperiod induction of flowering have been well studied in the two model species, *A. thaliana* and rice (Fornara *et al.*, 2010; Song *et al.*, 2015; Wei *et al.*, 2020). In *A. thaliana*, *GIGANTEA* (*GI*) activates flowering by promoting *FLOWERING LOCUS T* (*FT*) expression during LD through post-transcriptional inactivation of *CONSTANS* (*CO*) transcriptional repressors (Park *et al.*, 1999; Suárez-López *et al.*, 2001; Sawa *et al.*, 2007; Sawa and Kay, 2011). *CO*, a zinc-finger transcription factor, encoding the BBX domain protein, and the mRNA transcript is regulated by the circadian clock (Putterill *et al.*, 1995; Suárez-López *et al.*, 2001; Shim *et al.*, 2017). In LD, *GI* interacts with the blue light photoreceptor *FLAVIN-BINDING*, *KELCH REPEAT*, *F-BOX1* (*FKF1*) to form a dimeric E3 ubiquitin ligase complex that disrupts *CYCLING DOF FACTORs* (*CDFs*), which subsequently bind to the *CO* promoter region to promote *CO* transcription in the afternoon and towards dusk (Kim *et al.*, 2007, 2013; Sawa *et al.*, 2007; Fornara *et al.*, 2009; Hwang *et al.*, 2019). Besides the major *GI-FKF1-CDFs* module, the effects of *CO* expression and rhythmicity of *FT* mRNA abundance are mediated at the transcriptional and post-transcriptional level, including regulation by transcription factors (Ito *et al.*, 2012), alternative splicing (Gil *et al.*, 2017), photoreceptors (Valverde *et al.*, 2004; Song *et al.*, 2014; Hwang *et al.*, 2019), post-translational modifications (Sarid-Krebs *et al.*, 2015). The *GI-CO-FT* regulate model parallels an evolutionarily conserved system among the Gramineae species, such as rice, sorghum and maize (Yano *et al.*, 2000; Izawa *et al.*, 2011; Yang *et al.*, 2014*b*; Liu *et al.*, 2015; Abdul-Awal *et al.*, 2020). In rice,

OsGI upregulates *HEADING DATE 1* (*Hd1*), the ortholog of *CO*, which regulates the expression of *HEADING DATE 3a* (*Hd3a*), the ortholog of *FT,* to promote flowering in SD and delay flowering in LD (Hayama *et al.*, 2003; Izawa *et al.*, 2011; Goretti *et al.*, 2017).

Nevertheless, many photoperiod-pathway genes have in common between the grasses and *A*. *thaliana*, some key differences still exist. Another flowering time regulated model is *GRAIN NUMBER*, *PLANT HEIGHT AND HEADING DATE 7-EARLY HEADING DATE 1* (*Ghd7-Ehd1*), which has been found in rice, sorghum and maize but is absent from *A*. *thaliana* (Doi *et al.*, 2004; Xue *et al.*, 2008; Yang *et al.*, 2013; Murphy *et al.*, 2014; Song *et al.*, 2015). *Ghd7* is considered to be an evolutionarily new gene in the grasses as no orthologs were found in *A*. *thaliana* (Xue *et al.*, 2008; Yang *et al.*, 2012). *Ghd7*, encoding a CO, CO-LIKE AND TIMING OF CAB1 (CCT) domain protein, has been considered to be a key regulator of rice specific flowering pathway and also contributes to rice yield potential (Xue *et al.*, 2008). Enhanced expression of *Ghd7* under LD morning strongly inhibits the flowering activator *Ehd1* (Xue *et al.*, 2008; Itoh *et al.*, 2010). *Ehd1* encoding a type-B response regulator (RR) is a unique gene that induces the expression of *Hd3a* and *RICE FLOWERING LOCUS T1* (*RFT1*), two florigens in rice (Doi *et al.*, 2004; Itoh *et al.*, 2010; Zhao *et al.*, 2015). *Ehd1* is also regulated by multiple genes, including *Hd1*, *GI*, *Ghd7*, *PSEUDORESPONSE REGULATOR PROTEIN 37* (*PRR37*), and *GRAIN NUMBER*, *PLANT HEIGHT AND HEADING DATE 8* (*Ghd8*) (Xue *et al.*, 2008; Wei *et al.*, 2010; Yan *et al.*, 2011; Gao *et al.*, 2014; Gómez-Ariza *et al.*, 2015). Besides, *Ghd7* is also known as *VERNALIZATION 2* (*VRN2*) in wheat and barley is a negative regulator of

flowering that is down-regulated by *VERNALIZATION 1* (*VRN1*)/*FRUITFULL* (*FUL1*) during vernalization, specifically in core Pooideae taxa (Yan *et al.*, 2004; Woods *et al.*, 2016).

In sorghum, there is a similar but not identical flowering time pathway. Though sorghum has orthologs of major components of the *GI-Hd1*/*CO-Hd3a* pathway in rice, it lacks *RFT1*. *SbGI* plays a role in regulating *SbCO* and *SbEhd1* expression, subsequently promoting expression of florigen genes that serves to accelerate flowering under both LD and SD (Abdul-Awal *et al.*, 2020). *FT*, acting as florigen genes, belongs to the phosphatidylethanolamine binding protein (PEBP) gene family (Tamaki *et al.*, 2007; Danilevskaya *et al.*, 2008). Of the PEBP-family genes in sorghum, sorghum *CENTRORADIALIS 15* (*SbCN15*), the ortholog of rice *Hd3a*, may modify flowering time in a photoperiod-insensitive manner (Murphy *et al.*, 2011, 2014; Yang *et al.*, 2014*a*). Another two PEBP-family genes in sorghum are *SbCN8* and *SbCN12*, the co-linear ortholog of maize *ZCN8* and *ZCN12*, which function as floral activators and are involved in photoperiod sensitivity in maize (Meng *et al.*, 2011; Castelletti *et al.*, 2020). *SbCO* acts as an activator of flowering in SD by inducing the expression of *SbEhd1*, *SbCN8* and *SbCN12*, whereas in LD, *SbCO* activity is inhibited by *SbPRR37* (Yang *et al.*, 2014*b*). *SbPRR37* [*Maturity1*(*Ma1*)] and *SbGhd7* (*Ma6*), which are promoted by sorghum *PHYTOCHROME B* (*SbPhyB*), inhibit flowering by decreasing the expression of *SbEhd1*, *SbCN8 SbCN12* and *SbCN15* under LD (Murphy *et al.*, 2011, 2014; Yang *et al.*, 2014*a*). *Ma2* delayed flowering in LD by selectively enhancing the expression of *SbPRR37* and *SbCO* (Casto *et al.*, 2019).

Moreover, in rice, several flowering time genes have recently been identified to participate in either of the two main independent signaling pathways or even link them. *Ghd7* and *Ghd8* are two major flowering genes and allelic variations in both genes play crucial roles in the wide adaptability of cultivated rice around the world (Xue *et al.*, 2008; Yan *et al.*, 2011; Fujino *et al.*, 2013, 2019; Li *et al.*, 2015; Zhang *et al.*, 2015; Zong *et al.*, 2021). When rice expanded to northern areas, weak or non-functional alleles of the genes were generated and selected on the domestication. Loss-of-function *Ghd7* in rice, *Ghd7-0a*, caused by premature stop codons, is important for extremely early flowering time for adaptability to northern areas (Xue *et al.*, 2008; Fujino *et al.*, 2019; Fujino and Yamanouchi, 2020). *Ghd8*/*DTH8*, encoding a CCAAT-box binding factor, known as a HEME ACTIVATOR PROTEIN 3/NUCLEAR FACTOR YB (HAP3/NF-YB) protein, is identified as a major effect locus affecting flowering with the dual function to inhibit flowering under LD conditions and promote flowering under SD conditions by regulating *Ehd1* and *Hd3a* (Wei *et al.*, 2010; Yan *et al.*, 2011; Dai *et al.*, 2012). The non-functional allele of *Ghd8* (*DTH8*/*LHD1*/*Hd5*/*LH8*) with a 19 bp deletion has been selected in rice and used widely for breeding early heading varieties in Hokkaido, Japan (Fujino *et al.*, 2013), and also contributes to early flowering varieties in the Northeastern China (Li *et al.*, 2015; Zhang *et al.*, 2015). Recent studies in rice showed that *Hd1*, harboring the CCT-domain protein, can be switched-out to form a heterotrimeric complex with Ghd8/OsNF-YB11 and OsHAP5/NF-YC, subsequently, this complex activates the transcription of *Ghd7* by binding directly to the promoter region of *Ghd7*, and further suppress expression of *Ehd1* and *Hd3a,* delaying flowering under LD (Wang *et al.*, 2019*a*; Zong *et al.*,

2021). As there are multiple targets for the Ghd8-OsHAP5b-Hd1 complex, and this complex fine tune the spatiotemporal regulation of gene expression in the flowering network (Wang *et al.*, 2019*a*). Besides, in SD conditions, *Hd1* may also form a complex with a blue light-responsive flavin mononucleotide-binding protein OsHAL3, which promotes the expression of *Hd3a* (and possibly *RFT1*) by binds to their promoter region (Su *et al.*, 2016).

1.4 Progress on genetics of flowering regulation in *Miscanthus*

To date, information on the genetics of flowering regulation in *Miscanthus* is in its infancy (Atienza *et al.*, 2003; Jensen *et al.*, 2011, 2013; Gifford *et al.*, 2015; Dong *et al.*, 2021). Genetic linkage maps revealed nineteen flowering time quantitative trait loci (QTLs) in *M. sinensis* (Atienza *et al.*, 2003; Gifford *et al.*, 2015; Dong *et al.*, 2018; Jensen *et al.*, 2021). Atienza *et al*. (2003) located five putative flowering QTLs in *M. sinensis*, which might be age-dependent or interaction between genotype and environment, and only QTL F12 on linkage group (LG) 1 was detected in both years whereas the rest were only detected in one year. Gifford *et al.*, (2015) found a *Miscanthus* QTL that corresponded to sorghum maturity gene *Ma5* (*PHYTOCHROME C, PhyC*). Dong *et al*. (2018) found one *Miscanthus* flowering QTL on LG 2 that corresponded to sorghum maturity gene *Ma3* (*PhyB*) (Yang *et al.*, 2014*a*) and another located on LG 1 that corresponded to the *ASYMMETRIC LEAVES-like1* gene, which controls proximal-distal patterning in *A. thaliana* petals (Chalfun-Junior *et al.*, 2005). Jensen *et al*. (2021) reported eleven flowering QTLs on LG 4 in *M. sinensis*, three of which were robust QTLs related to the age-dependent flowering pathway (*SQUAMOSA PROMOTER BINDING*

PROTEIN-LIKE and *MADS-box SEPELLATA2*) and the GA pathway (gibberellin-responsive bHLH137). However, the functions of these candidate flowering time genes in *Miscanthus* QTLs have yet to be verified, and allelic sequence variation for these genes has yet to be described. At present, *Hd1*/*CO* is the only candidate flowering time gene that has been screened in *M. sinensis* for sequence diversity and its geographic distribution, with large differences observed between accessions from Asian mainland and Japanese archipelago (Nagano *et al.*, 2015).

1.5 Research objectives

Previous studies have demonstrated the importance of *Ghd8* and *Ghd7* to ensure flowering time control of rice in response to photoperiod. The central theme of present research was to identify two major flowering genes *Ghd8* and *Ghd7* in *M. sinensis*, and characterize their gene expression patterns in response to LD and SD among twelve *M. sinensis* that originate from different latitudes and represent different genetic groups. Based on allelic and deduced amino acid diversity, the aim of the present study is to investigate their geographic distribution, and the relationship between the variants and day to flower. To reveal the possible flowering regulation pathway in *M. sinensis*, the mRNA transcript level of partial up- and downstream flowering genes of *Ghd8* and *Ghd7* were examined in *M. sinensis* under LD and SD.

Chapter 2

Characterization of the *Ghd8* flowering time gene

2.1 Introduction

The genus *Miscanthus* is a rhizomatous, self-incompatible, C_4 perennial grass that has a natural distribution from the tropics to ~50 °N in East Asia and Oceania (Hodkinson *et al.*, 2015), including *M. sinensis*, *M. floridulus* and *M. sacchariflorus*, and is closely related to sugarcane and sorghum. Owing to its environmental adaptability, *Miscanthus* is used as forage for livestock feed, as an ornamental for landscapes, and as a bioenergy crop that provides high yields with low nutrient requirements (Greef and Deuter, 1993; Heaton *et al.*, 2008). Controlling flowering time of *Miscanthus* is benefit for obtaining high-biomass yield under different environments (Jensen *et al.*, 2013; Robson *et al.*, 2013), improving biomass quality (Iqbal *et al.*, 2017), assisting intra- and interspecific hybridizations between *Miscanthus* and *Saccharum* (Dong *et al.*, 2021), and developing seed-based hybrid cultivars of *Miscanthus* to reduce the cost of establishment and accelerate domestication relative to the current standard approach of vegetatively propagating rhizomes of *M* × *g* (Jensen *et al.*, 2011; Hastings *et al.*, 2017). To date, the understanding of flowering time regulation in *Miscanthus* at the molecular level is still limited (Atienza *et al.*, 2003; Jensen *et al.*, 2011, 2013; Dong *et al.*, 2021). Genetic linkage maps revealed nineteen flowering time quantitative trait loci (QTLs) in *M. sinensis* (Atienza *et al.*, 2003; Gifford *et al.*, 2015; Dong *et al.*, 2018; Jensen *et al.*, 2021). While, these candidate flowering time genes in *M. sinensis* QTLs have yet to be confirmed. Until now, *Hd1*/*CO* is the only candidate flowering time gene

that has been screened in *M. sinensis* for sequence diversity and its geographic distribution, with large differences found among accessions from the Asian mainland relative to those from the Japanese archipelago (Nagano *et al.*, 2015).

Though flowering regulation in *M. sinensis* was demonstrated to be complex affected by multiple factors (Jensen *et al.*, 2011), recently, *M. sinensis* was reported as a facultative SD plant and photoperiod has a strong effect on flowering time of *M. sinensis* (Dong *et al.*, 2021). Thus, it would be desirable to study photoperiod regulation of flowering time in *M. sinensis*. The major regulatory genes for photoperiod control of flowering have been evolutionarily conserved in flowering plants but their specific effects can vary greatly among genera and species (Song *et al.*, 2015). To date, the photoperiod regulation of flowering has been extensively investigated in the SD plant rice, and two independent genetic pathways have been identified (Song *et al.*, 2015). One is the rice *OsGI-Hd1-Hd3a* pathway (Hayama *et al.*, 2003; Izawa *et al.*, 2011; Goretti *et al.*, 2017), which has been conserved in the SD plant sorghum (Yang *et al.*, 2014*b*), and is orthologous with the *GI-CO-FT* pathway in the LD plant *A. thaliana* (Song *et al.*, 2015). Another flowering time pathway is *Ghd7-Ehd1-Hd3a*, which has been found in rice, sorghum and maize but is absent from *A. thaliana* (Doi *et al.*, 2004; Xue *et al.*, 2008; Yang *et al.*, 2013; Murphy *et al.*, 2014; Song *et al.*, 2015).

Recently, *Ghd8* (*DTH8*/*LHD1*/*Hd5*/*LH8*) has been found to be a key regulator of the *Ghd7-Ehd1-Hd3a* pathway in rice (Wang *et al.*, 2019*a*). *Ghd8* was initially identified as *HAP3b* in *A. thaliana*, which can promote flowering in *A. thaliana* by enhancing the expression of key flowering time genes, such as *FT* and

SOC1, under LD (Cai *et al.*, 2007). In rice, *Ghd8*, has a dual function to inhibit flowering under LD and promote flowering under SD by regulating *Ghd7*, *Ehd1*, *RFT1* and *Hd3a* (Wei *et al.*, 2010; Yan *et al.*, 2011). In particular, *Ghd8* encodes a protein transcription factor, HAP3/NF-YB, that in rice binds to CCAAT motif in the promoter region of *Ghd7*, as part of a complex with HD1 and OsHAP5b (Wang *et al.*, 2019*a*). In rice, a 19 bp deletion in *Ghd8*, causes a loss-of-function that confers early flowering and thus adaptation to high latitudes; this allele is widely distributed among cultivars from Northern China and Japan (Fujino *et al.*, 2013; Li *et al.*, 2015), and has been selected and used widely for breeding early heading varieties in Hokkaido (Fujino *et al.*, 2013). Therefore, *Ghd8* plays a key role in the domestication and adaptation of rice in Hokkaido. It is worthwhile to investigate if a similar process occurred in *M. sinensis* during its migration northward after the last glacial maximum. *Ehd1* in rice is induced by blue light in the morning, and *Ghd7* suppression of *Ehd1* is induced by red-light in the morning under LD thereby suppressing flowering, whereas under SD, the peak of *Ghd7* expression shifts to night, and this misaligned timing allows *Ehd1* to induce *Hd3a* and promote flowering (Song *et al.*, 2015). Genomic synteny and collinearity are common features in the Poaceae (Gale and Devos, 1998; Glémin and Bataillon, 2009), and has also been confirmed among rice, sorghum, switchgrass and *M. sinensis* genomes (Salse *et al.*, 2009; Kim *et al.*, 2012; Swaminathan *et al.*, 2012; Dong *et al.*, 2018; Mitros *et al.*, 2020; Jensen *et al.*, 2021). Previous studies have identified genes/QTLs under parallel evolution across grass species (Ming *et al.*, 2002; Hu *et al.*, 2003; Yang *et al.*, 2012; Liu *et al.*, 2015; Nagano *et al.*, 2015; van der Weijde *et al.*, 2017*b*; Jensen *et al.*, 2021). To date, there have been no reports of *Ghd8* in

C_4 bioenergy crops such as sorghum, switchgrass and *Miscanthus*. Thus, a key question of this chapter seeks to answer is the following: does *M. sinensis* have a functional *Ghd8* that contributes to the regulation of flowering time? Moreover, we expect that if *Ghd8* regulates flowering in *M. sinensis*, the gene's expression in the day will follow a pattern of differential flowering times under LD relative to SD. In this study, the ortholog of *OsGhd8* in a mini-core collection of *M. sinensis* was cloned with the aim to 1) characterize allelic and deduced amino acid sequence diversity and geographic distribution, and 2) determine expression patterns in response to photoperiod and relate these to previously obtained data on days to first flower under LD and SD.

2.2 Materials and methods

2.2.1 Plant materials and growth conditions

Twelve *Miscanthus* accessions (clones maintained by vegetative propagation) were studied for gene sequence variation and expression over time in response to two photoperiod treatments (15 h, LD; 12.5 h, SD) (Table 2.1). The twelve accessions included eleven *M. sinensis* from known locations in China and Japan, representing latitudes ranging from 18 °N to 45 °N, and one *M. floridulus* from 20.9 °S in New Caledonia (*M. floridulus* was considered to be conspecific with *M. sinensis* (Clark *et al.*, 2014, 2015) and hereafter refer to the entire panel as *M. sinensis*). The *M. sinensis* accessions represent six genetic groups that were previously identified by Clark *et al*. (2014, 2015). Dong *et al*. (2021) previously evaluated the same twelve accessions for days to first flowering under day lengths of 15, 12.5 and 10 h in controlled environment chambers, and observed strong flowering time responses

that varied by latitude of origin. In the current study, six pots of each accession were established by planting rhizomes in 2 L plastic pots containing soilless medium consisting of compost, vermiculite, calcined clay, and peat moss (Forex Mori Sangyo Co., Ltd., Hokkaido, Japan) and growing these in a greenhouse at Hokkaido University in Sapporo, Japan (43.1 °N, 141.3 °E), with natural photoperiod.

After 40 days of establishment in the greenhouse, the *Miscanthus* plants were cut to 5 cm above the soil surface and moved into growth chambers (BioTRON LH-350S, NK Systems, Nippon Medical & Chemical Instruments Co., Ltd., Osaka, Japan) under constant LD (15 h light/9 h dark). Pots were rotated randomly inside and between the chambers on a daily basis to minimize between-chamber and within-chamber environmental effects. The growth chambers provided 400 ± 50 μmol m^{-2} s^{-1} of photosynthetically active radiation with fluorescent lamps (Hitachi FLR40S-EX-N/M/36-A, Hitachi, Ltd., Tokyo, Japan), as measured with a quantum sensor (MIJ-14PARII, Environmental Measurement, Fukuoka, Japan). After 30 days of establishment in the chambers, the plants were subjected to one of two day-length treatments: LD (15 h light/9 h dark) and SD (12.5 h light/11.5 h dark), with three pots per accession given LD and three given SD. The temperature in the chambers was a constant 23 °C for the duration of the experiment. At planting and again at the start of each experiment, 15 g of 12-9-12 compound fertilizer (Kumiai Grassland No. 8; Hokkaido Fertilizer Co., Ltd., Japan) was added to each pot. Irrigation was provided to each pot each day. At day 38, one week after commencement of the LD or SD treatment, the three topmost leaves from each of the three pots per accession within each treatment were harvested and pooled at Zeitgeber times (ZT) of 3, 9, 15 and 21 h for one 24 hour light-dark cycle.

Table 2.1 Provenance, flowering time under short or long days, and amino acid sequence diversity for two homoeologous *Ghd8* loci in a mini-core panel of 11 *Miscanthus sinensis* and one *Miscanthus floridulus* accessions.

Genotypes	Ploidy	Lat	Long	Genetic group†	Genetic group color code†	Days to first flowering‡		Variant types classified by predicted protein in *Ghd8* homoeologs			
						12.5 h	15 h	MsiGhd8A		MsiGhd8B	
M. sinensis 'Teshio'	2x	44.9	141.9	Northern Japan	Blue		66	A1	A2	B1	B2
M. sinensis 'Sugadaira'	2x	36.0	138.1	Southern Japan	Yellow		96	A3	A4	B3	B4
M. sinensis 'Miyazaki'	2x	31.8	131.4	Southern Japan	Yellow	61	167	A5	A6	B5	B6
M. sinensis 'PMS-436'	2x	41.3	123.7	Korea/North China	Red		115	A3	A7	B7	B8
M. sinensis 'PMS-164'	2x	37.3	114.3	Yangtze-Qinling	Green		130	A8	A8	B8	B8
M. sinensis 'PMS-306'	2x	29.9	118.8	Yangtze-Qinling	Green	84	173	A8	A9	B8	B8
M. sinensis 'PMS-226'	2x	26.6	106.8	Sichuan Basin	Orange	76	189	A1	A7	B9	B10
M. sinensis 'Onna-1a'	2x	26.5	126.8	SE China plus tropical	Purple		274	A10		B11	B12
M. sinensis 'PMS-359'	2x	22.9	112.3	SE China plus tropical	Purple	81	179	A11	A7	B8	B13
M. sinensis 'PMS-375'	2x	19.6	110.3	SE China plus tropical	Purple	142		A11		B9	B9
M. sinensis 'PMS-382'	2x	18.9	109.5	SE China plus tropical	Purple	184		A11	A12	B1	B1
M. floridulus 'US56-0022-03'	2x	-20.9	165.3	SE China plus tropical	Purple	114		A11	A13	B14	B15

†*M. sinensis* genetic groups determined by Clark *et al*. (2014, 2015). ‡Days to first flowering under short days (12.5h) or long days (15h) by Dong *et al*. (2021); empty cells indicate flowering did not occur. A or B prefix indicates putatively functional alleles types based on predicted amino acid sequence variants in the A and B subgenomes, respectively. Cells with different colors represent different variant types that occurred in more than one accession; variant types that were observed only once have a gray background; blank cells of *MsiGhd8A* indicated that only one allele type was detected in Onna-1a and PMS-375, and therefore these two accessions were homozygous at *MsiGhd8A*.

2.2.2 Genomic DNA extraction and isolation of *Ghd8* in *Miscanthus*

Genomic DNA was isolated from young, healthy leaves by the modified cetyltrimethylammonium bromide (CTAB) (Scholin *et al.*, 1994) protocol using the DNeasy Plant Mini Kit (Qiagen, Tokyo, Japan) according to the manufacturer's instructions. Gene-specific primers (Forward primer 1: 5'-GAAAGGCGATTAA GAGGAGAAT-3'; Forward primer 2: 5'-CACCATAAGCTAGCTGACTAGCT-3'; Reverse primer 1: 5'-GCAAGTATCGTTTGTCGTCGTCTT-3') for *Ghd8* were designed by aligning multiple sequences retrieved from the *Miscanthus sinensis* v7.1 genome (Mitros *et al.*, 2020) and its close relative sorghum using the *Sorghum bicolor* v3.1 genome (McCormick *et al.*, 2018) from Phytozome v13 (https://phytozome-next.jgi.doe.gov). Amplification of *Ghd8* was accomplished by polymerase chain reactions (PCRs) containing 30 ng of total genomic DNA as a template and LA Taq polymerase (TaKaRa Bio, Shiga, Japan). Amplification conditions were 1 min at 95 °C, followed by 30 cycles of 30 sec at 95 °C, 30 sec at suitable primer temperature and 1 m 30 sec at 72 °C. PCR products were separated on 0.8 % agarose gels by electrophoresis. Amplified bands of desired molecular weight were eluted from the agarose gel with the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nager GmbH & Co. KG, Düren, Germany) and cloned into a pGEM-T Easy vector (Promega, Madison, WI, USA) using the TA-Blunt Ligation Kit (Nippon Gene Co., Ltd., Toyama, Japan) following the manufacturer's instructions. Positively transformed colonies were selected on blue-white selection on ampicillin/IPTG/X-Gal LB plates, and plasmids were purified using a High Pure Plasmid Isolation Kit (Roche, Sigma-Aldrich, Tokyo, Japan). About 20 plasmid clones of each genotype were sequenced in both directions with a BigDye

Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) via an ABI PRISM 3130 Genetic Analyzer (Life Technologies, Carlsbad, USA) according to the manufacturer's instructions. To identify true alleles and to limit the potential for misidentifying point mutations and indels resulting from PCR and sequencing errors as true alleles, we set a quality-control threshold of at least three colonies with the identical sequence for inclusion in further analysis and reporting.

2.2.3 RNA isolation and quantitative reverse transcription-PCR analysis

Leaves were sampled from fully expanded healthy leaves at ZT 3, 9, 15 and 21 h in the growth chamber. All samples were immediately frozen in liquid nitrogen and stored at -80 °C until analysis. Total RNA was isolated from frozen leaves with a Favorgen® Plant Total RNA Extraction Mini Kit (Favorgen Biotech Corp., Taiwan) and treated with DNase I (TaKaRa Bio, Shiga, Japan) to remove contaminating genomic DNA. cDNA was synthesized from purified RNA using an oligo (dT) 20 primer and random hexamer primers with Invitrogen™ M-MLV Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA) according to Dwiyanti *et al*. (2011). The transcript levels for *Ghd8* were determined by quantitative real-time PCR (qRT-PCR). The PCR reactions (20 μL) contained 4.6 μL of the cDNA synthesis reaction mixture diluted to 1/15 th of its original volume, 5 μL of 1.2 μM primer premix, 0.4 μL ROX Reference Dye (50 ×) and 10 μL of TB Green® Premix Ex Taq™ II (Tli RNaseH Plus) (TaKaRa Bio, Shiga, Japan). Expression levels were determined on a StepOnePlus™ Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) with cycling conditions of 95°C for 5 min, followed by 40 cycles of 95 °C for 10 sec, 60 °C for 20 sec and 72 °C for 30 sec. Values were

normalized to *ACTIN* (Misin17G008500) as an internal control. A reaction mixture without reverse transcriptase was also used as a control to confirm the absence of genomic DNA contamination. Amplification of a single DNA fragment was confirmed by melting-curve analysis of quantitative PCR and gel electrophoresis of the PCR products. Relative changes in gene expression were estimated following the $2^{-\Delta\Delta Ct}$ method (Bookout and Mangelsdorf, 2003). Averages and standard errors of relative expression levels were calculated for three independently synthesized cDNAs. The forward primer used for *ACTIN* (Misin17G008500) gene expression was 5'-AGGGCTGTTTTCCCTAGCATCG T-3', and the reverse primer was 5'-GGGTACTTGAGCGTGAGAATACCTC-3'. Primers were designed for *MsiGhd8* based on the putative functional alleles. The forward primer used for *MsiGhd8A* (Misin13G040800) gene expression was 5'-CTCAACCGCTACCGCGAGGTC-3', and the reverse primer was 5'-TCATCCGCCGCGCCATCT-3'. The forward primer used for *MsiGhd8B* (Misin07G127500) gene expression was 5'-ACGTCGGGCTCATGATGGGAGCA-3', and the reverse primer was 5'-ATACGACTTCCGTGCTGCCGT-3'.

2.2.4 Data analysis

The nucleotide sequences were assembled with ATGC v6 software (GENETYX Co., Tokyo, Japan). *O. sativa*, *S. bicolor*, *M. sinensis* genome sequences (Phytozome v13, 100kb) spanning homologs of *Ghd8* were used for microsynteny /collinearity analysis, which was determined and visualized by Genome Evolution Analysis (GEvo) (http://genomevolution.org/CoGe/GEvo.pl) and the high-resolution sequence analysis tool from Accelerating Comparative Genomics (CoGe)

toolkit (http://genomevolution.org/CoGe/). Multiple alignments of nucleotide and amino acid sequences were implemented in MEGA X (Zuckerkandl and Pauling, 1965; Kumar *et al*., 2018; Stecher *et al*., 2020), using ClustalW v2.1 (Thompson *et al*., 1994) with default settings. Phylogenetic trees were generated in MEGA X (Kumar *et al*., 2018; Stecher *et al*., 2020) using the Neighbor-Joining (NJ) method (Saitou and Nei, 1987) with the substitutional model of Kimura 2-parameter (Kimura, 1980). The corresponding sequences of rice and sorghum were used as an out-group. Support for each node was tested with 1,000 bootstrap repetitions (Felsenstein, 1985). The trees were edited and visualized in FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/). Relative changes in mean ± standard error of the mean (SE) gene expression were analyzed in Microsoft Excel (Microsoft Office 2016, Microsoft Inc., Seattle, WA, USA) and then exported to GraphPad Prism 9 software (GraphPad Software, San Diego, CA, USA) for visualization. Statistical tests for differences among means were conducted by Student's *t*-test or analyses of variances (ANOVAs) using GraphPad Prism 9 software (GraphPad Software, San Diego, CA, USA). The DNA sequences obtained are available from DDBJ (http://www.ddbj.nig.ac.jp/index-e.html) with the accession numbers LC598392 to LC598437.

2.3 Results and discussion

2.3.1 Characterization of *M. sinensis Ghd8*

In *M. sinensis*, two homoeologous *Ghd8* loci, *MsiGhd8A* located on chromosome 13 (Chr.13) and *MsiGhd8*B on chromosome 7 (Chr.07), were identified, with one on each of this paleo-allotetraploid species' subgenomes (Figure 2.1). A total of 46

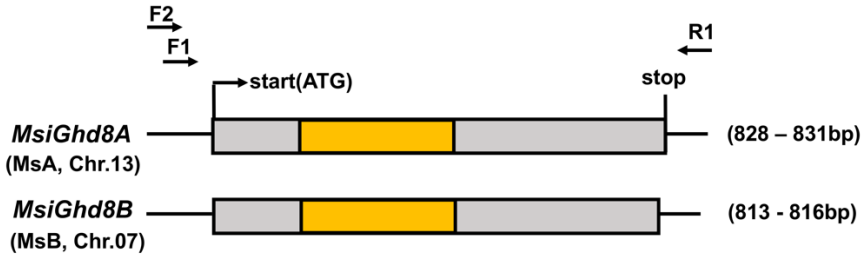
MsiGhd8 alleles were identified from the 12 wild-collected *M. sinensis* accessions (Figure 2.2, Tables S2.1 and S2.2). Sequence alignment indicated that the open reading frame (ORF) lengths of *M. sinensis Ghd8* ranged from 813 to 831 nucleotides, and contained one exon that coded for 270 to 276 amino acid residues (Figure 2.1). Multiple sequence blasting in Phytozome v13 (https://phytozome-next.jgi.doe.gov) revealed that the nucleotide sequences of *M. sinensis Ghd*8 were highly similar to those in other plant species, such as *S. bicolor* (Sobic.007G059500, 88.6% - 92.3%), *O. sativa* (LOC_Os08g07740, 72.2% - 73.3%), *Z. mays* (Zm0001d049485, 82.7% - 86.3%), and *A. thaliana* (AT5G47640, 32.0% - 32.9%). A microsynteny assessment of genomic regions adjacent to *Ghd8* in rice, sorghum and *M. sinensis* identified four colinear genes, including *Ghd8*, aligned with the same relative genomic order in a 100 kbp region, which was consistent with the identification of LOC_Os08g07740 as an ortholog of rice *Ghd8* (Wei *et al.*, 2010) (Figure S2.1 and Table S2.5), and in consistent with the known paleo-duplications (rice Chr.08- sorghum Chr.07, sorghum Chr.07- *Miscanthus* Chr.13 and Chr.07) from the ancestral grass chromosomal groups (Salse *et al.*, 2009; Dong *et al.*, 2018; Mitros *et al.*, 2020). Therefore, based on sequence similarity and gene collinearity, two homoeologous *Ghd8* loci in *M. sinensis* were designated as orthologs of *Ghd8* in rice and sorghum, and probable ortholog of *HAP3b* in *A. thaliana*. Neighbor-Joining (NJ) phylogenetic trees revealed a robust separation of clades representing *MsiGhd8A* (22 alleles) and *MsiGhd8B* (24 alleles) (Figure 2.2). The phylogenetic trees indicated that the sorghum *Ghd8* was more similar to *MsiGhd8B* than *MsiGhd8*A. Two accessions (Onna-1a and PMS-375, 16.7%) were homozygous at the *MsiGhd8A* locus and all accessions were heterozygous at the *MsiGhd8B* locus

(Table 2.1). Pairwise DNA sequence comparisons showed that the similarity of *MsiGhd8A* ORFs ranged from 98.7% (Teshio-Func2 vs. PMS-226-Func2, Teshio-Func2 vs. PMS-382-Func2, Teshio-Func2 vs. US56-0022-03-Func2) to 100% (Sugadaira-Func1 vs. PMS-436-Func1, PMS-164-Func1 vs. PMS-306-Func1, PMS-375-Func1 vs. PMS-382-Func1) (Table S2.1). Similarly, the nucleotide sequence similarity of *MsiGhd8B* ORFs varied from 97.9% (Sugadaira-Func4 vs. US56-0022-03-Func4) to 100% (PMS-164-Func3 vs. PMS-306-Func3, PMS-164-Func4 vs. PMS-436-Func4, PMS-226-Func3 vs. PMS-375-Func3) (Table S2.2).

Comparison of the 46 *MsiGhd8* alleles derived from the 12 wild-collected *M. sinensi*s accessions in this study with the alleles in the *Miscanthus sinensis* v7.1 genome (Mitros *et al.*, 2020), revealed 35 non-synonymous single nucleotide variants (nsSNVs), 36 synonymous single nucleotide variants (sSNVs) and two 3-bp insertions in the ORFs, with some accessions having more than one SNV per allele (Tables S2.1 and S2.2). Considering the fact that the nucleotide diversity cannot exactly represent the protein diversity owing to sSNVs in ORFs, Ghd8 protein variant types were analyzed in the present study (Tables 2.1, S2.3 and S2.4, Figures 2.2 and 2.3). Accounting for nsSNVs, 13 predicted amino acid sequence types of MsiGhd8A and 15 of MsiGhd8B (28 total) were identified from the 12 *M. sinensis* accessions (Tables 2.1, S2.3 and S2.4, Figure 2.2 and 2.3). The amino acid sequence similarity of putatively functional MsiGhd8A and MsiGhd8B variants ranged from 92.1 % to 94.2 %. Notably, the deduced amino acid sequence of Ghd8 in *M. sinensis* indicated that the gene products contain a HAP3/NF-YB DNA-binding domain located from position 53 to 146 (Figure 2.1b), which is critical for the transcription factor function of *Ghd8* gene products in rice (Wei *et al.*, 2010;

Yan *et al.*, 2011) and *A. thaliana* (Kumimoto *et al.*, 2008), and may have a conserved function to regulate flowering time. Though no putatively non-functional alleles were detected, four nsSNVs in the HAP3/NF-YB DNA-binding domain of MsiGhd8 (two in MsiGhd8A and two in MsiGhd8B) were observed in five accessions, with one nsSNV of MsiGhd8A found in each of two accessions (Sugadaira and PMS-436), one nsSNV of MsiGhd8A found in Teshio, and one nsSNV of MsiGhd8B in PMS-226 and another nsSNV of MsiGhd8B found in US56-0022-03 (Table S2.1 and S2.2). The nsSNVs identified in the domain might have an effect on protein stability and function (Liao and Lee, 2010), and additional research is required.

(a)



(b)

```
MsiGhd8A(Chr.13)   1 MKSRKGYGHQGHLLSPVGSPPSDNESDAAAAAATAGGGG-CGSSVG---YCGGGG-ESPAKEQDRFLPIANVSRIMKRSL 75
MsiGhd8B(Chr.07)   1 MKSRKGYGHQGHLLSPVGSPPSDNESGAAAAAATAGGGG-CGSSVG---YCGGGG-ESPAKEQDRFLPIANVSRIMKRSL 75
Sobic.007G059500   1 MKSRKGYGHQGHLLSPVGSPPSDNESGAVAAAATAGGGGGCGSSVG---YCGGGGGESPAKEQDRFLPIANVSRIMKRSL 77
LOC_Os08g07740     1 MKSRKSYG---HLLSPVGSPPLDNESGEAAAAAAAGGGG-CGSSAGYVVYGGGGGGDSPAKEQDRFLPIANVSRIMKRSL 76
Zm00001d049485     1 MKNRKGYGHQGHLLSPVGSPLSDNESG---AAAAAGGGG-CGSSVG---YCGGGGGESPAKEQDRFLPIANVSRIMKRSL 73
AT5G47640          1 ----------------MGD--SDRDSG--------GGQN--GNNQN-------GQSSLSPREQDRFLPIANVSRIMKKAL 45

MsiGhd8A(Chr.13)  76 PANAKISKEAKETVQECVSEFISFVTGEASDKCQREKRKTINGDDLLWAMTTLGFEAYVAPLKSYLNRYREVEGEKAAVL 155
MsiGhd8B(Chr.07)  76 PANAKISKEAKETVQECVSEFISFVTGEASDKCQREKRKTINGDDLLWAMTTLGFEAYVAPLKSYLNRYREAEGEKAAVL 155
Sobic.007G059500  78 PANAKISKEAKETVQECVSEFISFVTGEASDKCQREKRKTINGDDLLWAMTTLGFEAYVSPLKSYLNRYREAEGEKAAVL 157
LOC_Os08g07740    77 PANAKISKESKETVQECVSEFISFVTGEASDKCQREKRKTINGDDLLWAMTTLGFEAYVGPLKSYLNRYREAEGEKADVL 156
Zm00001d049485    74 PANAKISKEAKETVQECVSEFISFVTGEASDKCQREKRKTINGDDLLWAMTTLGFEAYVAPLKSYLNRYREAEGEKAAVL 153
AT5G47640         46 PANAKISKDAKETMQECVSEFISFVTGEASDKCQKEKRKTINGDDLLWAMTTLGFEDYVEPLKVYLQRFREIEGE----- 120

MsiGhd8A(Chr.13) 156 GAGP---GARHGDG--AADD-GPLAAAGVVVAPPSGGAGDRAGHDGAADTGGHDAHVGLMMG-ASSVVGFGAGGG-AAPS 227
MsiGhd8B(Chr.07) 156 GGG-----ARHGDG--VADD-GPLAAGGVIAPPSGGGAGDRAGHDGAADA---DAHVGLMMG-ASSVVGFGAGSA-AAAS 222
Sobic.007G059500 158 GG------ARHGDG--AVDD-GPLAAGGAVPP-SGAAGGDRAGHDGAADA---DAHVGLMMGGASSVVGFGAGGG-AAAS 223
LOC_Os08g07740   157 GGAGGAAAARHGEGG-CCGGGGGGADGVVIDGHYPLAGGLSHSHHGHQQQDG-GGDVGLMMGGGDAGVGYNAGAGSTTTA 234
Zm00001d049485   154 GGG-----ARHGEGGGAADDAGPLAAGG-----GAGDGVDRAGHDD-------DAHVGLMMG--ASSVGFGSGGG-AAPS 213
AT5G47640        121 ---------RTGLG-------------------RPQTGGEVGEHQR--------DAVG-------DGGGFYGGGG--GMQ 155

MsiGhd8A(Chr.13) 228 YY-----AAAARKAYG-GEGS-KVMEFEGIGGEEENGGAQRG--RGFGS-HLHGAVQW-------- 275
MsiGhd8B(Chr.07) 223 YY------TAARRSYGAGEAS-KVMEFEGIGGEEDNGGVQRG--RGFAN-HLHGAVQW-------- 270
Sobic.007G059500 224 YYT----ATAARKAYGAGEGSTKVMEFEGIAGEEENGGVQRG--RGFANNHLHGAVQW-------- 275
LOC_Os08g07740   235 FYAPAATAASGNKAYCGGDGS-RVMEFEGIGGEEESGGGGGGGERGFAG-HLHG-VQWFRLKRNTN 297
Zm00001d049485   214 YY-----AAASRKAYGAGEGS-KVMEFE---GEEENGGVQRG--RGFAS-HLHGAVQW-------- 259
AT5G47640        156 YHQHHQFLHQQNHMYGATGGG----------SDSGGGAASG--RTRT----------------- 190
```

Figure 2.1 Gene structure and multiple alignment analysis of *Miscanthus sinensis Ghd8* homoeologs and their comparison with orthologs from four other plant species. (a) Gene structure of *MsiGhd8A* and *MsiGhd8B*. F, forward primer; R, reverse primer; the primer pairs F1/R1 and/or F2/R1 were used to detect open reading frames (ORFs) for *Ghd8*. The start codon (ATG) and stop codon (TGA) are highlighted in black. The yellow box represents the HAP3/NF-YB domain. (b) Multiple amino acid sequence alignments for Ghd8 from *M. sinensis* (this study), *Sorghum bicolor* (Sobic.007g059500), *Oryza sativa* (LOC_Os08g07740), *Zea mays* (Zm0001d049485) and *Arabidopsis thaliana* (AT5G47640). The HAP3/NF-YB domain is boxed in red. The *M. sinensis* sequence used for alignment are from accession PMS-382.

2.3.2 Geographical distribution of naturally occurring MsiGhd8 protein variants

Some of the MsiGhd8 protein variants were found over a broad geographic range, whereas others had restricted patterns of occurrence (Table 2.1 and Figures 2.2, 2.3 and S2.1). In the A subgenome, variant A1 was the most broadly distributed, with occurrence in accessions that originated from the mid and highest latitudes in this study (PMS-226 from Sichuan Basin and Teshio from Northern Hokkaido Japan) but it was infrequently observed (16.7% of accessions). In contrast, A7 was distributed widely and the second-most frequently observed variant (25% of accessions). A3 was limited to two accessions, one in Northern China and one in Central Japan; however, DNA sequence analysis indicated that A3 and A7 are closely related (Table S2.3) and thus represent a broadly distributed group in mainland Asia and Japan. A11 had a restricted distribution from New Caledonia to Guangdong China with a latitude ranging from 20.9 °S to 22.9 °N and was the most frequent variant (33.3% of accessions), three of which couldn't flower in LD and flower late in SD, but was absent from mid and high latitudes in mainland Asia and Japan. However, phylogenetic analysis of the DNA sequence revealed that A11 and A1 protein variants were closely related and thus represented a widely distributed group from east to west and from north to south. A8 was limited to mid latitudes in mainland Asia. The other eight variants, were each observed in only one accession. A2 and A3, which encode one additional amino acid resulting from the same 3-bp insertion in the nucleotide sequence, were limited to Northern Japan and China., displaying non-flowering under SD but flower earlier in LD (Tables 2.1 and S2.3).

In the B subgenome, variant B1 was observed from Hainan to Hokkaido but infrequently (16.7% of accessions). In mainland Asia, B8 was also broadly distributed from low to high latitude and frequent (25% of accessions). B9 was observed in two accessions, one in Sichuan Basin and one in Southern China. The other twelve variants, were each observed in only one accession. Phylogenetic analyses of DNA sequence indicated the following closely related protein variant groups: B7 and B8; B9 and B10; B1, B4 and B13; B3, B6, B11 and B12 (Figures 2.2 and 2.3). The natural variations of *Ghd8* have been confirmed to contribute greatly to rice adaptation and genetic improvement (Fujino *et al.*, 2013; Li *et al.*, 2015; Zhang *et al.*, 2015; Fujino, 2020). These variants identified in the present study would provide us a theoretical clue to make the utmost of *Ghd8* in modern molecular breeding of *M. sinensis* in future. Allelic-specific molecular makers would be appropriately developed for selection of the favorable genotypes to meet the demand for varieties in different ecotypes.

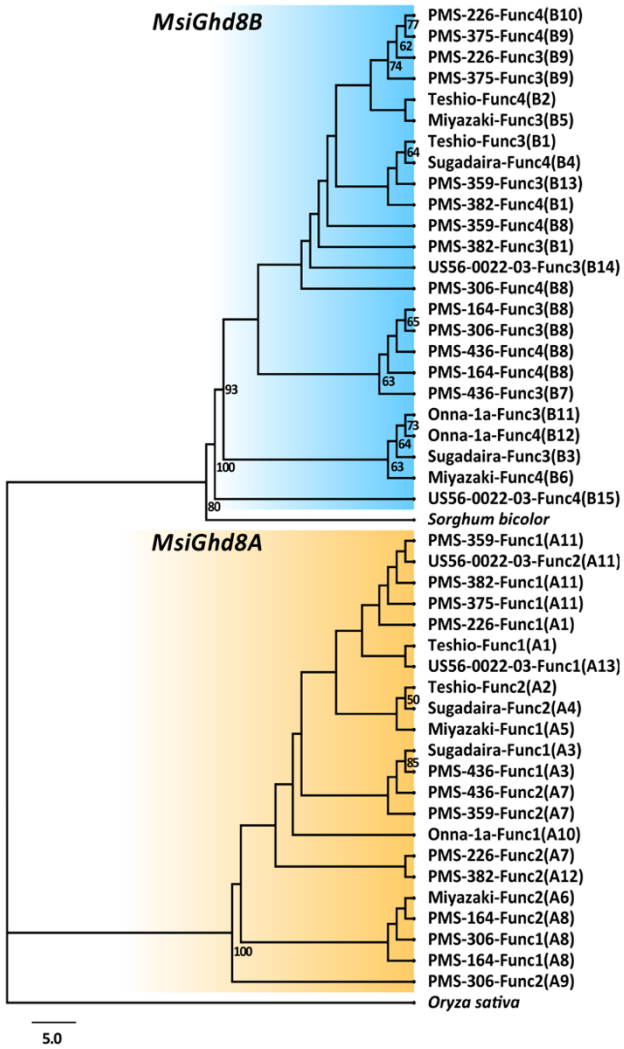


Figure 2.2 Phylogenetic tree inferred by neighbor-joining method for nucleotide sequences of 42 *Ghd8* alleles from 11 accessions of *Miscanthus sinensis* and, four alleles from one *Miscanthus floridulus* accession. *Sorghum bicolor* (Sobic.007g059500) and *Oryza sativa* (LOC_Os08g07740) were used as an out-group. The phylogenetic tree was divided into two clusters, which were classified as *MsiGhd8A* and *MsiGhd8B*, one for each of the two subgenomes. Bootstrap values for nodes supported in >50% of 1000 bootstrap replicates are shown. Allele names with A or B prefix indicate putatively functional alleles types based on predicted amino acid sequence variants, which are named in parentheses and correspond to the names in Figures 2.3 and S2.2, Tables 2.1, S2.1, S2.2, S2.3 and S2.4.

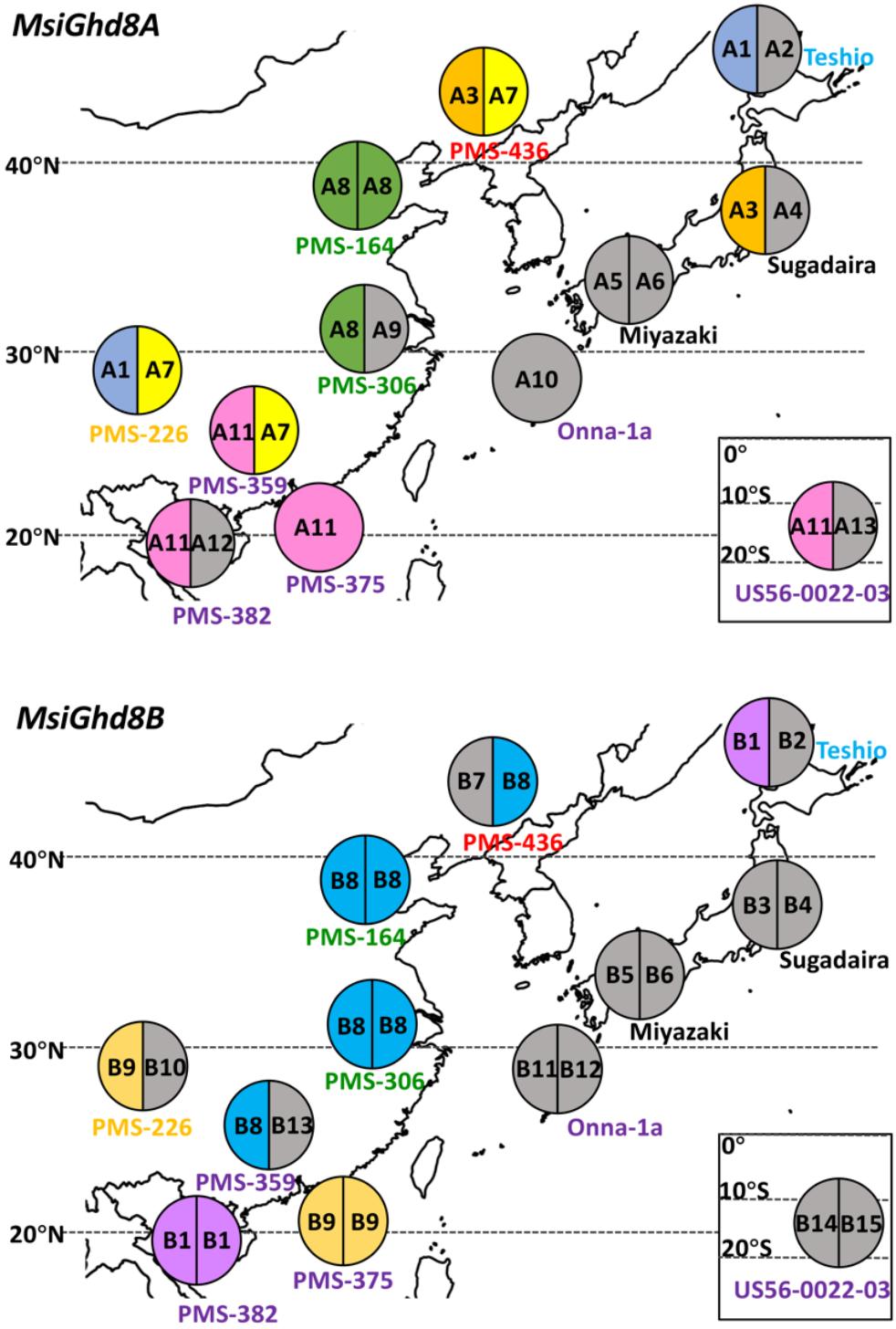


Figure 2.3 Geographical distribution of MsiGhd8A and MsiGhd8B predicted amino acid sequence variant types in *Miscanthus sinensis*. Pie charts with one to two sections represent the number of detected alleles. A or B prefix indicates putatively functional alleles types based on predicted amino acid sequence variants.

Different colors in pie charts represent different variant types that occurred in more than one accession; variant types that were observed only once have a gray background, corresponding to Table 2.1. Accessions' names are colored to represent *M. sinensis* genetic groups previously described by Clark *et al*. (2014, 2015); Sugadaira and Miyazaki are changed from yellow to black for making the map clear. The map image is taken from Wikimedia Commons: https://commons.m.wikimedia.org/wiki/File.

2.3.3 Expressions patterns of *M. sinensis Ghd8*

For each of the *M. sinensis* accessions, two *Ghd8* homoeologs expressed under both LD and SD, and expression of *Ghd8* (assessed as the ratio of *Ghd8*/*ACTIN* mRNA transcript abundance) from the B subgenome was one to two orders of magnitude greater than for the A subgenome (Figure 2.4 and 2.5). The lower expression observed for *MsiGhd8A* than *MsiGhd8B*, was consistent with a previously observed *M. sinens*is genome-wide expression bias in favor of the B subgenome, with ~10% more pairs of genes having higher expression in the B subgenome (Mitros *et al.*, 2020). Thus, *MsiGhd8A* may be a case of reduced- or neo-functionalization, which is common in organisms with duplicated genomes (Lallemand *et al.*, 2020).

The two accessions with the highest day-expression of *MsiGhd8B* under LD (Teshio and Onna-1a) also had the highest expression of *MsiGhd8A* (Figures 2.4 and 2.5). Interestingly, under LD, Onna-1a was the latest flowering of the accessions but Teshio was the earliest flowering, and neither flowered under SD. The three tropical accessions (PMS-375, PMS-382 and US56-0022-03) were among the only three accessions in the panel that did not flower under LD (Figure 2.5, Table 2.1) with *Ghd8* lower expression did not flower under 15 h but did flower under 12.5 h day length. Overall, three patterns of diurnal *MsiGhd8* expression were observed: day peak, night peak, and no clear peak (Figure 2.5). The most common diurnal *MsiGhd8* expression pattern observed was a day peak at ZT9 or ZT15 (Figure 2.5), which is later than the dawn peak that has been reported for rice (Wei *et al.*, 2010; Yan *et al.*, 2011), suggesting that optimal timing may differ between *M. sinensis* and rice.

Given that grasses have multiple pathways to regulate flowering time, including two known major pathways for photoperiod regulation of flowering time that each has multiple modifiers, flowering time predominantly could not be conferred by any one gene, including *Ghd8*. Moreover, in *A. thaliana* and rice, through using yeast and animal systems, it has been demonstrated that HAPs, a CCAAT-box-binding transcription factor, form a heterotetramer or heterotrimer for transcription activation (Ben-Naim *et al.*, 2006; Wenkel *et al.*, 2006). In rice, Ghd8 could interact with OsHAP5b and Hd1, forming a complex under LD, then induce *Ghd7* expression to inhibit *Ehd1* and delay flowering (Li *et al.*, 2015; Zhang *et al.*, 2015, 2019; Nemoto *et al.*, 2016; Du *et al.*, 2017; Liu *et al.*, 2020; Zong *et al.*, 2021). However, *Hd1*/*CO* also competes with the complexes to promote *Hd3a*/*RFT1* expression, creating a tradeoff relationship for photoperiod sensitive flowering under SD conditions (Zong *et al.*, 2021). Thus, the regulatory network controlling flowering time is complex and quantitative, which likely accounts for the great plasticity of this trait in diverse populations. Whether *MsiGhd8* protein can bind these flowering related genes (*Hd1*/*CO* and *Ghd7*) products forming NF-Y complexes as described in rice remains to be confirmed in further studies. It would be desirable to analyze how *Ghd8* regulated the downstream genes in response to day length in *M. sinensis*.

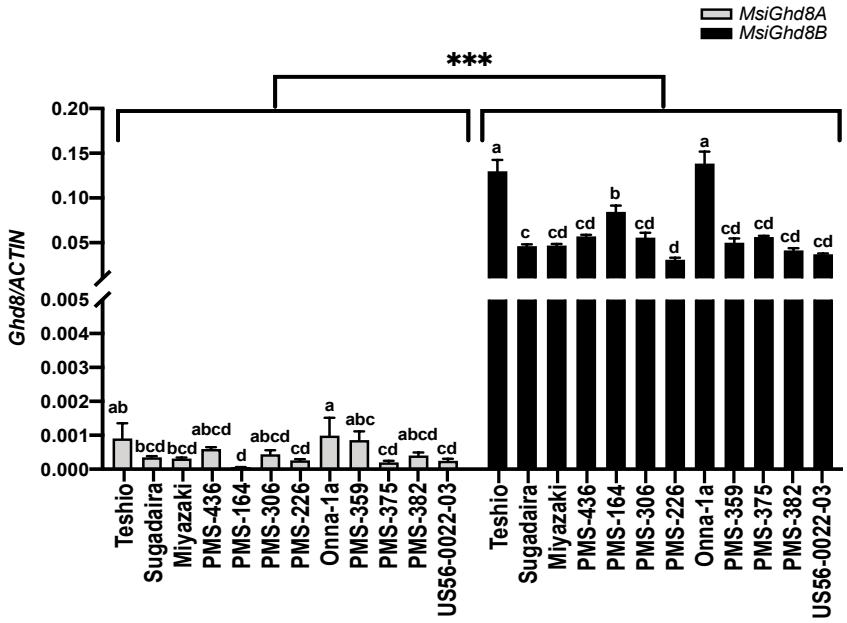


Figure 2.4 Expression of *MsiGhd8* at Zeitgeber time 9 for 12 *Miscanthus sinensis* accessions under long days (15 h, LD). Grey and black represent *MsiGhd8A* and *MsiGhd8B*, respectively. Relative mRNA levels are expressed as the ratios to *ACTIN* transcript levels. Mean ± 1SE for three replications are given for each data point. A different letter on top of a bar indicates significant difference between accessions within each subgenome according to the Tukey HSD (95% family-wise confidence level) multiple comparison tests. *** shown between the two subgenomes indicates a significant difference at $P < 0.001$ according to the Student's *t*-test.

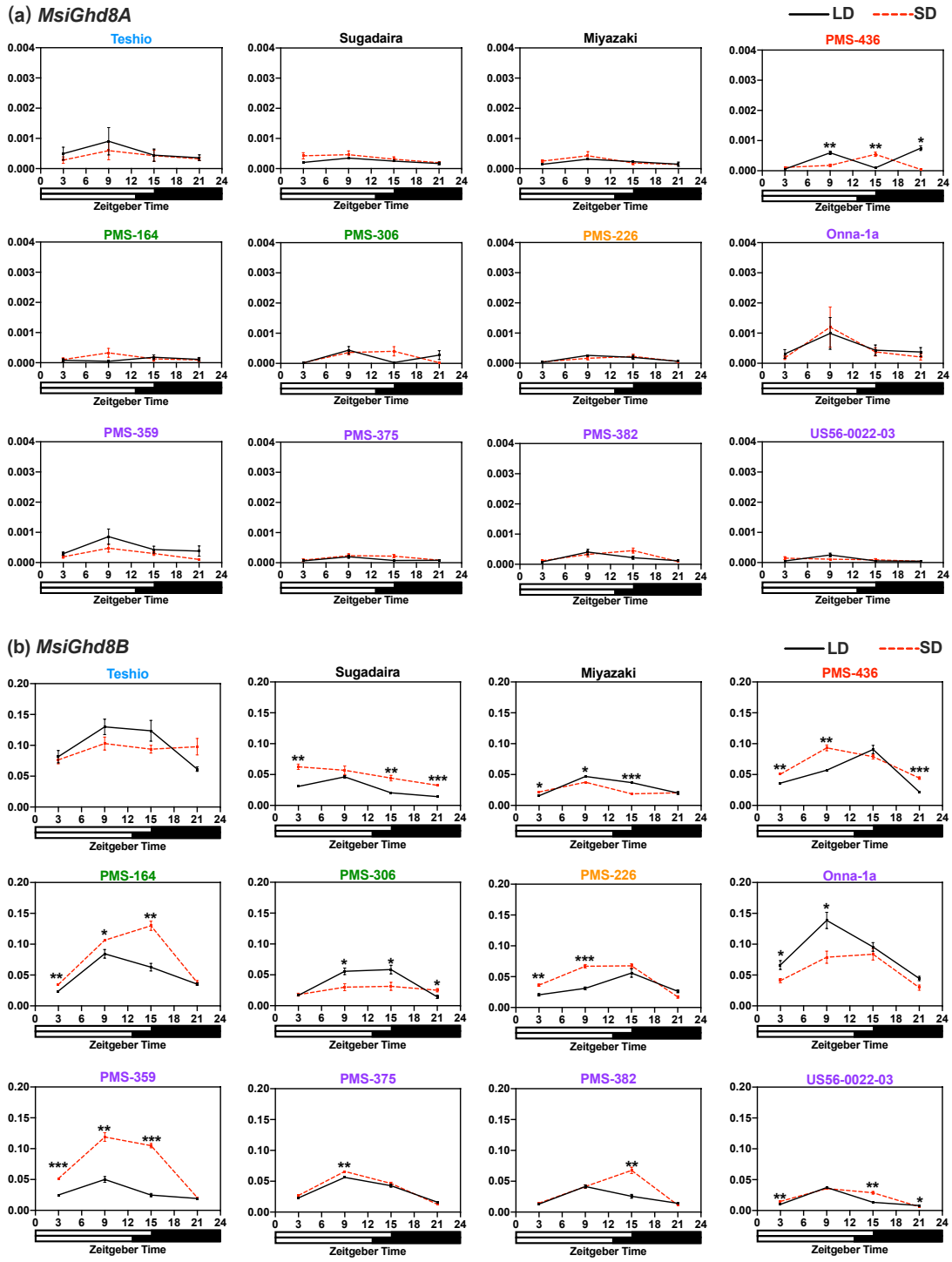


Figure 2.5 Diurnal expression of *MsiGhd8* in 12 *Miscanthus sinensis* genotypes under long days (15 h, LD; solid black lines) and short days (12.5 h, SD; red dashed line). (a) *MsiGhd8A* and (b) *MsiGhd8B*. The *y*-axis represents relative mRNA levels normalized to *ACTIN* transcript levels. The numbers below the *x*-axis indicate

Zeitgeber times (ZT) of the day. The white bar at the bottom of each graph indicates the light period and the black bar indicates the dark period. Mean ± 1SE for three replications are given for each data point. Asterisks indicate significant difference between the two means under LD and SD at the same ZT of the day (Student's *t*-test, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$). No asterisk indicates the difference between the two means is not statistically significant ($P > 0.05$). Accessions' names are colored to represent *M. sinensis* genetic groups previously described by Clark *et al*. (2014, 2015); Sugadaira and Miyazaki are changed from yellow to black for making it clear.

Supplementary materials for Chapter2:

Supplementary Figure S2.1: Chromosome organization of *Ghd8* gene orthologous regions (100 kbp) from *Oryza sativa*, *Sorghum bicolor* and *Miscanthus sinensis*.

Supplementary Figure S2.2: Geographical distribution of *MsiGhd8A* and *MsiGhd8B* alleles in *Miscanthus sinensis*.

Supplementary Table S2.1: A summary of polymorphic sites in open reading frames (ORFs) of *MsiGhd8A* alleles from 12 *Miscanthus sinensis*.

Supplementary Table S2.2: A summary of polymorphic sites in open reading frames (ORFs) of *MsiGhd8B* alleles from 12 *Miscanthus sinensis.*

Supplementary Table S2.3: Protein types of MsiGhd8A from 12 *Miscanthus sinensis*.

Supplementary Table S2.4: Protein types of MsiGhd8B from 12 *Miscanthus sinensis*.

Supplementary Table S2.5: Colinear genes near homologs of *Ghd8* in *Miscanthus sinensis*, *Sorghum bicolor* and *Oryza sativa*.

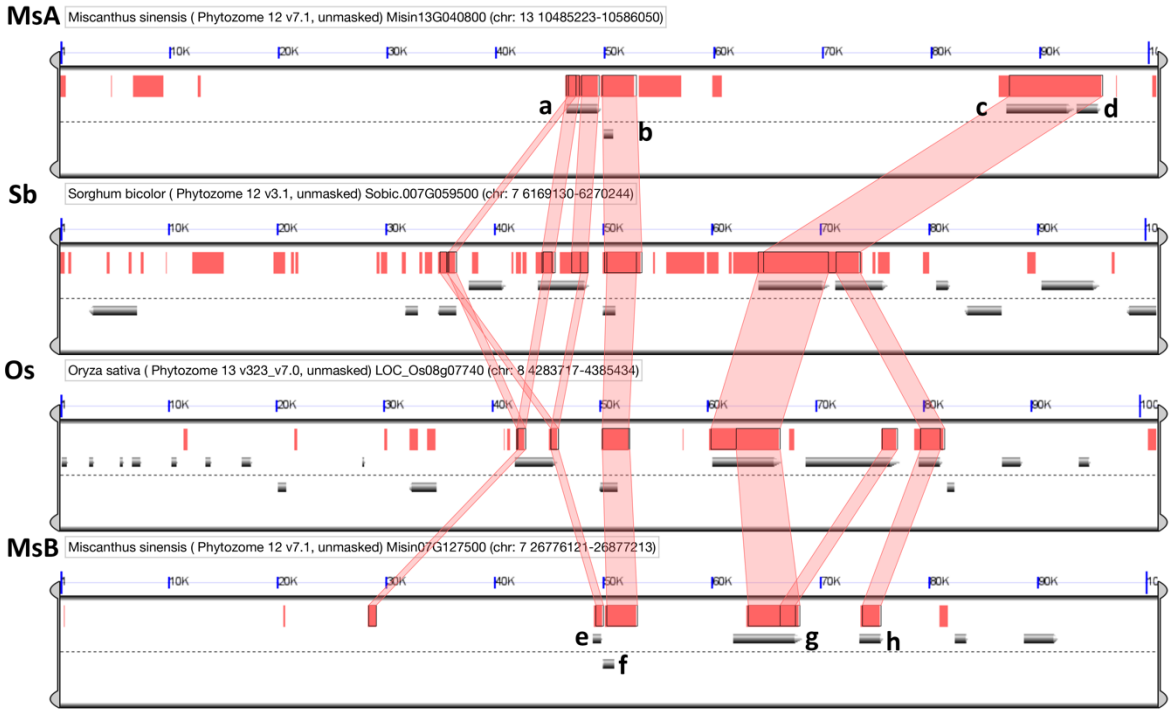


Figure S2.1 Chromosome organization of *Ghd8* gene orthologous regions (100 kbp) from *Oryza sativa*, *Sorghum bicolor* and *Miscanthus sinensis*. Only high-scoring sequence pairs (HSPs) between adjacent regions are drew in the red boxes. The dashed line in the middle of each region represents the division between the top (5′ on left) and the bottom (5′ on right) strand. The full gene models are drawn as gray arrows directly above or below this line. Colinear genes within the aligned region are connected by red lines. a-h represent colinear genes among *Oryza sativa*, *Sorghum bicolor*, *Miscanthus sinensis* (Misin13G040700 - Misin13G041000, MsA; Misin07G127400 - Misin07G127700, MsB) including *MsiGhd8A* (Misin13G040800, b) and *MsiGhd8B* (Misin07G127500, f), corresponding to Table S2.2.

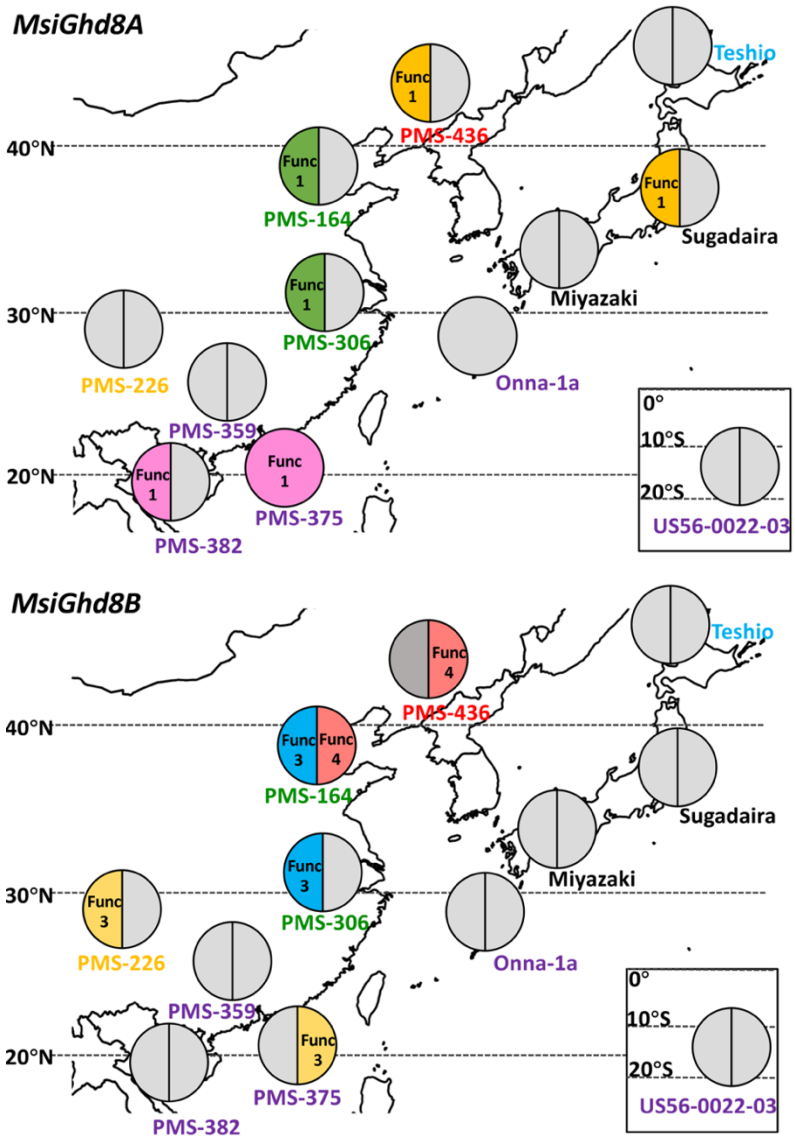


Figure S2.2 Geographical distribution of *MsiGhd8A* and *MsiGhd8B* alleles in *Miscanthus sinensis*. Pie charts with one to two sections represent the number of detected alleles. Func: putatively functional alleles, corresponding to the names in Figure 2.2. Pie charts with both the same color and Func number represent the same allele; alleles that were observed only once have a light gray background. Accessions' names are colored to represent *M. sinensis* genetic groups previously described by Clark *et al*. (2014, 2015); Sugadaira and Miyazaki are changed from yellow to black for making the map clear. The map image is taken from Wikimedia Commons: https://commons.m.wikimedia.org/wiki/File.

Table S2.1 A summary of polymorphic sites in open reading frames (ORFs) of *MsiGhd8A* alleles from 12 *Miscanthus sinensis*

Polymorphic Sites(SNVs/instertions)	SNV1	SNV2	SNV3	SNV4	SNV5	SNV6	SNV7	SNV8	SNV9	SNV10	SNV11	SNV12	SNV13	SNV14	SNV15	SNV16	SNV17	SNV18	SNV19	SNV20	SNV21	SNV22	SNV23	SNV24	SNV25	SNV26	SNV27	SNV28	SNV29	SNV30	SNV31	SNV32	Insertion	Protein Types
Position(bp)	33	48	54	80	95	96	98	105	114	145	153	221	261	279	282	359	387	408	439	489	495	516	521	544	545	568	584	586	597	633	776	792	806	
Ref(Misin13G040800)	C	G	C	A	C	G	C	C	C	G	C	C	A	G	C	A	C	G	G	C	T	G	C	G	C	A	G	G	C	A	C	G	—	
A.A Substitution				D27G	A32E		A33V			G49R		S74F				D120G			V148I	H163Q			A174V	A182P	P182Q	S190R	G195A	A196T			A259G		—/Q	
Teshio-MsiGhd8AFunc1	C	G	C	A	C	G	C	C	C	G	C	C	G	G	C	A	C	G	G	C	T	G	C	C	C	C	G	G	C	A	G	A	—	A1
Teshio-MsiGhd8AFunc2	C	G	C	A	C	G	C	C	A	G	C	C	G	G	C	G	C	G	G	C	T	G	C	C	C	C	C	G	C	A	C	A	AGC	A2
Sugadaira-MsiGhd8AFunc1	C	T	C	G	C	G	C	C	C	G	C	T	G	G	C	A	C	G	G	C	T	G	C	C	C	C	G	G	C	A	C	A	—	A3
Sugadaira-MsiGhd8AFunc2	C	G	C	A	C	G	T	C	C	G	C	C	G	G	C	A	C	G	G	C	T	G	C	G	C	C	G	G	C	A	C	A	AGC	A4
Miyazaki-MsiGhd8AFunc1	C	G	C	A	C	G	C	C	C	G	C	C	G	G	C	A	C	G	G	C	T	G	C	C	C	C	G	A	C	A	C	A	—	A5
Miyazaki-MsiGhd8AFunc2	C	G	C	G	A	G	C	C	C	G	C	C	G	G	C	A	C	G	G	C	T	G	C	C	C	C	G	G	C	A	G	A	—	A6
PMS-436-MsiGhd8AFunc1	C	T	C	G	C	G	C	C	C	G	C	T	G	G	C	A	C	G	G	C	T	G	C	C	C	C	G	G	C	A	C	A	—	A3
PMS-436-MsiGhd8AFunc2	G	G	C	G	C	G	C	C	C	G	C	C	G	G	C	A	C	G	G	C	T	G	C	C	C	C	G	G	C	A	C	A	—	A7
PMS-164-MsiGhd8AFunc1	C	G	C	G	C	G	C	C	C	G	C	C	G	G	C	A	C	G	G	C	T	G	C	C	C	C	G	G	C	A	G	A	—	A8
PMS-164-MsiGhd8AFunc2	C	G	C	G	C	G	C	C	C	G	G	C	G	G	C	A	C	G	G	C	T	G	C	C	C	C	G	G	C	A	G	A	—	A8
PMS-306-MsiGhd8AFunc1	C	G	C	G	C	G	C	C	C	G	C	C	G	G	C	A	C	G	G	C	T	G	C	C	C	C	G	G	C	A	G	A	—	A8
PMS-306-MsiGhd8AFunc2	C	G	C	G	C	G	C	A	C	G	C	C	G	G	C	A	C	G	A	C	T	G	C	C	C	C	G	G	C	A	C	G	—	A9
PMS-226-MsiGhd8AFunc1	C	G	G	A	C	T	C	C	C	G	C	C	G	G	C	A	C	G	G	C	T	G	C	C	C	C	G	G	C	A	G	A	—	A1
PMS-226-MsiGhd8AFunc2	C	G	C	G	C	G	C	G	C	G	C	C	G	G	C	A	C	G	G	C	C	A	C	C	C	C	G	G	C	T	C	A	—	A7
Onna-1a-MsiGhd8AFunc1	C	G	G	G	C	G	C	C	C	G	C	C	G	G	C	A	C	G	G	A	T	G	C	C	C	C	G	G	T	A	C	A	—	A10
PMS-359-MsiGhd8AFunc1	C	G	G	A	C	G	C	C	C	G	C	C	A	G	C	A	C	G	G	C	T	G	C	C	C	C	G	G	C	A	C	A	—	A11
PMS-359-MsiGhd8AFunc2	C	G	C	G	C	G	C	C	C	G	C	C	G	G	C	A	C	G	G	C	T	G	C	C	C	C	G	G	C	A	C	A	—	A7
PMS-375-MsiGhd8AFunc1	C	G	G	A	C	G	C	C	C	G	C	C	G	G	C	A	C	G	G	C	T	G	C	C	C	C	G	G	C	A	C	A	—	A11
PMS-382-MsiGhd8AFunc1	C	G	G	A	C	G	C	C	C	G	C	C	G	G	C	A	C	G	G	C	T	G	C	C	C	C	G	G	C	A	C	A	—	A11
PMS-382-MsiGhd8AFunc2	C	G	C	G	C	G	C	G	C	C	C	C	G	G	C	A	C	A	G	C	T	G	T	C	C	C	G	G	C	A	C	A	—	A12
US56-0022-03-MsiGhd8AFunc1	C	G	C	A	C	G	C	C	C	G	C	C	G	G	C	A	C	G	G	C	T	G	C	C	A	C	G	G	C	A	G	A	—	A13
US56-0022-03-MsiGhd8AFunc2	C	G	G	A	C	G	C	C	C	G	G	C	G	T	T	A	T	G	G	C	T	G	C	C	C	C	G	G	C	A	C	A	—	A11

Note: SNVs: single nucleotide variants. Cells with light blue represent SNVs aligned with *MsiGhd8A* (Misin13G040800) in *Miscanthus sinensis* v7.1 genome (Mitros *et al.*, 2020). SNVs located in the domain were highlighted in the red. The A prefix with numbers in the right column were the protein types. Cells with different colors in the right column represent different variant types that occurred in more than one accession; variant types that were observed only once have a gray background, corresponding to Table 2.1 and Figures 2.2 and 2.3. The sequence similarity 100%: Sugadaira -MsiGhd8AFunc1 vs. PMS-436 -MsiGhd8AFunc1, PMS-164 -MsiGhd8AFunc1 vs. PMS-306 -MsiGhd8AFunc1, PMS-375 -MsiGhd8AFunc1 vs. PMS-382 -MsiGhd8AFunc1.

Table S2.2 A summary of polymorphic sites in open reading frames (ORFs) of *MsiGhd8B* alleles from 12 *Miscanthus sinensis*

Polymorphic Sites(SNVs/Insertions)	SNV1	SNV2	SNV3	SNV4	SNV5	Insertion	SNV6	SNV7	SNV8	SNV9	SNV10	SNV11	SNV12	SNV13	SNV14	SNV15	SNV16	SNV17	SNV18	SNV19	SNV20	SNV21	SNV22	SNV23	SNV24	SNV25	SNV26	SNV27	SNV28	SNV29	SNV30	SNV31	SNV32	SNV33	SNV34	SNV35	SNV36	SNV37	SNV38	SNV39	Ptotein Types
Position(bp)	19	50	60	101	120	155	181	212	246	366	441	456	472	478	479	484	497	510	511	516	518	522	525	537	546	552	555	563	590	594	597	623	649	660	661	672	686	699	750	767	
Ref(Misin07G127500)	T	T	G	C	C	—	T	T	C	T	A	C	G	G	G	G	C	G	C	C	T	C	C	G	G	C	C	G	C	C	G	G	A	G	G	C	A	G	C	G	
A.A Substitution	Y7H	V17A	—	T34M	—	—/G	F61L	M71R	—	—	—	—	G158C	G160R	R160L	G162R	A166V	—	L171F	—	V173A	—	—	—	—	—	—	R188L	A197V	—	—	S208N	S217G	—	A221S	—	K229R	—	D250E	R256T	
Teshio-MsiGhd8BFunc3	T	T	G	C	C	—	T	T	C	C	A	C	G	C	G	G	C	G	C	T	C	C	G	G	T	C	C	G	C	C	G	G	A	G	G	C	G	G	C	G	B1
Teshio-MsiGhd8BFunc4	T	T	G	C	C	—	T	T	C	C	A	C	T	C	G	G	C	G	C	C	C	C	G	T	G	C	C	G	C	C	G	G	A	G	G	C	A	A	C	G	B2
Sugadaira-MsiGhd8BFunc3	T	T	G	C	C	—	T	T	C	C	A	C	G	G	G	G	T	G	C	C	C	C	C	G	G	C	C	G	C	C	G	G	A	G	T	C	A	G	C	G	B3
Sugadaira-MsiGhd8BFunc4	T	T	G	C	C	GCG	T	T	C	C	A	C	G	C	G	G	C	G	C	T	C	C	G	G	G	C	C	G	C	C	G	G	A	G	G	C	G	G	C	G	B4
Miyazaki-MsiGhd8BFunc3	C	T	G	C	C	—	T	T	C	C	A	C	G	C	G	G	C	G	C	C	C	C	G	G	G	C	C	G	C	C	G	G	A	G	G	C	G	A	C	G	B5
Miyazaki-MsiGhd8BFunc4	T	T	G	C	C	—	T	T	C	C	A	C	G	C	G	G	C	T	C	C	C	C	C	G	G	C	C	G	C	C	G	G	A	G	T	C	A	G	C	G	B6
PMS-436-MsiGhd8BFunc3	T	T	G	C	C	—	T	T	C	C	A	C	G	C	G	G	C	G	C	C	C	C	C	G	G	C	C	G	C	C	A	A	A	G	G	C	A	G	C	G	B7
PMS-436-MsiGhd8BFunc4	T	T	G	C	C	—	T	T	C	C	A	C	G	C	G	G	C	G	C	C	C	C	C	G	G	C	C	G	C	C	A	G	A	G	G	C	A	G	C	G	B8
PMS-164-MsiGhd8BFunc3	T	T	A	C	C	—	T	T	C	C	A	C	G	C	G	G	C	G	C	C	C	C	C	G	G	C	C	G	C	C	A	G	A	G	G	C	A	G	C	G	B8
PMS-164-MsiGhd8BFunc4	T	T	G	C	C	—	T	T	C	C	A	C	G	C	G	G	C	G	C	C	C	C	C	G	G	C	C	G	C	C	A	G	A	G	G	C	A	G	C	G	B8
PMS-306-MsiGhd8BFunc3	T	T	A	C	C	—	T	T	C	C	A	C	G	C	G	G	C	G	C	C	C	C	C	G	G	C	C	G	C	C	A	G	A	G	G	C	A	G	C	G	B8
PMS-306-MsiGhd8BFunc4	T	T	G	C	C	—	T	T	C	C	A	C	G	C	G	G	C	G	C	C	C	C	C	G	G	C	A	G	C	T	G	G	A	G	G	T	A	G	C	G	B8
PMS-226-MsiGhd8BFunc3	T	T	G	C	C	—	T	T	A	C	A	C	G	C	G	G	C	G	T	C	C	C	G	G	G	C	C	G	C	C	G	G	A	G	G	C	A	A	C	G	B9
PMS-226-MsiGhd8BFunc4	T	T	G	C	C	—	T	G	A	C	A	C	G	C	G	G	C	G	T	C	C	C	G	G	G	C	C	G	C	C	G	G	A	G	G	C	A	A	C	G	B10
Onna-1a-MsiGhd8BFunc2	T	T	G	C	C	—	T	T	C	C	A	C	G	G	G	G	C	G	C	C	C	C	C	G	G	C	C	G	C	C	G	G	A	G	T	C	A	G	C	C	B11
Onna-1a-MsiGhd8BFunc3	T	T	G	C	C	—	T	T	C	C	A	C	G	G	G	G	C	G	C	C	C	C	C	G	G	C	C	G	C	C	G	G	A	G	T	C	A	G	C	G	B12
PMS-359-MsiGhd8BFunc3	T	C	G	C	C	—	T	T	C	C	A	C	G	C	G	C	C	G	C	C	C	C	G	G	G	C	C	G	C	C	G	G	A	G	G	C	G	G	C	G	B13
PMS-359-MsiGhd8BFunc4	T	T	G	C	C	—	T	T	C	C	A	C	G	C	G	G	C	G	C	C	C	C	G	G	G	C	C	G	C	C	G	G	A	G	G	C	A	G	C	G	B8
PMS-375-MsiGhd8BFunc2	T	T	G	C	C	—	T	T	A	C	A	C	G	C	G	G	C	G	T	C	C	C	G	G	G	C	C	G	C	C	G	G	A	G	G	C	A	A	C	G	B9
PMS-375-MsiGhd8BFunc3	T	T	G	C	C	—	T	T	A	C	A	C	G	C	G	G	C	G	T	C	C	C	G	G	G	C	C	G	C	C	G	G	A	G	G	C	A	G	C	G	B9
PMS-382-MsiGhd8BFunc3	T	T	G	C	C	—	T	T	C	C	A	C	G	C	G	G	C	G	C	C	C	C	C	G	G	T	C	G	C	C	G	G	A	G	G	C	G	G	C	G	B1
PMS-382-MsiGhd8BFunc4	T	T	G	C	C	—	T	T	C	C	A	C	G	C	G	G	C	G	C	C	C	C	G	A	G	C	C	G	C	C	G	G	A	G	G	C	G	G	C	G	B1
US56-0022-03-MsiGhd8BFunc3	T	T	G	C	C	—	T	T	C	C	A	C	G	C	T	G	C	G	C	C	C	C	C	G	G	C	C	G	C	C	G	G	A	G	G	C	A	G	C	G	B14
US56-0022-03-MsiGhd8BFunc4	T	T	G	T	T	—	C	T	C	C	C	A	G	C	G	G	C	G	C	C	C	T	C	G	G	C	C	T	T	C	G	G	G	A	G	C	A	G	G	G	B15

Note: SNVs: single nucleotide variants. Cell with light blue represent SNVs aligned with MsiGhd8B (Misin07G127500) in *Miscanthus sinensis* v7.1 genome (Mitros *et al.*, 2020). SNVs located in the domain were highlighted in the red. The B prefix with numbers in the right column were the protein types. Cells with different colors in the right column represent different variant types that occurred in more than one accession; variant types that were observed only once have a gray background, corresponding to Table 2.1 and Figures 2.2 and 2.3. The sequence similarity 100%: PMS-164-MsiGhd8BFunc3 vs. PMS-306-MsiGhd8BFunc3, PMS-164-Ghd8BFunc4 vs. PMS-436-MsiGhd8BFunc4, PMS-226-MsiGhd8BFunc3 vs. PMS-375 MsiGhd8BFunc3).

Table S2.3 Protein types of MsiGhd8A from 12 *Miscanthus sinensis*

Polymorphic Sites(SNVs/insertions)	SNV4	SNV5	SNV7	SNV10	SNV12	SNV16	SNV19	SNV20	SNV23	SNV24	SNV25	SNV26	SNV27	SNV28	SNV31	Insertion
Position(bp)	80	95	98	145	221	359	439	489	521	544	545	568	584	586	776	806
Ref(Misin13G040800)	A	C	C	G	C	A	G	C	C	G	C	A	G	G	C	—
A.A Substitution / Protein Types	D27G	A32E	A33V	G49R	S74F	D120G	V148I	H163Q	A174V	A182P	P182Q	S190R	G195A	A196T	A259G	—/Q
A1	A	C	C	G	C	A	G	C	C	C	C	C	G	G	G	—
A2	A	C	C	G	C	G	G	C	C	C	C	C	C	G	C	AGC
A3	G	C	C	G	T	A	G	C	C	C	C	C	G	G	C	—
A4	A	C	T	G	C	A	G	C	C	G	C	C	G	G	C	AGC
A5	A	C	C	G	C	A	G	C	C	C	C	C	G	A	C	—
A6	G	A	C	G	C	A	G	C	C	C	C	C	G	G	G	—
A7	G	C	C	G	C	A	G	C	C	C	C	C	G	G	C	—
A8	G	C	C	G	C	A	G	C	C	C	C	C	G	G	G	—
A9	G	C	C	G	C	A	A	C	C	C	C	C	G	G	C	—
A10	G	C	C	G	C	A	G	A	C	C	C	C	G	G	C	—
A11	A	C	C	G	C	A	G	C	C	C	C	C	G	G	C	—
A12	G	C	C	C	C	A	G	C	T	C	C	C	G	G	C	—
A13	A	C	C	G	C	A	G	C	C	C	A	C	G	G	G	—

Note: SNVs: single nucleotide variants. Cells with light blue represent SNVs aligned with MsiGhd8A (Misin13G040800) in *Miscanthus sinensis* v7.1 genome (Mitros *et al.*, 2020). SNVs located in the domain were highlighted in the red.

Table S2.4 Protein types of MsiGhd8B from 12 *Miscanthus sinensis*

Polymorphic Sites(SNVs/Insertions)	SNV1	SNV2	SNV4	Insertion	SNV6	SNV7	SNV12	SNV13	SNV14	SNV15	SNV16	SNV18	SNV20	SNV27	SNV28	SNV31	SNV32	SNV34	SNV36	SNV38	SNV39
Position(bp)	19	50	101	155	181	212	472	478	479	484	497	511	518	563	590	623	649	661	686	750	767
Ref(Misin07G127500)	T	T	C	—	T	T	G	G	G	G	C	C	T	G	C	G	A	G	A	C	G
A.A Substitution / Protein Types	Y7H	V17A	T34M	—/G	F61L	M71R	G158C	G160R	R160L	G162R	A166V	L171F	V173A	R188L	A197V	S208N	S217G	A221S	K229R	D250E	R256T
B1	T	T	C	—	T	T	G	C	G	G	C	C	C	G	C	G	A	G	G	C	G
B2	T	T	C	—	T	T	T	C	G	G	C	C	C	G	C	G	A	G	A	C	G
B3	T	T	C	—	T	T	G	G	G	G	T	C	C	G	C	G	A	T	A	C	G
B4	T	T	C	GCG	T	T	G	C	G	G	C	C	C	G	C	G	A	G	G	C	G
B5	C	T	C	—	T	T	G	C	G	G	C	C	C	G	C	G	A	G	G	C	G
B6	T	T	C	—	T	T	G	C	G	G	C	C	C	G	C	G	A	T	A	C	G
B7	T	T	C	—	T	T	G	C	G	G	C	C	C	G	C	A	A	G	A	C	G
B8	T	T	C	—	T	T	G	C	G	G	C	C	C	G	C	G	A	G	A	C	G
B9	T	T	C	—	T	T	G	C	G	G	C	T	C	G	C	G	A	G	A	C	G
B10	T	T	C	—	T	G	G	C	G	G	C	T	C	G	C	G	A	G	A	C	G
B11	T	T	C	—	T	T	G	G	G	G	C	C	C	G	C	G	A	T	A	C	C
B12	T	T	C	—	T	T	G	G	G	G	C	C	C	G	C	G	A	T	A	C	G
B13	T	C	C	—	T	T	G	C	G	C	C	C	C	G	C	G	A	G	G	C	G
B14	T	T	C	—	T	T	G	C	T	G	C	C	C	G	C	G	A	G	A	C	G
B15	T	T	T	—	C	T	G	C	G	G	C	C	C	T	T	G	G	G	A	G	G

Note: SNVs: single nucleotide variants. Cells with light blue represent SNVs aligned with MsiGhd8B (Misin07G127500) in *Miscanthus sinensis* v7.1 genome (Mitros *et al.*, 2020). SNVs located in the domain were highlighted in the red.

Table S2.5 Colinear genes near homologs of *Ghd8* in *Miscanthus sinensis*, *Sorghum bicolor* and *Oryza sativa*

Colinear genes	Locus ID in *Miscanthus sinensis**	Locus ID in *Sorghum bicolor*	Locus ID in *Oryza sativa*	Putative function
a, e	Misin13G040700 Misin07G127400	Sobic.007G059400	LOC_Os08g07730	Transferase family protein
b, f	Misin13G040800 Misin07G127500	Sobic.007G059500	LOC_Os08g07740	CCAAT-box-binding transcription factor
c, g	Misin13G040900 Misin07G127600	Sobic.007G059600	LOC_Os08g07760	BRASSINOSTEROID INSENSITIVE 1-associated receptor kinase 1 precursor
d, h	Misin13G041000 Misin07G127700	Sobic.007G059700	LOC_Os08g07790	CRS2-associated factor 2, mitochondrial precursor

* Gene Locus IDs in *Miscanthus sinensis*, *Sorghum bicolor* and *Oryza sativa* are from Phytozome v13 (https://phytozome-next.jgi.doe.gov)

Chapter 3

Characterization of the *Ghd7* flowering time gene

3.1 Introduction

Flowering time is crucial for reproduction and geographic expansion of plants. And flowering regulation is a complex process involving many flowering QTLs/genes. Of which, *Ghd7* and *Ghd8* are two major flowering genes identified in rice (Xue *et al.*, 2008; Wei *et al.*, 2010; Yan *et al.*, 2011). Recently, the combinations and interaction between *Ghd7* and *Ghd8* were identified that largely affect rice flowering and then determine the ecogeographical adaptation and yield potential of cultivated rice (Zhang *et al.*, 2015, 2019; Wang *et al.*, 2019*a*; Zong *et al.*, 2021). *Ghd7* is thought to be an evolutionarily new gene in the grass family, because it was absent from *A. thaliana*, and the protein sequence encodes a CCT domain (Griffiths *et al.*, 2003; Xue *et al.*, 2008; Lu *et al.*, 2012). In rice, by the comparison of the predicted protein sequences, five alleles were identified in rice and showed a clear geographic distribution (Xue *et al.*, 2008). The functional alleles with strong effects allow rice plants to delay flowering under LD condition in areas with long growing seasons, thus producing large panicles and increasing yield (Xue *et al.*, 2008). The reduced function of *Ghd7* is important for extremely early flowering time with adaptation of rice to regions with low temperatures and short growth seasons (Xue *et al.*, 2008; Gómez-Ariza *et al.*, 2015; Li *et al.*, 2015; Fujino *et al.*, 2019; Fujino and Yamanouchi, 2020). Comparative sequence analysis revealed high gene collinearity among rice, sorghum, switchgrass and *M. sinensis* genomes (Salse *et al.*, 2009; Kim *et al.*, 2012; Lu *et al.*, 2012; Swaminathan *et al.*, 2012;

Dong *et al.*, 2018; Mitros *et al.*, 2020; Jensen *et al.*, 2021). *Ghd7* was also identified in sorghum and alleles of *SbGhd7* affected photoperiod sensitivity and flowering times (Murphy *et al.*, 2014).

To date, there have been no report of *Ghd7* in *Miscanthus*. Thus, the purpose of this chapter is to investigate whether the grass-specific gene, *Ghd7*, exist in *M. sinensis* or not? Does it have a functional *Ghd7* that contributes to the regulation of flowering time and was there the difference in gene expression patterns of *Ghd7* regulates flowering in *M. sinensis* under LD and SD? In this study, the ortholog of *SbGhd7* in a mini-core collection of *M. sinensis* was cloned with the objective to 1) characterize allelic and deduced amino acid sequence diversity and geographic distribution, and 2) determine expression patterns of *Ghd7* in response to photoperiod.

3.2 Materials and methods

3.2.1 Plant materials and growth conditions

The same twelve *M. sinensis* accessions were planted and applied for isolation of *Ghd7* and gene expression analysis, and the details were described in Section 2.2.1 in Chapter 2.

3.2.2 Genomic DNA extraction and cloning of *Ghd7* in *M. sinensis*

Genomic DNA was extracted from young, healthy leaves of each line using the DNeasy Plant Mini Kit (Qiagen, Tokyo, Japan), and qualified in 0.8 % agarose gels by electrophoresis and quantified by the NanoDrop1000 instrument. *Miscanthus sinensis* v7.1 genome (Mitros *et al.*, 2020) and its close relative sorghum using the

Sorghum bicolor v3.1 genome (McCormick *et al.*, 2018) from Phytozome v13 (https://phytozome-next.jgi.doe.gov) were queried using a BLAST search and the candidate nucleotide region detected with about 2kb on the located chromosome was used for primer design. To obtain the full length of sequence and intron/exon junction in *M. sinensis* accessions, *Ghd7* sequence was amplified from both cDNA and genomic DNA and subjected to capillary sequencing. As a repetitive sequence inserted into the intron, resulting in extreme difficulty in cloning highly repetitive DNA into plasmid vectors (Godiska *et al.*, 2009), instability of the structure of DNA (Holder *et al.*, 2015). Therefore, genomic *Ghd7* sequence in *M. sinensis* were divided into three fragments for amplification and alignment. PCRs was applied in amplification of *Ghd7* and it contained 30 ng of total genomic DNA as a template and LA Taq polymerase (TaKaRa Bio, Shiga, Japan). Cycling conditions were as follows: 1 min at 95 °C, 30 cycles of 30 sec at 95 °C, 30 sec at suitable primer temperature and at 72 °C for suitable amplified length. PCR products were separated on 0.8 % agarose gels by electrophoresis. Targeted fragments were eluted from the agarose gel with the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nager GmbH & Co. KG, Düren, Germany) and cloned into the pGEM-T Easy vector (Promega, Madison, WI, USA) using the TA-Blunt Ligation Kit (Nippon Gene Co., Ltd., Toyama, Japan) following the manufacturer's instructions. Subsequently, the ligation products were transformed into Escherichia coli JM109 competent cells for white-blue screening. Bacteria were grown in Luria-Bertani (LB) broth or on LB agar supplemented with Ampicillin at 100 mg/ml, Kanamycin at 50 mg/ml, chloramphenicol at 25 mg/ml for selecting the positively transformed colonies, and plasmids were purified using High Pure Plasmid Isolation Kit (Roche,

Sigma-Aldrich, Tokyo, Japan). About 20 plasmid clones of each genotype were sequenced in both directions with a BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) via an ABI PRISM 3130 Genetic Analyzer (Life Technologies, Carlsbad, USA) according to the manufacturer's instructions. To identify true alleles, quality assessment was conducted and the only sequence represented at least three times that also aligned to the physical map were retained.

3.2.3 RNA extraction and diurnal gene expression analysis

The protocol for RNA extraction and gene expression analysis were described in Section 2.2.3 in Chapter 2. The primer sets used for *MsiGhd7* expression in two subgenomes were 5'-TCAAAGAGACAACCCTGACCGACGA-3' (forward primer) and 5'-GGTTACCTTAGCAAAGCGGCCTC-3' (reverse primer) for A subgenome, and 5'- TCAAGGAGCCAACCCTGACCGATGG-3' (forward primer) and 5'-TCGGTTACCTTGGCAAAGCGGCCTT-3' (reverse primer) for B subgenome. The absence of genomic DNA contamination was verified by running no-template control qRT-PCR reactions for each sample.

3.2.4 Data analysis

The nucleotide sequences were assembled with ATGC v.6 software (GENETYX Co., Tokyo, Japan). The data analysis was conducted similarly to Section 2.2.4 in Chapter 2.

3.3 Results and discussion

3.3.1 Features of *Ghd7* in *M. sinensis* accessions

A total of 56 alleles were identified from the 12 wild-collected *M. sinensis* accessions. Three to six different alleles were obtained from each *M. sinensis* accession, indicating that *MsiGhd7* consisted of at least three loci in *M. sinensis* genome. Blasting to *Miscanthus sinensis* v7.1 genome (Mitros *et al.*, 2020), with paleo-allotetraploid species' subgenomes, one locus was located on Chromosome 11 (A subgenome, Chr.11:1998112..2000968), regarded as *MsiGhd7A*, and other loci were located on Chromosome 12 (Chr.12, B subgenome) (Figure 3.1). In the B subgenome, the alleles showed high similarity to two segments on Chr.12, with one located ranged from 1952462 to 1953905 and another at the region from 1803910 to 1804245, as only the partial sequence were available, all *MsiGhd7* alleles located at the B subgenome were named as *MsiGhd7B*. Neighbor-Joining (NJ) phylogenetic trees revealed a robust separation of clades representing *MsiGhd7A* (22 alleles) and *MsiGhd7B* (34 alleles) (Figure 3.2). Therefore, *MsiGhd7* identified by this study in *M. sinensis* might have been caused by the whole genome duplication (*MsiGhd7A* and *MsiGhd7B*) and also by local gene duplications (multiple *MsiGhd7B* loci).

Based on the results of RT-PCR, both *MsiGhd7* homoeologs consisting of two exons and one intron in the PCR fragment was identified (Figure 3.1), and the splicing sites of *MsiGhd7* matched to those of sorghum and rice (Xue *et al.*, 2008; Murphy *et al.*, 2014). The length of exons and intron in *MsiGhd7A* ranged from 2270 to 2398 nucleotides, and *MsiGhd7B* ranged from 2235 to 2340 nucleotides,

which contained a different length of repetitive region varying from 14 base pair (bp) to 172 bp (Figure 3.1). Multiple sequence blasting revealed that the nucleotide sequences of *M. sinensis Ghd7* showed a high similarity to those in other plant species, especially for sorghum (Sobic.006G004400.2, 89.5% - 91.2%), maize (Zm00001d024909 _T001, 76.2%-77.3%), *Panicum hallii* (Pahal.7G008400.1, 75.6%-76.6%), rice (LOC_Os07g15770.1, 62.0%-64.1%). A microsynteny assessment of genomic regions adjacent to *Ghd7* in rice, sorghum and *M. sinensis* indicated that *MsiGhd7* is orthologous to *OsGhd7* (Xue *et al.*, 2008) and *SbGhd7* (Murphy *et al.*, 2014) (Figure S3.1).

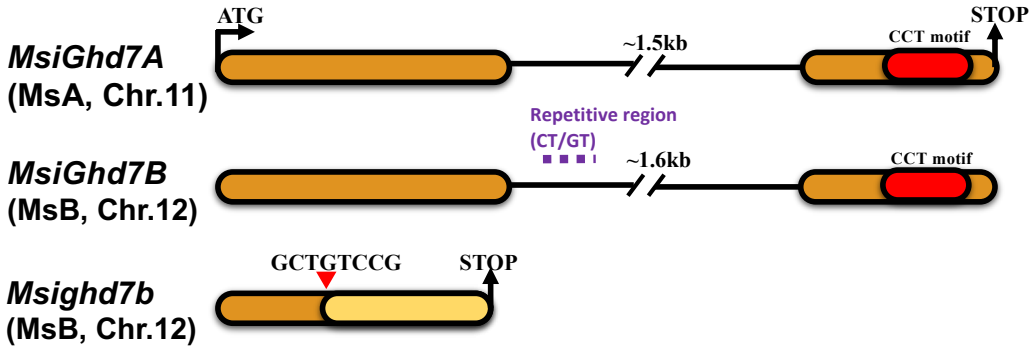


Figure 3.1 The structure of *Miscanthus sinensis Ghd7* homoeologous loci, *MsiGhd7A* and *MsiGhd7B*. *Msighd7b* allele derived from PMS-164 contains an eight-base insertion in the Exon 1, causing a frameshift and a premature stop. Exon are showed as boxes, and intron as solid lines. Red boxes indicate CCT domain; pale yellow indicates extensive missense.

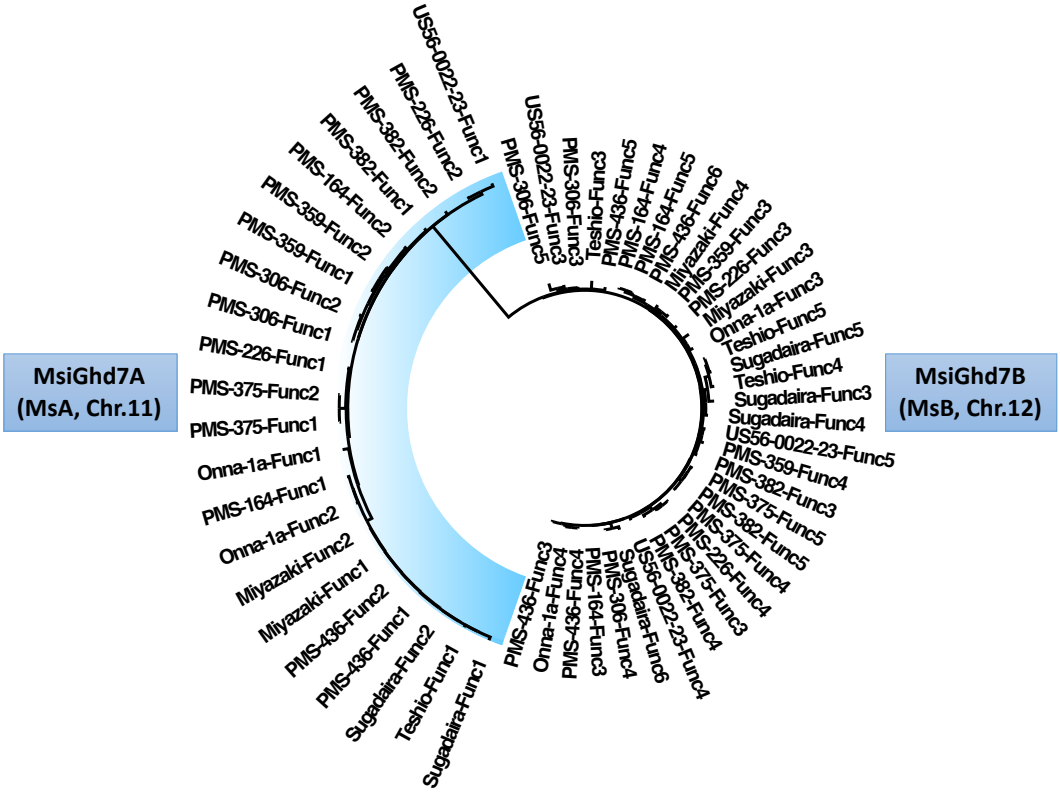


Figure 3.2 Phylogenetic tree inferred by neighbor-joining method for nucleotide sequences of 52 *Ghd7* alleles from 11 accessions of *Miscanthus sinensis* and four alleles from one *Miscanthus floridulus* accession. The phylogenetic tree was divided into two clusters, which were classified as *MsiGhd7A* and *MsiGhd7B*, one for each of the two subgenomes of *M. sinensis*.

Of 56 *MsiGhd7* alleles derived from the 12 wild-collected *M. sinensis* accessions in this study, the ORFs lengths of *MsiGhd7A* ranged from 732 to 735 nucleotides, encoding for 243 to 244 amino acid residence, while ORFs of *MsiGhd7B*, consists of 735 to 738 bp translating to from 244 to 245 amino acid product. The amino acid sequence similarity of putatively functional MsiGhd7A and MsiGhd7B variants ranged from 92.1% to 94.2%. In addition, similar to rice and sorghum (Xue *et al.*, 2008; Murphy *et al.*, 2014), MsiGhd7 does not have an obvious B-box zinc finger structure and were identified 43 contiguous amino acids having significant identity with the CCT domain of Ghd7 protein in the other species (Figure 3.1 and 3.3). Phylogenetic tree generated based on the amino acid sequence among the closest orthologs of Ghd7 in seven plant species (Figure 3.4) is in consistent with the evolution of *Ghd7* in the grasses (Ream *et al.*, 2012; Yang *et al.*, 2012; Woods *et al.*, 2016). CCT-domain proteins have been reported to have crucial roles in regulating processes such as photoperiodic flowering (Putterill *et al.*, 1995; Yano *et al.*, 2000; Turner *et al.*, 2005; Hung *et al.*, 2012; Murphy *et al.*, 2014; Yang *et al.*, 2014*b*), vernalization (Yan *et al.*, 2004), circadian rhythms (Strayer *et al.*, 2000; Salomé *et al.*, 2006), and light signaling (Kaczorowski and Quail, 2003; Murphy *et al.*, 2011). Only one putative loss-of-function allele identified in *MsiGhd7B*, *Msighd7b*, was found in PMS-164 from Northern China and was characterized by an eight-base insertion in the first exon, upstream of the CCT domain (Figure 3.1). This mutation resulted in a frameshift and eventual premature termination of the protein, lacking CCT domain (Figure 3.1).

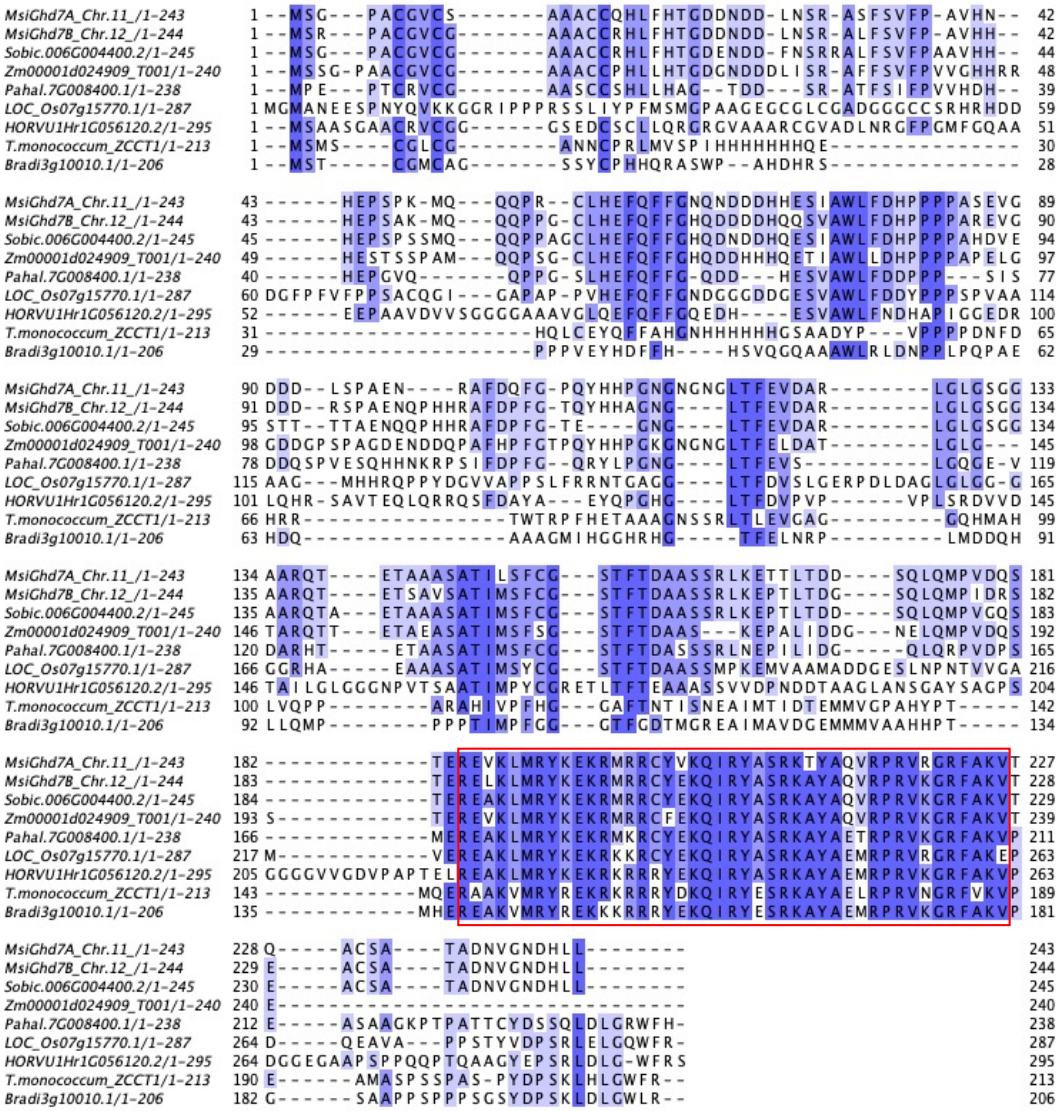


Figure 3.3 Alignment of MsiGhd7 with its closest orthologs in seven plant species. *Sorghum bicolor* (Sobic.006G004400.2, SbGhd7), *Zea mays* (Zm00001d024909_T001, ZmCCT), *Panicum hallii* (Pahal.7G008400.1, PhGhd7), *Oryza sativa* (LOC_Os07g15770.1, OsGhd7), *Hordeum vulgare* (HORVU1Hr1G056120.2, HvVRN2), *Triticum monococcum* (T. monococcum_ZCCT1) and *Brachypodium distachyon* (Bradi3g10010.1, CCT). The Ghd7 amino acid sequence in *Triticum monococcum* was obtained from UniProt (https://www.uniprot.org). The *M. sinensis* Ghd7 amino acid sequence used for alignment was derived from Sugadaira. Amino acid residues were colored

indicate residues of strongly conserved properties, while residues uncolorful by a period indicate residues with more weakly similar properties. The CCT domain region is highlighted in red.

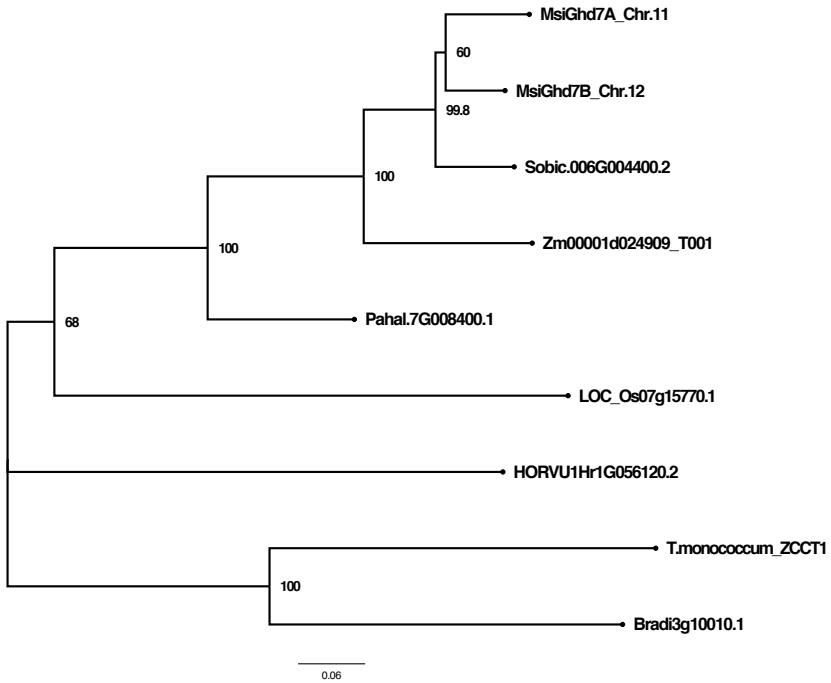


Figure 3.4 Neighbor-joining tree generated based on the amino acid in the closest orthologs of MsiGhd7 among several plant species. Bootstrap values for nodes supported in > 50% of 1000 bootstrap replicates are shown. *Sorghum bicolor* (Sobic.006G004400.2, SbGhd7), *Zea mays* (Zm00001d024909_T001, ZmCCT), *Panicum hallii* (Pahal.7G008400.1, PhGhd7), *Oryza sativa* (LOC_Os07g15770.1, OsGhd7), *Hordeum vulgare* (HORVU1Hr1G056120.2, HvVRN2) and *Triticum monococcum* (T. monococcum_ZCCT1) and *Brachypodium distachyon* (Bradi3g-10010.1, CCT). The Ghd7 amino acid sequence in *T. monococcum* was obtained from UniProt (https://www.uniprot.org). The *Miscanthus sinensis* Ghd7 amino acid sequence used was derived from Sugadaira.

3.3.2 Geographical distribution of Ghd7 protein variants in *M. sinensis*

Compared to ORFs of alleles in the *Miscanthus sinensis* v7.1 genome (Mitros *et al.*, 2020), 41 non-synonymous single nucleotide variants (nsSNVs), 20 synonymous single nucleotide variants (sSNVs) and two 3-bp insertions and one 8-bp insertion in ORFs, with some accessions having more than one SNV per allele were observed (Tables S3.2). Considering the fact that the nucleotide diversity cannot precisely represent the protein diversity owing to synonymous SNVs in ORFs, Ghd7 protein variant types were analyzed in the present study (Figure 3.5 and Table S3.2). Accounting for nsSNVs, 11 predicted amino acid sequence types of MsiGhd7A and 21 of MsiGhd7B (33 total) were identified from the 12 *M. sinensis* accessions (Figure 3.5 and Table S3.2).

In the A subgenome, variant ‘a1’ had a restricted distribution in Japan with a latitude ranging from 31.8 °N and 44.9 °N, and in northeastern China but was absent from mid and low latitudes in mainland Asia, was the most frequent variant (33.3% of accessions), indicated that potentially their contribution to the adaptability of a short growing season. Six variants (a3, a4, a5, a6, a9 and a11) with 3-bp nucleotide insertion in the Exon1, encoding for more one amino acid than other types, were absent from Japan. Variants ‘a3’ was widely distributed in accessions from south to north in China, with a latitude ranging from 20.9 °S to 37.3 °N and frequently observed (25 % of accessions). Variants ‘a4’, differing from ‘a3’ by one amino acid, was found in accessions from Central and Southern China, with a latitude < 30 °N, and frequently observed (25 % of accessions).

By contrast, as at least two loci existed in the B subgenome, harboring much more variants types, some of the MsiGhd7 protein variants were found over a broad geographic range, whereas others had restricted patterns of occurrence. Variant ‘b2’ had a limited distribution in Northern Japan. ‘b12’ was observed in low latitude in Japan as well as from Hainan to Northeastern China, and was the most frequently (25% of accessions). In mainland Asia, ‘b9’ and ‘b11’was also broadly distributed from low to high latitude but infrequent (16.6% of accessions). ‘b15’ was found in central and southeastern China but also infrequent (16.6% of accessions). ‘b20’ had a restricted distribution in Hainan island China. The *M. floridulus* accession (US56-0022-03) from New Caledonia contains three types, of which ‘b14’ and ‘b19’ were found in PMS-306 and PMS-359 from low latitude in China, respectively. The other thirteen variants were each observed in only one accession. The only non-functional variant type ‘b13’ occurred, resulting from a premature termination in the predicted coding region, was found in one Chinese accession PMS-164 from 37.3 °N. Though this mutation is rare, it would be worth to elucidate its distribution under population groups in future study.

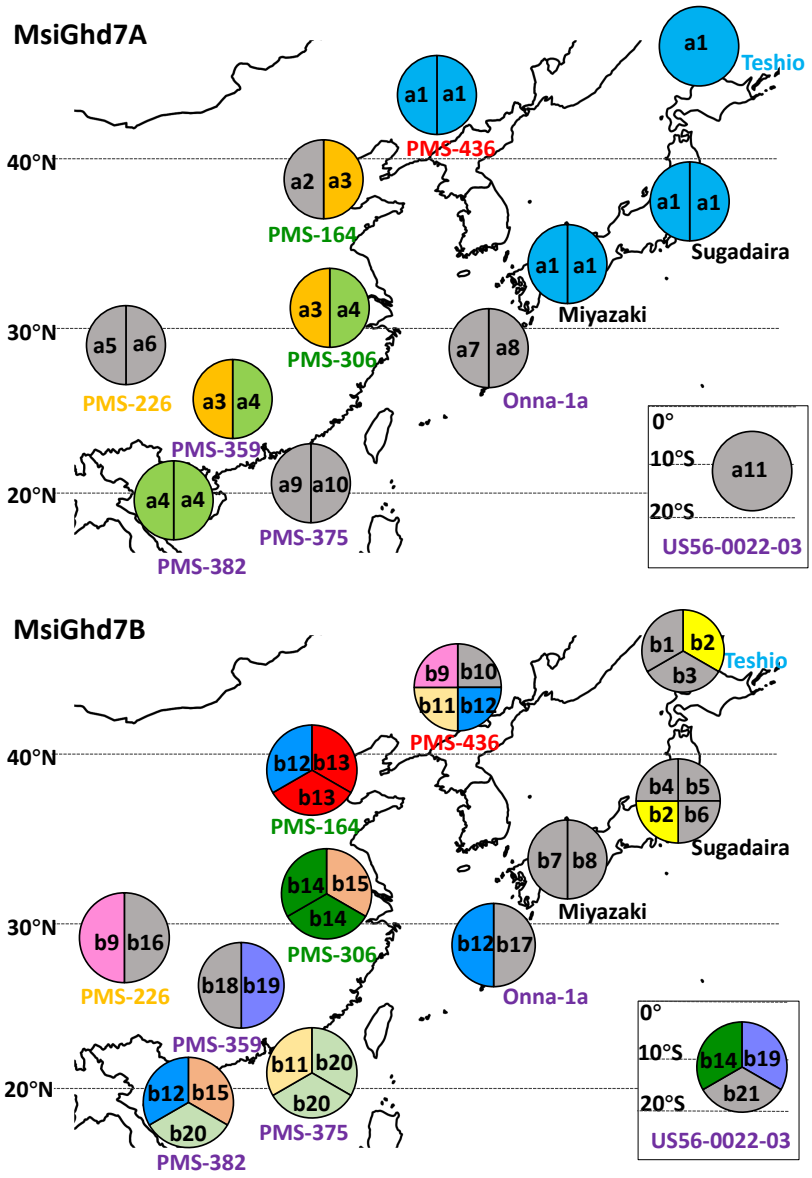


Figure 3.5 Geographical distribution of MsiGhd7A and MsiGhd7B predicted amino acid sequence variant types in *Miscanthus sinensis*. Pie charts with one to two sections represent the number of detected alleles. 'a' or 'b' prefix indicates putatively functional alleles types based on predicted amino acid sequence variants, corresponding to the names in Table S3.2. Different colors in pie charts represent different variant types that occurred in more than one accession; variant types that were observed only once have a gray background, corresponding to Table S3.2. Accessions' names are colored to represent *M. sinensis* genetic groups previously

described by Clark *et al*. (2014, 2015); Sugadaira and Miyazaki are changed from yellow to black for making the map clear. The map image is taken from Wikimedia Commons: https://commons.m.wikimedia.org/wiki/File.

3.3.3 Response of *Ghd7* homoeologous genes to photoperiod

The expression of *Ghd7* in rice and sorghum have been shown to be responsive to light signals and circadian clock (Xue *et al.*, 2008; Murphy *et al.*, 2014). In order to gain the insight into the molecular mechanism of *MsiGhd7* in response to photoperiod in *M. sinensis*, its expression was measured by qRT-PCR in young leaves of twelve *M. sinensis* accessions under LD and SD. Two *MsiGhd7* homoeologs expressed in all accessions and expression level varied in individuals. Notably, mRNA transcript levels of both *MsiGhd7* were clearly modulated by photoperiod (Figure 3.6). For most accessions, the expression of both *MsiGhd7* genes was much higher in LD condition than SD, and *MsiGhd7* mRNA transcript was more abundant during the light period than in the dark period, similar to rice (Xue *et al.*, 2008), sorghum (Murphy *et al.*, 2014) and *P. hallii* (Weng *et al.*, 2019). In rice, *OsGhd7* showed a single peak of clock-gated expression in the morning of LD (Xue *et al.*, 2008; Itoh *et al.*, 2010). However, in *M. sinensis*, the expression of *Ghd7* showed a frequent peak at the early morning (ZT3) for most accessions, and the *Ghd7* mRNA abundance for some accessions had two peaks at 3h and 15h after initial light exposure in the morning (Figure 3.6), similar to *SbGhd7* (Murphy *et al.*, 2014) and *P. hallii* (Weng *et al.*, 2019). However, in the sugarcane (Glassop and Rae, 2019), a perennial plant, which displayed an obvious peak and a smaller peak for sugarcane *Ghd7*, but not significantly. Dual peaks of *VRN2* (the ortholog of *Ghd7*) also observed in wheat (Shaw *et al.*, 2020), the second peak was also in the light and dark transition, whereas, the first peak (at ZT8) was later than that in *M. sinensis*, sorghum and *P. hallii*.

By contrast, in SD, these peaks in expression were almost diminished for most accessions, and the new obvious peak at ZT9 in *MsiGhd7B* of PMS-306 and PMS-359, which flowered both LD and SD conditions. In *M. sinensis*, sorghum, *P. hallii*, rice and wheat (data are not available in sugarcane), there is no second peak when the plants are grown under photoperiod inductive conditions. These conserved expression patterns suggested that the *Ghd7* regulation pathway also existed in *M. sinensis*, and furthermore, this dual peak of *Ghd7* tends to potentially modify the flowering regulation model versus rice. Obviously, the underlying mechanisms should be further investigated.

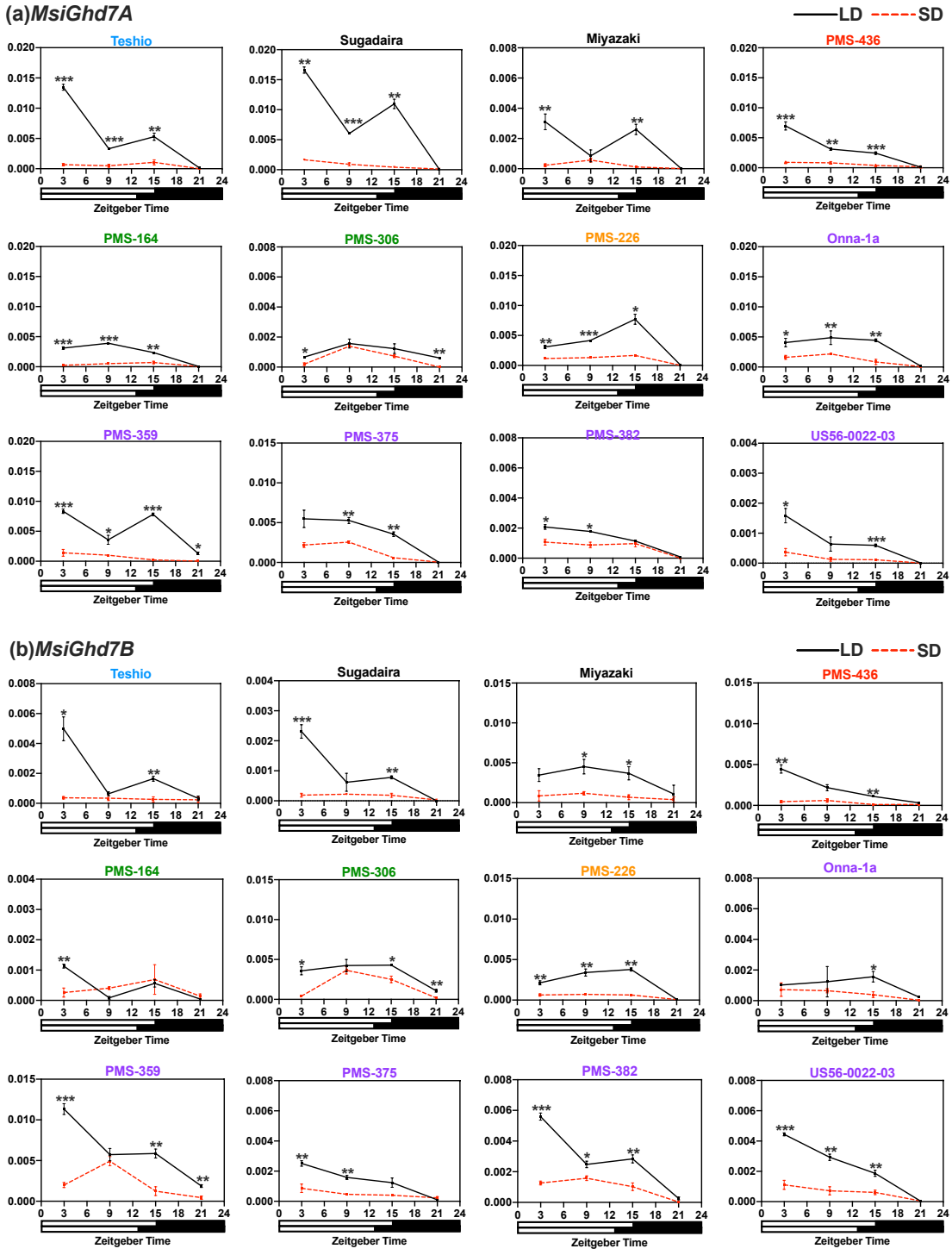


Figure 3.6 Diurnal expression of *MsiGhd7* in 12 *Miscanthus sinensis* genotypes under long days (15 h, LD; solid black lines) and short days (12.5 h, SD; red dashed line). (a) *MsiGhd7A* and (b) *MsiGhd7B*. The *y*-axis represents relative mRNA levels normalized to *ACTIN* transcript levels. The numbers below the *x*-axis indicate

Zeitgeber times (ZT) of the day. The white bar at the bottom of each graph indicates the light period and the black bar indicates the dark period. Mean ± 1SE for three replications are given for each data point. Asterisks indicate significant difference between the two means under LD and SD at the same ZT of the day (Student's *t*-test, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$). No asterisk indicates the difference between the two means is not statistically significant ($P > 0.05$). Accessions' names are colored to represent *M. sinensis* genetic groups previously described by Clark *et al*. (2014, 2015); Sugadaira and Miyazaki are changed from yellow to black for making it clear.

Supplementary materials for Chapter 3:

Supplementary Table S3.1: Primer sequences used for *Ghd7* homoeologous loci amplification and sequencing in *Miscanthus sinensis.*

Supplementary Table S3.2: A summary of polymorphic sites in MsiGhd7A protein from 12 *Miscanthus sinensis*.

Supplementary Table S3.3: A summary of polymorphic sites in MsiGhd7B protein from 12 *Miscanthus sinensis.*

Supplementary Figure S3.1: Chromosome organization of near *Ghd7* gene regions (100 kbp) from *Oryza sativa*, *Sorghum bicolor* and *Miscanthus sinensis*.

Table S3.1 Primer sequences used for *Ghd7* homoeologous loci amplification and sequencing in *Miscanthus sinensis*

Segment	Amplification Primer Sequence (5' → 3')	Primer sequences for sequencing (5' → 3')
Segment 1 Amplification (Across Exon1) [F1R1(573bp); F2R2(577bp)]	F1: CACACAGCAGACCTCTCAAGATC R1: GAACGTGAAGCAACTGAAGGCA F2: CACACCTCCACACAGCAG R2: GGTCTAGTCTAGAATAAAAGGCAG	>For all segments CTCACTATAGGGCGAATTGG GGTGCACTATAGAATACTCTAAGC
Segment 2 Amplification (Across Exon2) [F3R3(492bp); F4R3(967bp); F5R4(615bp); F5R5(615bp); F6R6(615bp)]	F3: CACAACCTCTTTCTGGCCCATT F4: TTCCATGCATATGTCCAACTGG F5: CGCTACCAAGGCAATATCTGATGTT F6: GAGGACTGAGAGATATTAACAAG R3: TATAGCAGTCGGCAATTGAGAGAC R4: GTCAAATTAACCAGTGCCCAATATC R5: CAACTTTTGGGGTTACTGTTTG R6: GTAGTCAAATTAACCAGTGCC	>For all segments CTCACTATAGGGCGAATTGG GGTGCACTATAGAATACTCTAAGC >F6R6 segment GCAGGCACTCCAGATTTGTATGGT
Segment 3 Amplification (Across Intron) [F7R7(2068bp); F8R8(1568bp)]	F7: TGCACGAGTTCCAGTTCTTTGGC F8: CTCACCTTTGAGGTGGACGCCAGGCT R7: CAACGTTGTCTGCTGTGGCGGAGC R8: AATGGGCCAGAAAGAGGTTGTG	>For all segments CTCACTATAGGGCGAATTGG GGTGCACTATAGAATACTCTAAGC >F7R7 segment AAGTACTCCTAGTAGGGAGCATGG TGCCTTCAGTTGCTTCACGTTC GCAGGCACTCCAGATTTGTATGGT CCAGTTGGACATATGCATGGAA CTTGTTAATATCTCTCAGTCCTCG >F8R8 segment GAGGACTGAGAGATATTAACAAG GGATGGTCAGGGCTCATGTT ACCATACAAATCTGGAGTGCCTGC

F: forward primer; R: reverse primer

Table S3.2 A summary of polymorphic sites in MsiGhd7A protein from 12 *Miscanthus sinensis*

Polymorphic Sites(SNVs/instertions)	SNV1	SNV2	SNV3	SNV4	SNV5	SNV6	insertion 1	SNV7	SNV8	SNV9	SNV10	SNV11	SNV12	SNV13	SNV14	SNV15	SNV16	SNV17	SNV18	SNV19	SNV20	SNV21	SNV22	Protein Types
Position(bp)	33	60	66	94	124	157	161	167	259	261	415	420	422	1964	2025	2029	2057	2063	2074	2164	2174	2204	2208	
Ref(Misin11G0003800.2)	G	T	G	G	A	C	—	G	G	G	G	C	C	C	T	C	G	T	G	G	C	C	C	
A.A Substitution	—	—	—	A32S	N42H	P53A	R54P*; —/G	C55S	E87A	E87D	E139K	—	A139V	L148M	L168Q	—	D179N	S181P	—	Q214H	—	Q228E	A229G	
Teshio-MsiGhd7AFunc1	G	T	G	G	A	C	—	G	G	G	G	C	C	C	T	C	G	T	G	G	C	C	C	a1
Sugadaira-MsiGhd7AFunc1	G	T	G	G	A	C	—	G	G	G	G	C	C	C	T	C	G	T	G	G	C	C	C	a1
Sugadaira-MsiGhd7AFunc2	G	T	G	G	A	C	—	G	G	G	G	C	T	C	T	C	G	T	G	G	C	C	C	a1
Miyazaki-MsiGhd7AFunc1	G	T	G	G	A	C	—	G	G	G	G	C	C	C	T	C	G	T	G	G	C	C	C	a1
Miyazaki-MsiGhd7AFunc2	G	T	G	G	A	C	—	G	G	G	G	C	C	C	T	C	G	T	G	G	C	C	C	a1
PMS-436-MsiGhd7AFunc1	G	T	G	G	A	C	—	G	G	G	G	C	C	C	T	C	G	T	G	G	C	C	C	a1
PMS-436-MsiGhd7AFunc2	G	C	G	G	A	C	—	G	G	G	G	C	C	C	T	C	G	T	G	G	C	C	C	a1
PMS-164-MsiGhd7AFunc1	G	T	G	G	A	C	—	G	G	G	G	C	C	A	T	C	G	T	G	G	C	C	C	a2
PMS-164-MsiGhd7AFunc2	G	T	T	G	A	G	CGG	G	G	G	G	C	C	A	T	C	G	T	G	G	C	G	C	a3
PMS-306-MsiGhd7AFunc1	G	T	T	G	A	C	CGG	G	G	G	G	C	C	A	T	C	G	T	G	G	C	G	C	a4
PMS-306-MsiGhd7AFunc2	A	T	T	G	A	G	CGG	G	G	G	G	C	C	A	T	C	G	T	G	G	C	G	C	a3
PMS-226-MsiGhd7AFunc1	G	T	C	G	C	C	CGG	G	G	C	G	C	C	A	T	C	G	C	G	G	C	G	C	a5
PMS-226-MsiGhd7AFunc2	G	T	C	T	A	C	CGG	C	G	G	G	C	C	A	T	C	G	T	G	G	A	G	C	a6
Onna-1a-MsiGhd7AFunc1	G	T	G	G	A	C	—	C	G	G	G	C	C	A	T	C	G	C	G	C	C	C	G	a7
Onna-1a-MsiGhd7AFunc2	G	T	G	G	A	C	—	C	G	G	G	T	C	A	T	C	G	T	G	G	C	G	C	a8
PMS-359-MsiGhd7AFunc1	G	T	T	G	A	G	CGG	C	G	G	G	C	C	A	A	C	G	T	G	G	C	G	C	a3
PMS-359-MsiGhd7AFunc2	G	T	C	G	A	C	CGG	C	A	G	G	C	C	A	A	C	G	T	G	G	C	G	C	a4
PMS-375-MsiGhd7AFunc1	G	T	C	G	A	C	CGG	C	G	G	G	C	C	A	T	T	G	C	A	G	C	G	C	a4
PMS-375-MsiGhd7AFunc2	G	T	C	G	A	C	—	C	G	G	G	C	C	A	T	T	G	C	A	G	C	G	C	a4
PMS-382-MsiGhd7AFunc1	G	T	C	G	A	C	CGG	C	G	G	G	C	C	A	T	C	G	T	G	G	C	G	C	a9
PMS-382-MsiGhd7Func2	G	T	C	G	A	C	CGG	C	G	G	G	C	C	A	T	C	A	T	G	G	C	G	C	a10
US56-0022-03-MsiGhd7Func1	G	T	C	T	C	C	CGG	C	G	C	A	C	C	A	T	C	C	T	G	G	C	G	C	a11

Note: SNVs: single nucleotide variants. Cells with light blue represent SNVs aligned with *MsiGhd7A* (Misin11G003800.2) in *Miscanthus sinensis* v7.1 genome (Mitros *et al.*, 2020). SNVs located in the domain were highlighted in the red. The 'a' prefix with numbers in the right column were the protein types. Cells with different colors in the right column represent different variant types that occurred in more than one accession; variant types that were observed only once have a gray background, corresponding to Figure 3.5.

Table S3.3 A summary of polymorphic sites in MsiGhd7B protein from 12 *Miscanthus sinensis*

Polymorphic Sites(SNVs/insertions)	SNV1	SNV2	SNV3	SNV4	SNV5	SNV6	SNV7	SNV8	SNV9	SNV10	Insertion 1	SNV12	SNV13	SNV14	SNV15	SNV16	SNV17	SNV18	SNV19	SNV20	SNV21	SNV22	SNV23	SNV24	SNV25	SNV26	SNV27	SNV28	SNV29	SNV30	SNV31	SNV33	insertion 2	SNV35	SNV36	SNV37	SNV38	SNV39	SNV40	SNV41	SNV42	SNV43	Protein Types
Position(bp)	7	38	54	57	63	78	90	132	139	163	172	199	205	217	234	245	254	267	268	273	280	281	330	346	390	394	406	419	433	444	461	475	481	530	542	559	589	612	635	691	713	721	
Ref(Chr12 1726bp)	A	C	C	C	T	C	C	G	G	G	–	G	G	C	C	G	C	C	G	T	C	G	G	G	C	A	G	T	A				--										
Ref(Chr12 337bp)	G	C	C	C	C	C	C	A	C	G	–	G	G	G	C	A	A	C	G	C	C	G	G																				
A.A Substitution	R-G	A-V	I	I	I	I	I	I	A-P	G-R		D-N	D-N	Q-E	I	I	P-H	I	G-R	I	R-W	R-Q	I	A-P	I	S-G	A-T	V-E	S-R	I	N-S	A-T	"-/A"	M-T	R-Q	L-F	M-V	Q-H	A-V	C-S	V-G	D-N	
Teshio-MsiGhd7BFunc3	G	C	C	C	C	C	C	G	C	G	–	G	G	G	C	A	A	C	G	C	C	G	G	C	C	A	G	A	A	C	A	G	–	C	G	C	A	G	C	C	T	G	b1
Teshio-MsiGhd7BFunc4	A	C	C	C	T	C	C	G	G	C	–	G	G	C	C	G	C	T	G	T	C	G	G	G	C	A	G	T	A	C	A	G	–	T	G	C	A	G	C	C	T	G	b2
Teshio-MsiGhd7BFunc5	A	C	C	C	T	C	C	G	G	G	–	G	G	C	C	G	C	C	G	T	C	G	G	G	C	A	G	A	A	C	A	G	–	T	G	C	A	G	C	C	T	G	b3
Sugadaira-MsiGhd7BFunc3	A	C	C	C	T	C	C	G	G	G	–	G	G	C	C	G	C	C	G	T	C	G	G	G	C	A	G	A	A	C	G	G	–	T	G	C	A	G	C	C	T	G	b4
Sugadaira-MsiGhd7BFunc4	A	C	C	C	T	C	C	G	G	G	–	G	G	G	C	G	C	C	G	T	C	G	G	C	C	A	G	A	A	C	G	G	–	T	G	C	A	G	C	C	T	G	b5
Sugadaira-MsiGhd7BFunc5	A	C	C	C	T	C	C	G	G	G	–	G	G	C	C	G	C	C	G	T	C	G	G	G	C	A	G	T	A	C	A	G	–	T	G	C	A	G	C	C	T	A	b2
Sugadaira-MsiGhd7BFunc6	A	C	C	C	T	C	C	G	G	G	–	A	G	G	C	G	C	C	G	T	C	G	G	G	C	A	G	A	A	C	G	G	–	T	G	C	G	G	C	C	T	G	b6
Miyazaki-MsiGhd7BFunc3	G	C	C	C	C	C	C	G	C	G	–	G	G	G	C	G	C	C	G	T	C	G	G	C	C	A	G	A	A	C	G	G	–	T	G	C	A	C	C	C	T	G	b7
Miyazaki-MsiGhd7BFunc4	G	C	C	C	C	C	C	G	C	G	–	G	A	G	C	G	C	C	G	T	C	A	G	C	C	A	G	A	A	C	G	G	–	T	G	C	A	G	C	C	T	G	b8
PMS-436-MsiGhd7BFunc3	G	C	C	C	C	C	C	G	C	G	–	A	G	G	C	G	C	C	G	T	C	G	G	C	C	A	G	T	A	C	G	A	–	T	G	C	A	G	C	A	T	G	b9
PMS-436-MsiGhd7BFunc4	G	C	C	C	C	C	C	G	C	G	–	A	G	G	C	G	C	C	G	T	C	G	G	C	C	A	G	T	A	C	G	G	–	T	A	C	A	G	C	A	T	G	b10
PMS-436-MsiGhd7BFunc5	G	C	C	C	C	T	C	A	C	G	–	A	G	G	C	A	C	C	G	T	C	G	G	C	C	A	G	T	A	C	G	G	–	T	G	C	A	G	C	C	T	G	b11
PMS-436-MsiGhd7BFunc6	G	C	C	C	C	C	C	G	C	G	–	G	G	G	C	G	C	C	G	T	C	G	G	C	C	A	G	T	A	C	G	G	–	T	G	C	A	G	C	C	T	G	b12
PMS-164-MsiGhd7BFunc3	G	C	C	C	C	C	C	G	C	G	–	G	G	G	C	G	C	C	G	T	T	G	G	C	C	A	G	T	A	C	G	G	–	T	G	C	A	G	C	C	T	G	b12
PMS-164-MsiGhd7BFunc4	G	C	C	C	C	C	C	G	C	G	8bp	G	G		C	G	C	C	G	T	C	G	G	C	C	A	G	T	A	C	G	G	_	T	G	C	A	G	C	C	T	G	b13
PMS-164-MsiGhd7BFunc5	G	C	C	C	C	C	C	G	C	G	8bp	G	G		C	G	C	C	G	T	C	G	G	C	C	A	G	T	A	C	G	G	_	T	G	C	A	G	C	C	T	G	b13
PMS-306-MsiGhd7BFunc3	G	C	C	C	C	C	C	A	C	G	–	G	G	G	C	A	A	C	G	T	C	G	G	C	C	A	G	T	A	C	G	G	–	T	G	C	A	G	C	C	T	G	b14
PMS-306-MsiGhd7BFunc4	G	C	C	C	C	C	C	G	C	G	–	A	G	G	C	G	C	C	G	T	C	G	G	C	C	A	G	T	A	C	G	G	–	T	G	C	A	G	C	C	T	G	b15
PMS-306-MsiGhd7BFunc5	G	C	C	C	C	C	C	A	C	G	–	G	G	G	C	A	A	C	G	T	C	G	G	C	C	G	G	T	A	C	G	G	–	T	G	C	A	G	T	C	T	G	b14
PMS-226-MsiGhd7BFunc3	G	C	T	T	C	C	C	G	C	G	–	G	G	G	C	G	C	C	G	T	C	G	G	C	C	A	G	T	A	C	G	G	–	T	G	C	A	G	C	A	T	G	b9
PMS-226-MsiGhd7BFunc4	G	C	C	C	C	C	C	G	C	G	–	A	G	G	C	G	C	C	G	T	C	G	G	C	C	A	G	T	A	C	G	G	–	T	G	C	A	G	C	A	T	G	b16
Onna-1a-MsiGhd7BFunc3	G	C	C	C	C	C	C	G	C	G	–	G	G	G	C	G	C	C	G	T	C	G	G	C	C	A	G	T	A	C	G	G	–	T	G	C	A	G	C	C	T	G	b12
Onna-1a-MsiGhd7BFunc4	G	C	C	C	C	C	C	G	C	G	–	G	G	G	C	G	C	C	G	T	C	G	G	C	C	A	G	T	A	C	G	G	–	T	A	C	A	G	C	A	T	G	b17
PMS-359-MsiGhd7BFunc3	G	C	C	C	C	C	C	G	C	G	–	G	A	G	C	G	C	C	G	T	C	G	T	C	C	A	G	T	A	C	G	G	–	T	G	T	A	G	C	C	T	G	b18
PMS-359-MsiGhd7BFunc4	G	C	C	C	C	C	C	G	C	G	–	A	G	G	C	G	C	C	G	T	C	G	G	C	C	A	A	T	A	C	G	G	–	T	G	C	A	G	C	C	T	G	b19
PMS-375-MsiGhd7BFunc3	G	C	T	T	C	C	C	G	C	G	–	A	G	G	C	G	C	C	G	T	C	G	G	C	C	A	G	T	A	C	G	G	–	T	G	C	A	G	C	C	T	G	b11
PMS-375-MsiGhd7BFunc4	G	C	C	C	C	C	T	G	C	G	–	G	G	G	C	G	C	C	G	T	C	G	G	C	C	A	G	T	A	A	G	G	–	T	A	C	A	G	C	C	T	G	b20
PMS-375-MsiGhd7BFunc5	G	T	C	C	C	C	C	G	C	G	–	G	G	G	T	G	C	C	G	T	C	G	G	C	C	A	G	T	A	A	G	G	–	T	A	C	A	G	C	C	T	G	b20
PMS-382-MsiGhd7BFunc3	G	C	C	C	C	C	C	G	C	G	–	G	G	G	C	G	C	C	G	T	C	G	G	C	C	A	G	T	A	C	G	G	–	T	G	C	A	G	C	C	T	G	b12
PMS-382-MsiGhd7BFunc4	G	C	C	C	C	C	C	G	C	G	–	A	G	G	C	G	C	C	G	T	C	G	G	C	C	A	G	T	A	C	G	G	–	T	G	C	A	G	C	C	G	G	b15
PMS-382-MsiGhd7BFunc5	G	C	C	C	C	C	C	G	C	G	–	G	G	G	C	G	C	C	G	T	C	G	G	C	C	A	G	T	A	A	G	G	–	T	A	C	A	G	C	C	T	G	b20
US56-0022-03-MsiGhd7BFunc3	G	C	C	C	C	C	C	A	C	G	–	G	G	G	C	G	A	C	C	C	C	G	G	C	C	A	G	T	A	C	G	G	–	T	G	C	A	G	C	C	T	G	b14
US56-0022-03-MsiGhd7BFunc4	G	C	C	C	C	C	C	G	C	G	–	A	G	G	C	G	C	C	G	T	C	G	G	C	C	A	A	T	C	C	G	G	–	T	G	C	A	G	C	C	T	G	b19
US56-0022-03-MsiGhd7BFunc5	G	C	C	C	C	C	C	G	C	G	–	G	G	C	C	G	C	C	G	T	C	G	G	C	A	A	G	T	A	C	G	G	3bp	T	G	C	A	G	C	C	T	G	b21

Note: SNVs: single nucleotide variants. Cells with light blue represent SNVs aligned with *MsiGhd7B* (two segments) in *Miscanthus sinensis* v7.1 genome (Mitros *et al.*, 2020). SNVs located in the domain were highlighted in the red. The 'b' prefix with numbers in the right column were the protein types. Cells with different colors in the right column represent different variant types that occurred in more than one accession; variant types that were observed only once have a gray background, corresponding to Figure 3.5.

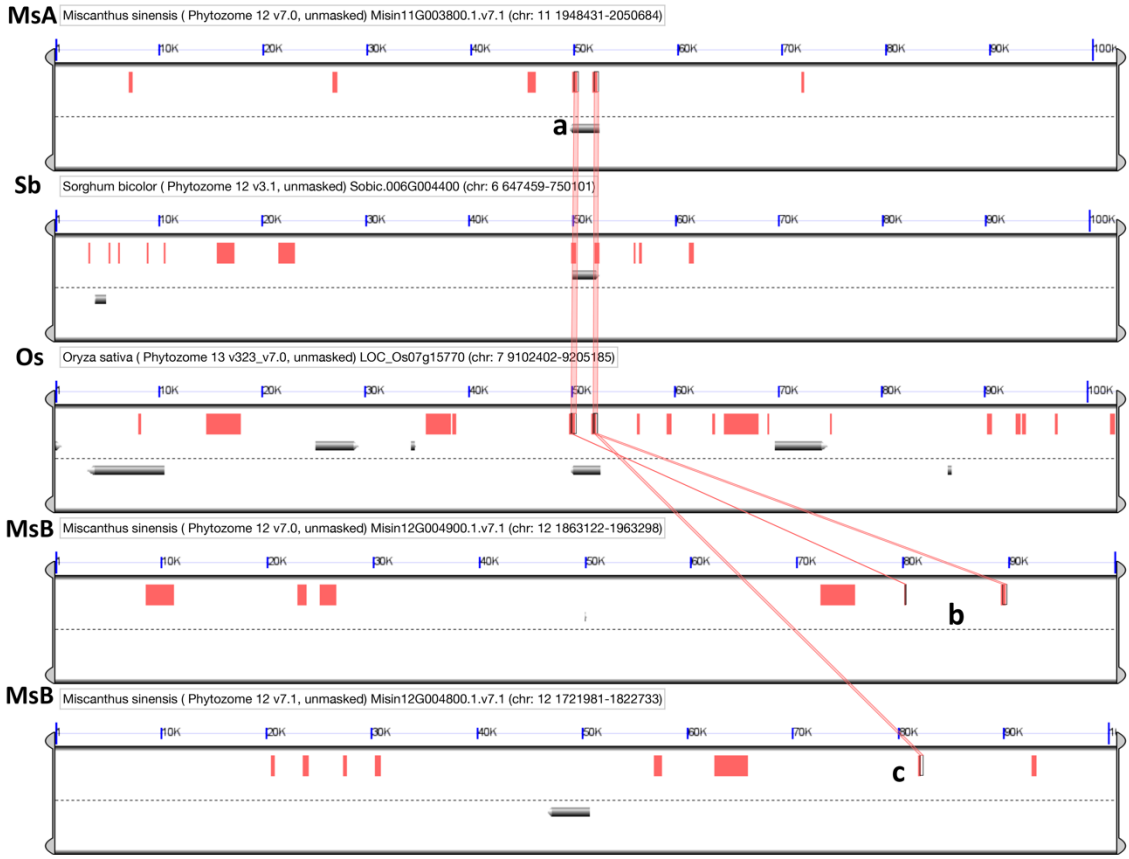


Figure S3.1 Chromosome organization near *Ghd7* gene regions (100 kbp) from *Oryza sativa*, *Sorghum bicolor* and *Miscanthus sinensis*. Only high-scoring sequence pairs (HSPs) between adjacent regions are drew in the red boxes. The dashed line in the middle of each region represents the division between the top (5′ on left) and the bottom (5′ on right) strand. The full gene models are drawn as gray arrows directly above or below this line. Colinear genes or highly similar fragments within the aligned region are connected by red lines. 'a' and 'b, c' represent the colinear gene and highly similar fragments, respectively, among *O. sativa* (LOC_Os07g15770), *S. bicolor* (Sobic.006G004400) and *M. sinensis* including *MsiGhd7A* (Misin11G003800, a; MsA) and *MsiGhd7B* (Chr.12: 1952462..1953905, b; Chr. 12: 1803910..1804245, c; MsB).

Chapter 4

Diurnal expression patterns of several flowering-related genes associated with photoperiod perception

4.1 Introduction

The consistent increase and decrease in the level of gene expression or protein over a cycle of 24 h, known as a biological rhythm, can be controlled by endogenous (internal biological circadian clock) or exogenous (external) stimuli (Webb, 2003; McWatters and Devlin, 2011). Endogenous rhythms that cycle over a period of time close to 24 h are also called circadian rhythms. One of the most common exogenous rhythms is the synchronization to the length of day and night, known as a diurnal rhythm; a key diagnostic is that these rhythms cease to persist when exposed to constant light or dark conditions (Schaffer *et al.*, 2001; Webb, 2003; Yeang, 2015). Although many gene expression profiles correlate with the day/night 24-h cycle, not all genes are directly affected by light and/or dark periods but are actually responding to fluctuations in photosynthate compounds or other internal rhythms. In the latest research, Dong *et al*. (2021) showed that relatively short days could accelerate floral induction of *M. sinensis*, but under a critical threshold, especially for genotypes adapted to high latitudes, otherwise it will remain its vegetative growth phase. Until now, only *CO/Hd1* in *M. sinensis* were investigated for sequence diversity (Nagano *et al.*, 2015), whereas, the gene expression pattern of *MsiHd1* is still not be identified.

In Chapters 2 and 3, the allelic diversity and gene expression patterns of *Ghd8* and *Ghd7* in twelve *M. sinensis* accessions were characterized, showing the photoperiodic perception. Given that the complex genetic basis of flowering regulation in *M. sinensis*, which may be controlled by multiple pathways, it would be worthwhile to investigate the difference in the expression throughout the main photoperiodic pathway and the relationship between these expression profiles and flowering time among *M. sinensis* accessions. In rice, the mRNA transcript level of *Ghd7* is regulated by *Ghd8* through the GHD8-HAP5b-HD1 complex (Wang *et al.*, 2019*a*). In other words, gene products of GHD8, HAP5b, and HD1 form a complex that acts as a transcription factor to bind to the specific CCAAT-box region in the *Ghd7* promoter and upregulate expression of *Ghd7*, which subsequently downregulates *Ehd1* and *Hd3a* (probably *RFT1*), the rice orthologs of *FT* that has the florigenic activity, and then lengthens the time to flower. The FT proteins are the members of PEBPs family, of which, two PEBP-family genes in sorghum, *SbCN8* and *SbCN12* are the colinear orthologs of maize *ZCN8* and *ZCN12*, which possess florigen activity (Meng *et al.*, 2011; Murphy *et al.*, 2011, 2014; Castelletti *et al.*, 2020). An additional PEBP-family gene *CN15*, the ortholog of rice *Hd3a*, is present in sorghum and also mediated by *PhyB* to regulate flowering time of sorghum (Yang *et al.*, 2014*a*). Collectively, *SbCN8*, *SbCN12* and *SbCN15* are regulated by *SbCO* and *SbEhd1* (Murphy *et al.*, 2014; Yang *et al.*, 2014*b*; Abdul-Awal *et al.*, 2020).

In light of this, in this chapter, five flowering genes (*Hd1*, *Ehd1*, *CN8*, *CN12* and *CN15*) correlated with *Ghd7* and *Ghd8* in this pathway were selected for qRT-PCR in 24 h light-dark cycle under artificial LD and SD, since the heading date of

twelve *M. sinensis* accessions could not be interpreted by these two major genes. The objective is to investigate whether the differentiation of their expression patterns exist in individuals from different regions. Furthermore, does this flowering regulation pathway conserved in *M. sinensis*? Do downstream genes have similar function relative to the rice and sorghum in response to the photoperiod?

4.2 Materials and methods

4.2.1 Plant materials and growth conditions

Twelve *M. sinensis* were studied and subjected to day length treatment experiments as mentioned in Section 2.2.1 in Chapter 2.

4.2.2 RNA extraction and cDNA synthesis

Leaves were sampled from fully expanded healthy leaves at ZT 3, 9, 15 and 21h in the growth chamber under LD and SD. The protocol for total RNA extraction and gene expression analysis were described in Section 2.2.3 in Chapter 2.

4.2.3 Candidate genes

Five *M. sinensis* genes homologous to sorghum genes, which were identified in association with photoperiod perception and floral induction in sorghum from previously published reports (Murphy *et al.*, 2011, 2014; Yang *et al.*, 2014*a*,*b*; Casto *et al.*, 2019), were selected as the candidate genes for gene expression analysis in the present study (Table S4.1). Gene locus ID of interest associated with *Ghd8* and *Ghd7* included: *CO*/*Hd1*, *Ehd1*, *CN8*, *CN12* and *CN15*, see Table S4.1 for *M. sinensis* and orthologs of rice, sorghum and maize. The housekeeping gene

used to normalize the qRT-PCR was *ACTIN* (Misin17G008500). Primers for *ACTIN*, *Ehd1* and *CO*/*Hd1*, were designed based on the obtained sequence in the present lab with published or unpublished data, other genes' primers were designed according to the annotated database of *Miscanthus sinensis* v7.1 (Mitros *et al.*, 2020) and *Sorghum bicolor* v3.1 (McCormick *et al.*, 2018) in Phytozome v13 (https://phytozome-next.jgi.doe.gov). All selected primers followed the criteria with (i) a high PCR efficiency and (ii) the solo product that confirmed by the melting-curve analysis of qRT-PCR and gel electrophoresis. Primer sequences for all genes were listed in Table S4.2.

4.2.4 Gene expression analysis

The transcript levels for candidate genes were determined by qRT-PCR. The PCR reactions (20 μL) contained 4.6 μL of the cDNA synthesis reaction mixture diluted to 1/15 th of its original volume, 5 μL of 1.2 μM primer premix, 0.4 μL ROX Reference Dye (50 ×) and 10 μL of TB Green® Premix Ex Taq™ II (Tli RNaseH Plus) (TaKaRa Bio, Shiga, Japan). Expression levels were determined on a StepOnePlus™ Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) with cycling conditions of 95°C for 5 min, followed by 40 cycles of 95 °C for 10 sec, 60 °C for 20 sec and 72 °C for 30 sec. Values were normalized to *ACTIN*. A reaction mixture without reverse transcriptase was used as a control to confirm the absence of genomic DNA contamination. Also, the controls without template showed no amplification for the results to be accepted (Nolan *et al.*, 2006). Amplification of a single DNA fragment was confirmed by melting-curve analysis

of qRT-PCR and 2 % NuSieve™ 3:1 Agarose (FMC BioProducts, Rockland, Maine, USA) gel electrophoresis.

4.2.4 Data analysis

The double delta threshold cycle (Ct) method was used to process qRT-PCR results (Bookout and Mangelsdorf, 2003; Nolan *et al.*, 2006). Averages and standard errors of relative expression levels were calculated for three independently synthesized cDNAs. Relative changes in mean ± standard error of the mean (SE) gene expression was analyzed in Microsoft Excel (Microsoft Office 2019, Microsoft Inc., Seattle, WA, USA) and then exported to GraphPad Prism 9 software (GraphPad Software, San Diego, CA, USA) for visualization. Statistical tests for differences among means were conducted by Student's *t*-test or analyses of variances (ANOVAs) using GraphPad Prism 9 software (GraphPad Software, San Diego, CA, USA).

4.3 Results and discussion

4.3.1 MsiGHD8-HD1 complex activate the transcription of *MsiGhd7* under long days

To investigate the possibility that rice GHD8-HD1 complex regulates *Ghd7* model in *M. sinensis*, *Hd1*/*CO* mRNA transcript abundance was analyzed. *Hd1*/*CO* is the central regulator for the flowering pathway in *Arabidopsis*, rice and sorghum (Yano *et al.*, 2000; Yang *et al.*, 2014*b*; Shim *et al.*, 2017), and was also identified in *M. sinensis* (Nagano *et al.*, 2015). As opposite to *CO* that promotes *A. thaliana* flowering under LD (Putterill *et al.*, 1995) and may represses flowering in SD

(Luccioni *et al.*, 2019), *Hd1* performs dual functions on the rice flowering that promotes heading under SD and inhibits under LD (Yano *et al.*, 2000; Izawa *et al.*, 2002). In *M. sinensis*, *Hd1*/*CO* expressed in both LD and SD photoperiods (Figure 4.1), in accordance with that of in rice and sorghum (Yano *et al.*, 2000; Izawa *et al.*, 2002; Yang *et al.*, 2014*b*). The relative expression of *MsiHd1*/*CO* increased from the afternoon and peaked after the onset of the dark period (ZT15 or ZT21), and then decreased to almost no detectable expression even 3h after the beginning of the light period under both daylengths (Figure 4.1). In *A*. *thaliana*, a similar increase in *CO* expression in the evening is observed due to the interaction between GI and blue light-activated FKF1 (Sawa *et al.*, 2007), which were also identified in rice (Izawa *et al.*, 2011) and sorghum (Yang *et al.*, 2014*b*; Abdul-Awal *et al.*, 2020). At ZT15, no significant difference of *MsiHd1*/*CO* mRNA transcripts were observed between LD and SD in four Japanese accessions (Teshio, Sugadaira, Miyazaki and Onna-1a), one *M. floridulus* (US56-0022-03) from New Caledonia, and three Chinese accessions (PMS-164, PMS-306 and PMS-226). By contrast, *MsiHd1*/*CO* in SD expressed significantly higher than in LD at the peak of ZT15 among three accessions (PMS-359, PMS-375 and PMS-382) from Southeastern China with latitude varied from 18.9 °N to 22.9 °N and one accession (PMS-436) from high latitude in the Northeastern China.

In Chapter 2, *MsiGhd8* expression patterns were observed, though frequent peak was at ZT9 or ZT15, it still expressed at night period (Figures 2.4 and 4.2). Under this condition, there is an opportunity that Ghd8 and Hd1 could form the complex to recognize the CCAAT-box element and/or CORE motif of the *Ghd7* promoter and then activate *MsiGhd7* expression. In Chapter 3, the mRNA transcript

level of *MsiGhd7* was frequently peaked in the morning at ZT3 (Figures 3.6 and 4.2). Besides, the CCAAT-box was detected in *Ghd7* promoter from reference genome of *Miscanthus sinensis* v7.1 (Mitros *et al.*, 2020). Considering the above facts, the results of gene expression profile models supported the hypothesis in present study, to some extent, the Ghd8-Hd1 complex that activates with *Ghd7* model might also be conserved in *M. sinensis*. Further studies are needed to confirm this likely molecular interaction by biochemistry analysis in *M. sinensis.*

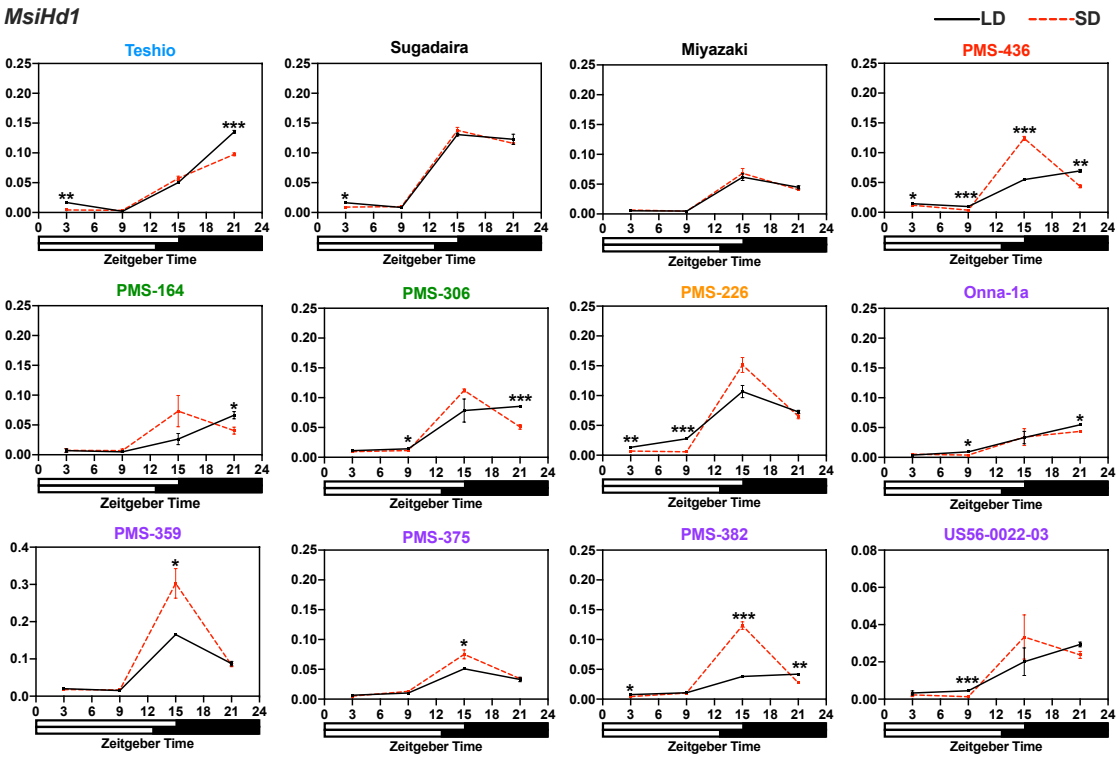


Figure 4.1 Diurnal expression of *MsiHd1* in 12 *Miscanthus sinensis* accessions under long days (15 h, LD; solid black lines) and short days (12.5 h, SD; red dashed line). The *y*-axis represents relative mRNA levels normalized to *ACTIN* transcript levels. The numbers below the *x*-axis indicate Zeitgeber times (ZT) of the day. The white bar at the bottom of each graph indicates the light period and the black bar indicates the dark period. Mean ± 1SE for three replications are given for each data point. Asterisks indicate significant difference between the two means under LD and SD at the same ZT of the day (Student's *t*-test, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). No asterisk indicates the difference between the two means is not statistically significant ($P > 0.05$). Accessions' names are colored to represent *M. sinensis* genetic groups previously described by Clark *et al.* (2014, 2015); Sugadaira and Miyazaki are changed from yellow to black for making it clear.

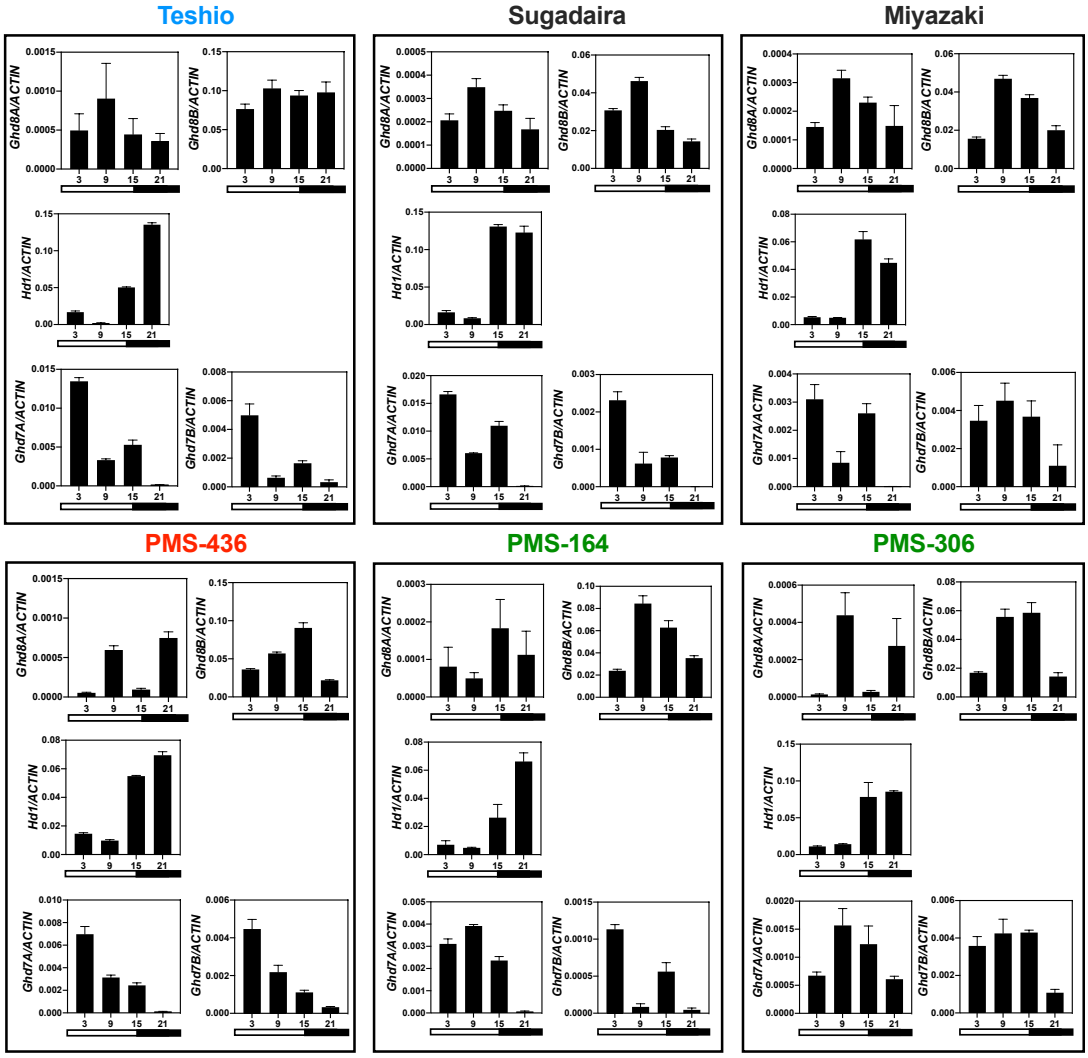


Figure 4.2 Expression patterns of *MsiGhd8*, *MsiHd1* and *MsiGhd7* under long days (15h, LD). The *x*-axis indicates Zeitgeber times (ZT) of the day and *y*-axis represents relative mRNA levels normalized to *ACTIN* transcript levels. The white and black bar at the bottom of each graph indicate the light and dark period, respectively. Accessions' names are colored to represent *Miscanthus sinensis* genetic groups previously described by Clark *et al*. (2014, 2015); Sugadaira and Miyazaki are changed from yellow to black for making it clear. **(Continued)**

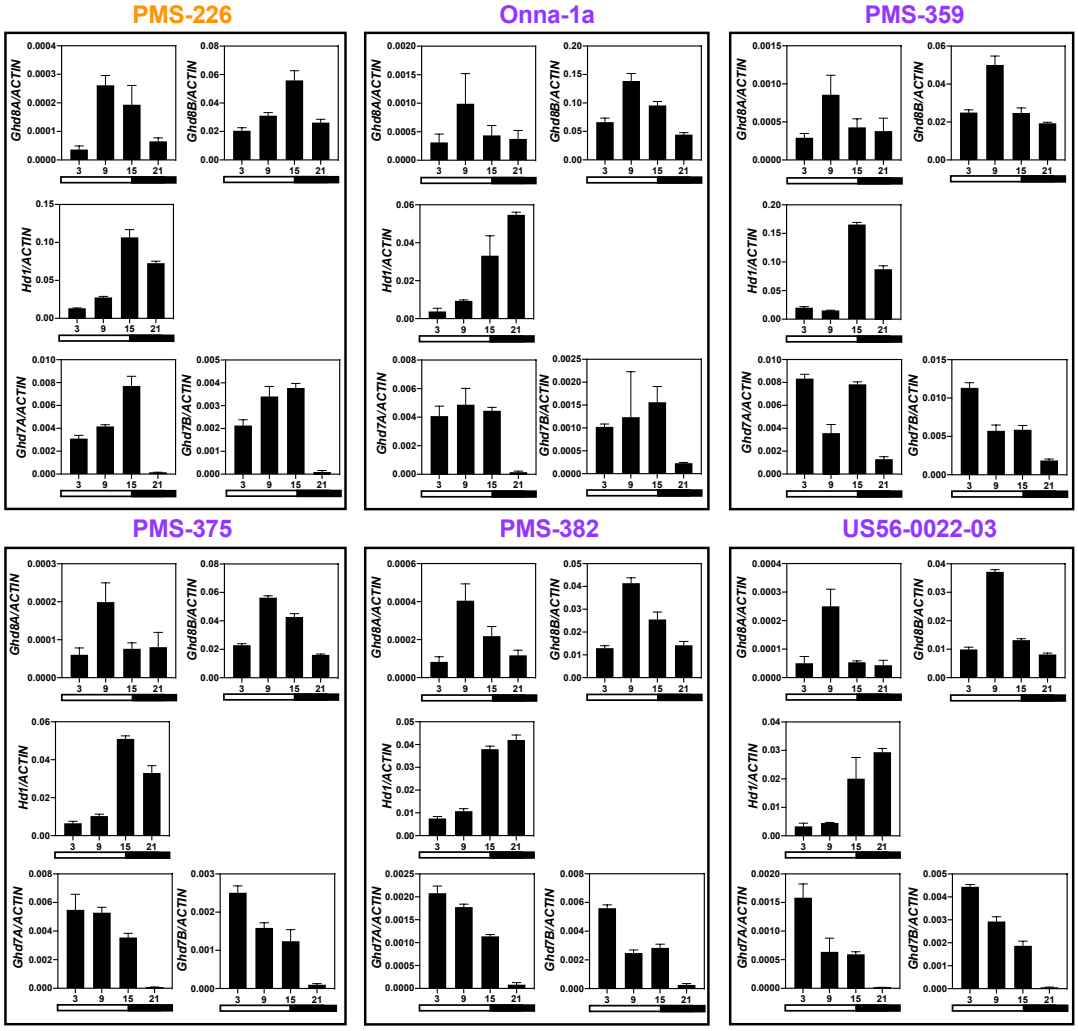


Figure 4.2 Expression patterns of *MsiGhd8*, *MsiHd1* and *MsiGhd7* under long days (15h, LD). The *x*-axis indicates Zeitgeber times (ZT) of the day and *y*-axis represents relative mRNA levels normalized to *ACTIN* transcript levels. The white and black bar at the bottom of each graph indicate the light and dark period, respectively. Accessions' names are colored to represent *Miscanthus sinensis* genetic groups previously described by Clark *et al*. (2014, 2015); Sugadaira and Miyazaki are changed from yellow to black for making it clear.

4.3.2 *Ehd1* expression is influenced by multiple genes

Ghd7 performs as one of the upstream genes of *Ehd1* to repress *Ehd1* expression in rice (Xue *et al.*, 2008; Itoh *et al.*, 2010) and sorghum (Murphy *et al.*, 2014; Yang *et al.*, 2014*a*) under LD condition. *Ehd1*, encoding a B-type response regulator transcription factor, is unique to the grasses that has been shown to promote flowering in rice and sorghum (Doi *et al.*, 2004; Murphy *et al.*, 2011), and also be found in *M. sinensis* genome (Table S4.1). To examine whether *MsiGhd7* regulated the expression of *Ehd1* in *M. sinensis*, *MsiEhd1* expression patterns in leaves of twelve accessions grown in LD and SD were quantified and compared (Figure 4.3). Except for PMS-226, which showed a relative high mRNA transcript level of *MsiEhd1* in LD compared to SD, the expression of *MsiEhd1* in LD for the other eleven accessions were greatly repressed relative to SD (Figure 4.3). Whereas, *MsiGhd7A* and *MsiGhd7B* for all twelve accessions had a much greater extent in LD than that in SD, and the transcript level of *MsiGhd7* increased especially in the early morning under LD condition (Figure 4.3). These conserved expression patterns suggested that the *Ghd7*-*Ehd1* regulation model also existed in *M. sinensis*. In LD, *MsiGhd7* played a role in inhibiting *MsiEhd1* expression, corresponding with rice (Xue *et al.*, 2008) and sorghum (Murphy *et al.*, 2014). In SD, the expression of *MsiEhd1* was induced in the dark period under SD condition (Figure 4.3), showing the similar expression pattern to *MsiHd1*/*CO* (Figure 4.1), consistent with sorghum *Ehd1* that upregulated by *SbCO* (Yang *et al.*, 2014*b*; Abdul-Awal *et al.*, 2020).

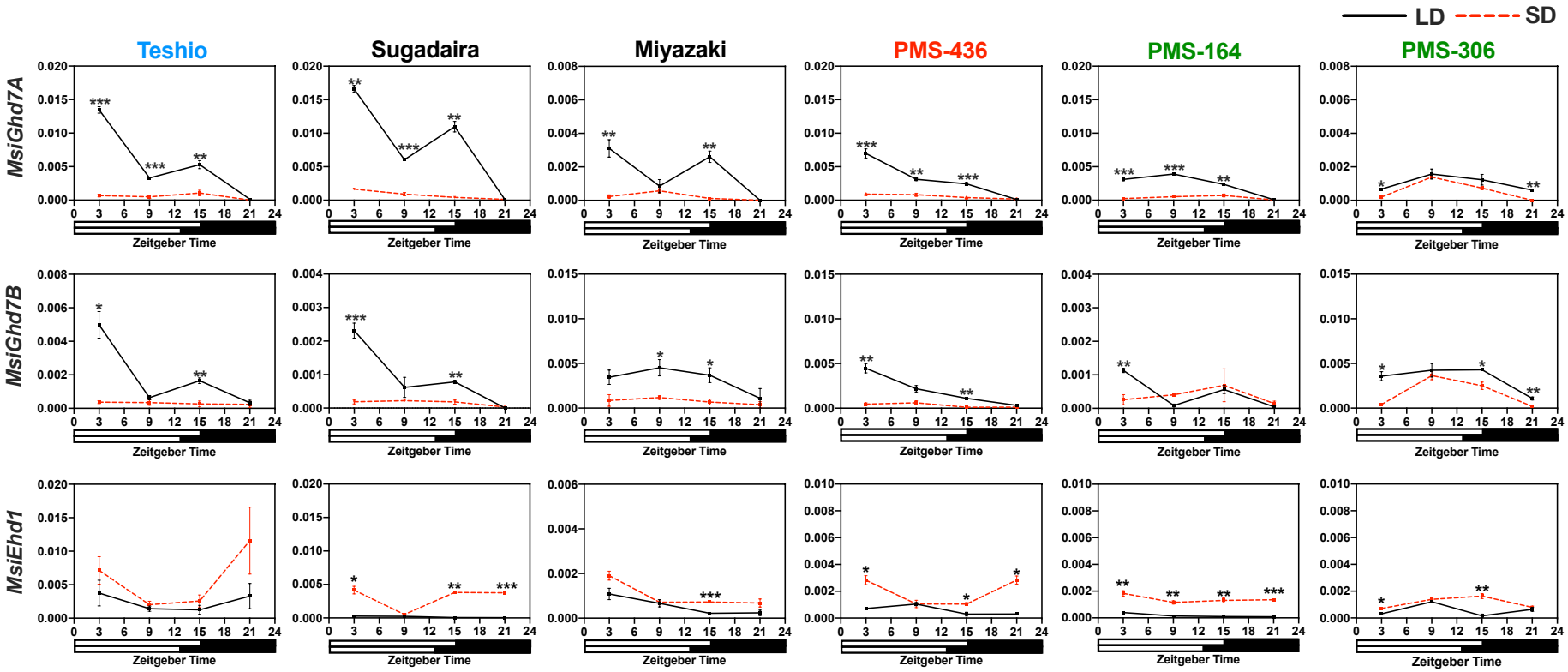


Figure 4.3 Diurnal expression of *MsiGhd7A*, *MsiGhd7B* and *MsiEhd1* in 12 *Miscanthus sinensis* accessions under long days (15h, LD, solid line) or short days (12.5h, SD, red dashed line) conditions. The *y*-axis represents relative mRNA levels normalized to *ACTIN* transcript levels. The numbers below the *x*-axis indicate Zeitgeber times (ZT) of the day. The white bar at the bottom of each graph indicates the light period and the black bar indicates the dark period. Mean ± 1SE for three replications are given for each data point. Asterisks indicate significant difference between the two means under LD and SD at the same ZT of the day (Student's *t*-test, *$P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). No asterisk indicates the difference between the two means is not statistically significant ($P > 0.05$). Accessions' names are colored to represent *M. sinensis* genetic groups previously described by Clark *et al*. (2014, 2015); Sugadaira and Miyazaki are changed from yellow to black for making it clear. **(Continued)**

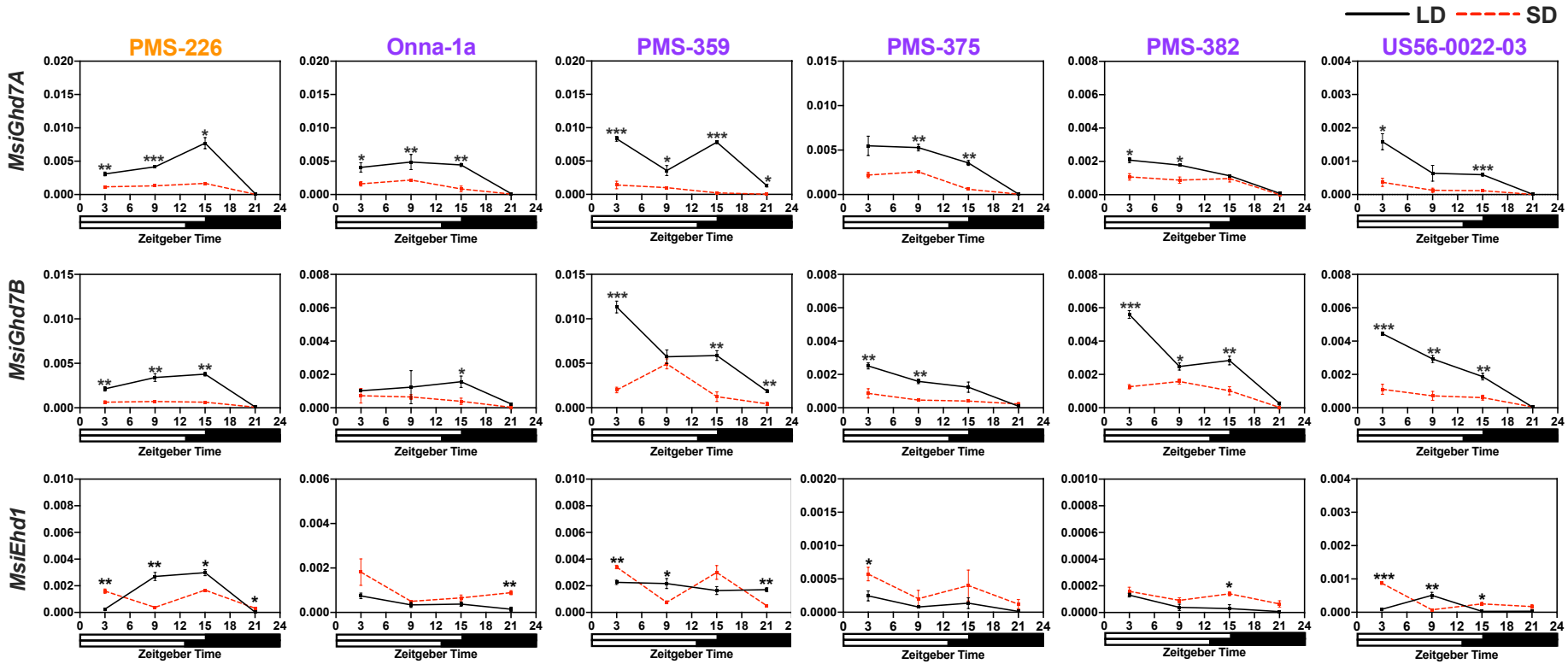


Figure 4.3 Diurnal expression of *MsiGhd7A*, *MsiGhd7B* and *MsiEhd1* in 12 *Miscanthus sinensis* accessions under long days (15h, LD, solid line) or short days (12.5h, SD, red dashed line) conditions. The *y*-axis represents relative mRNA levels normalized to *ACTIN* transcript levels. The numbers below the *x*-axis indicate Zeitgeber times (ZT) of the day. The white bar at the bottom of each graph indicates the light period and the black bar indicates the dark period. Mean ± 1SE for three replications are given for each data point. Asterisks indicate significant difference between the two means under LD and SD at the same ZT of the day (Student's *t*-test, *$P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). No asterisk indicates the difference between the two means is not statistically significant ($P > 0.05$). Accessions' names are colored to represent *M. sinensis* genetic groups previously described by Clark *et al*. (2014, 2015); Sugadaira and Miyazaki are changed from yellow to black for making it clear.

4.3.3 *Ehd1* induces the expression of downstream florigens

The metabolite responsible for signaling changes that control and/or trigger flowering has generally been called florigen. The florigen has been identified as the product of *FT*, whereby transcripts produced in the leaves are transported to the meristem, and the translated protein affects the transition from shoot apical meristem to floral meristem in conjunction with other proteins (Tamaki *et al.*, 2007; Danilevskaya *et al.*, 2008; Wolabu *et al.*, 2016). In rice, *Ehd1* directly up-regulates two florigens: *Hd3a* and *RFT1* (Doi *et al.*, 2004). However, the *RFT1* ortholog has not been detected in *M. sinensis* and sorghum genome. Moreover, similar to sorghum, *M. sinensis* possessed a set of *FT*-like genes, which shared high sequence similarity with *ZCN* sets in maize. One of these, *ZCN8*, was determined to be the candidate for florigen in maize because of its transcriptional response to photoperiod and its effects on the flowering time in the temperate maize (Meng *et al.*, 2011; Castelletti *et al.*, 2020). In Figure 4.4, Teshio originated from the highest latitude showed no difference in *MsiCN8* expression in response to day length, indicated that *MsiCN8* might be controlled by other unknown genes or mutation occurred in Teshio *MsiCN8*. Apart from Teshio, almost no detectable *MsiCN8* mRNA transcripts for other eleven accessions were observed in LD, similar to sorghum and maize *CN8* (Meng *et al.*, 2011; Murphy *et al.*, 2011, 2014; Castelletti *et al.*, 2020).

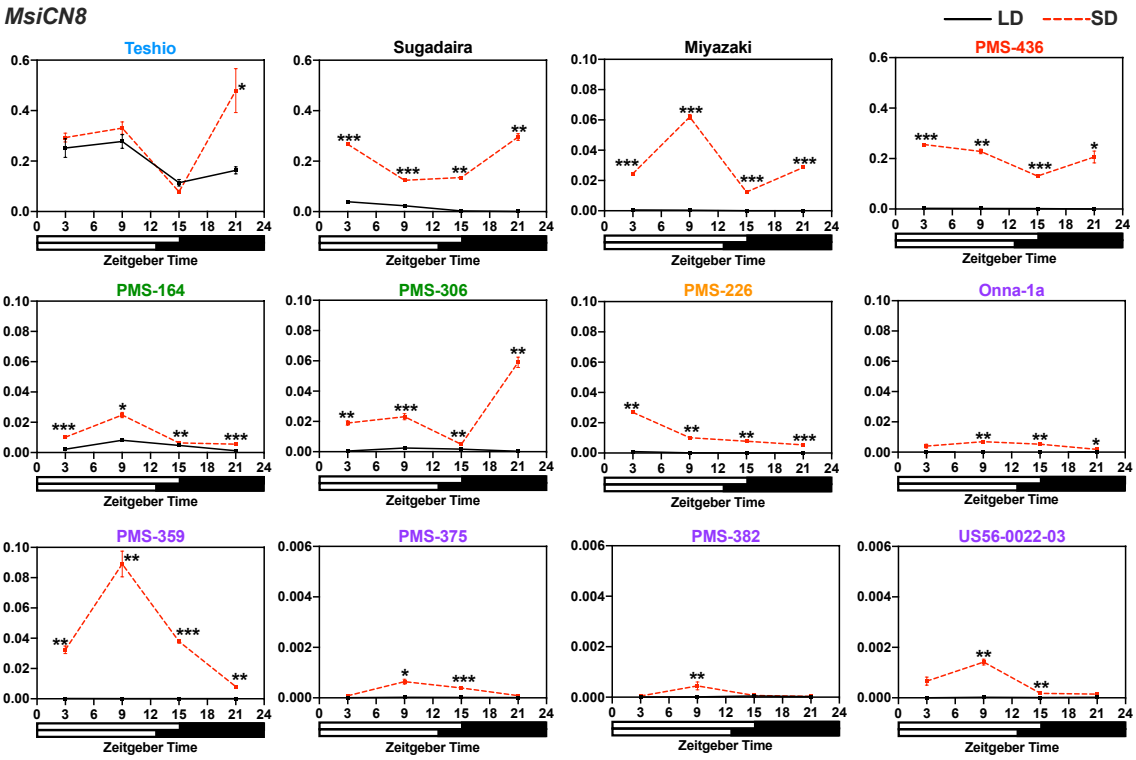


Figure 4.4 Diurnal expression of *MsiCN8* in 12 *Miscanthus sinensis* accessions under long days (15 h, LD; solid black lines) and short days (12.5 h, SD; red dashed line). The *y*-axis represents relative mRNA levels normalized to *ACTIN* transcript levels. The numbers below the *x*-axis indicate Zeitgeber times (ZT) of the day. The white bar at the bottom of each graph indicates the light period and the black bar indicates the dark period. Mean ± 1SE for three replications are given for each data point. Asterisks indicate significant difference between the two means under LD and SD at the same ZT of the day (Student's *t*-test, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). No asterisk indicates the difference between the two means is not statistically significant ($P > 0.05$). Accessions' names are colored to represent *M. sinensis* genetic groups previously described by Clark *et al.* (2014, 2015); Sugadaira and Miyazaki are changed from yellow to black for making it clear.

In maize, a second candidate for florigen was identified as *ZCN12*, a regulator of maize flowering, and co-expressed with *ZCN8* (Meng *et al.*, 2011; Castelletti *et al.*, 2020). In sorghum, *CN12* was expressed in response to SD prior to floral initiation, and with the pattern similar to it seen for *SbCN8*. By contrast, in *M. sin*ensis, unlike *MsiCN8*, not all accessions had a significant differential expression level of *MsiCN12* between LD and SD (Figure 4.5). Expressions of *MsiCN12* for two accessions (PMS-226 and US56-0022-03) displayed a relatively higher in LD than that in SD, and almost similar for another two accessions (Onna-1a and PMS-375) under both LD and SD.

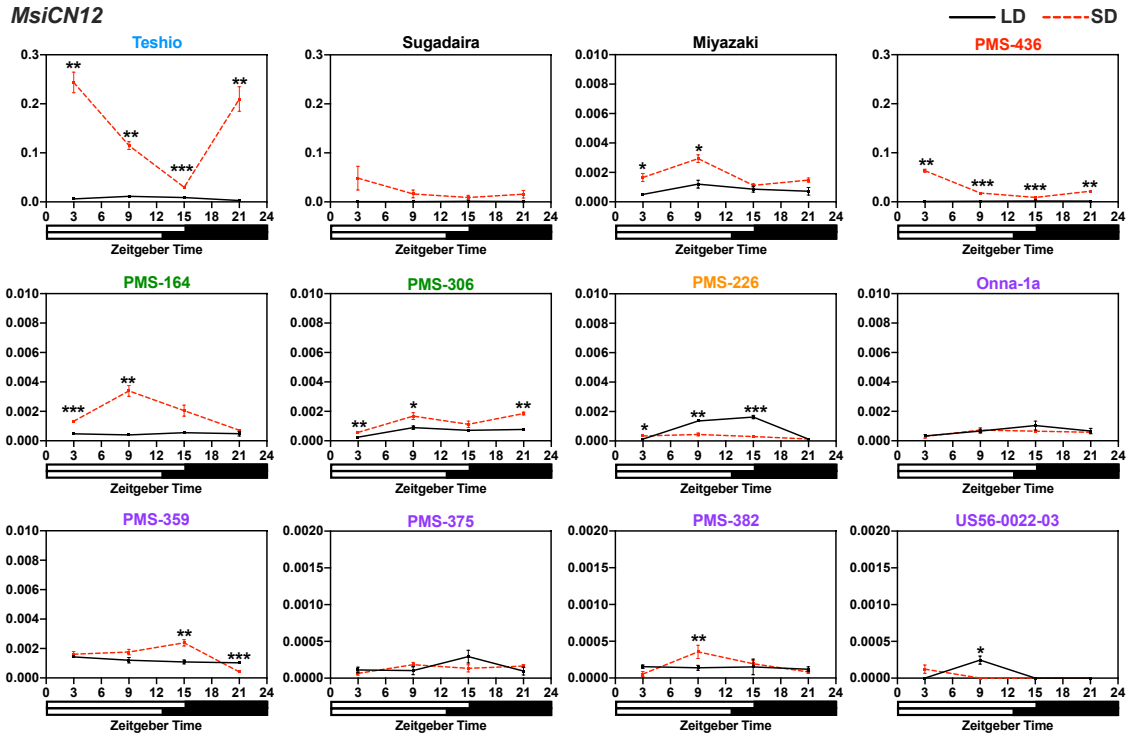


Figure 4.5 Diurnal expression of *MsiCN12* in 12 *Miscanthus sinensis* accessions under long days (15 h, LD; solid black lines) and short days (12.5 h, SD; red dashed line). The *y*-axis represents relative mRNA levels normalized to *ACTIN* transcript levels. The numbers below the *x*-axis indicate Zeitgeber times (ZT) of the day. The white bar at the bottom of each graph indicates the light period and the black bar indicates the dark period. Mean ± 1SE for three replications are given for each data point. Asterisks indicate significant difference between the two means under LD and SD at the same ZT of the day (Student's *t*-test, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). No asterisk indicates the difference between the two means is not statistically significant ($P > 0.05$). Accessions' names are colored to represent *M. sinensis* genetic groups previously described by Clark *et al.* (2014, 2015); Sugadaira and Miyazaki are changed from yellow to black for making it clear.

In rice, *Hd3a* is a major florigen signal in SD (Itoh *et al.*, 2010). *ZCN15* and *SbCN15*, the orthologs of rice *Hd3a*, were found in maize and rice, respectively. However, in maize, *ZCN15* expressed predominantly in kernels and not in leaves, ruling out its potential as maize florigen (Danilevskaya *et al.*, 2008). In sorghum, *SbCN15* was suggested as a minor target of photoperiod regulation because of a small extent by variation in photoperiod (Murphy *et al.*, 2011, 2014). In *M. sinensis*, *MsiCN15* expressed at a lower level relative to *MsiCN8* and *MsiCN12*, similar to sorghum (Murphy *et al.*, 2011, 2014), especially seven accessions originated from the latitude below 30 °N (Figure 4.6). Yang *et al*. (2014*b*) pointed out that *SbCN15* might also be responsible for early flowering induced by shading or GA pathway. Whether this potential function exists in *M. sinensis*, further study is needed.

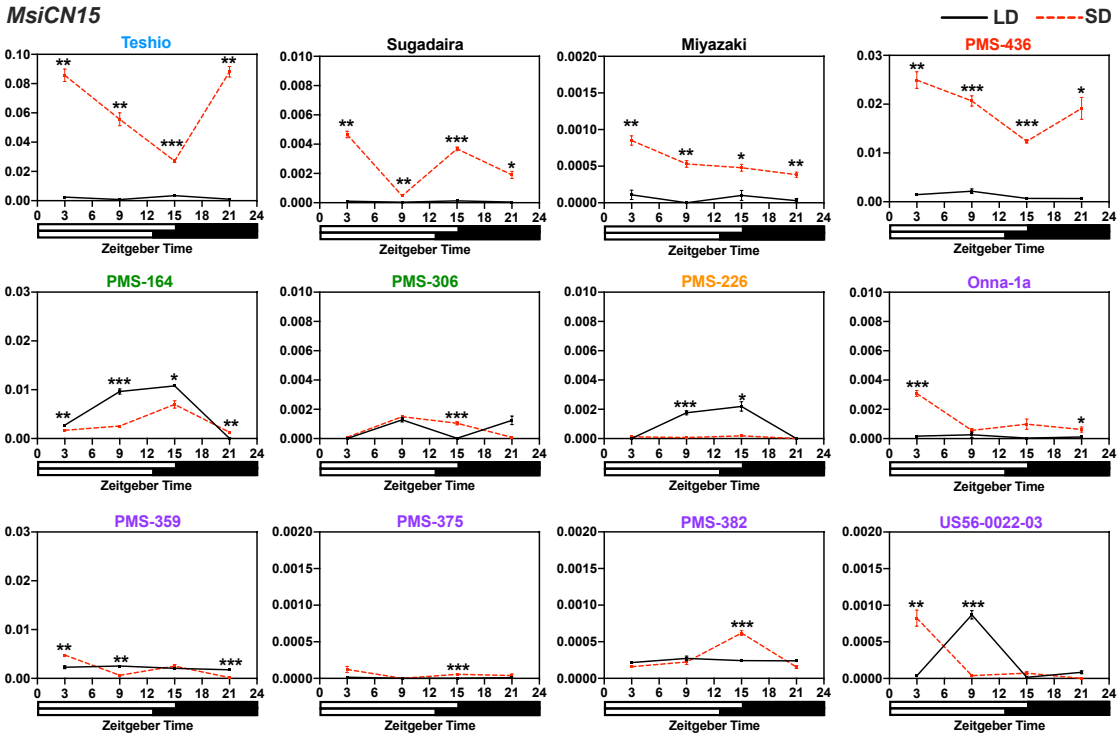


Figure 4.6 Diurnal expression of *MsiCN15* in 12 *Miscanthus sinensis* accessions under long days (15 h, LD; solid black lines) and short days (12.5 h, SD; red dashed line). The *y*-axis represents relative mRNA levels normalized to *ACTIN* transcript levels. The numbers below the *x*-axis indicate Zeitgeber times (ZT) of the day. The white bar at the bottom of each graph indicates the light period and the black bar indicates the dark period. Mean ± 1SE for three replications are given for each data point. Asterisks indicate significant difference between the two means under LD and SD at the same ZT of the day (Student's *t*-test, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). No asterisk indicates the difference between the two means is not statistically significant ($P > 0.05$). Accessions' names are colored to represent *M. sinensis* genetic groups previously described by Clark *et al.* (2014, 2015); Sugadaira and Miyazaki are changed from yellow to black for making it clear.

In Figure 4.7, for all accessions, when the expression of *MsiEhd1* showed a relatively high expression, at least one *FT* expressed at a high transcriptional level. Additionally, when *MsiEhd1* mRNA transcript level was lower, all the *FTs* displayed low expressions (PMS-375, PMS-382, US56-0022-03). A higher expression of *MsiEhd1* in LD relative to SD was observed in PMS-226 (Figure 4.3), and the expression of *MsiCN12* in this accession also showed a higher level in LD. Taken together, regardless of day length, *Ehd1* could induce the expression of these three *FTs* (*CN8*, *CN12* and *CN15*) in *M. sinensis*. As florigen that exists in all species acts as a floral integrator, transducing input from multiple pathways into one signal that acts in the meristem. Therefore, it is likely that *MsiCN8*, *MsiCN12*, *MsiCN15* are transcriptionally regulated by *MsiEhd1* dependent pathway but also by other pathways.

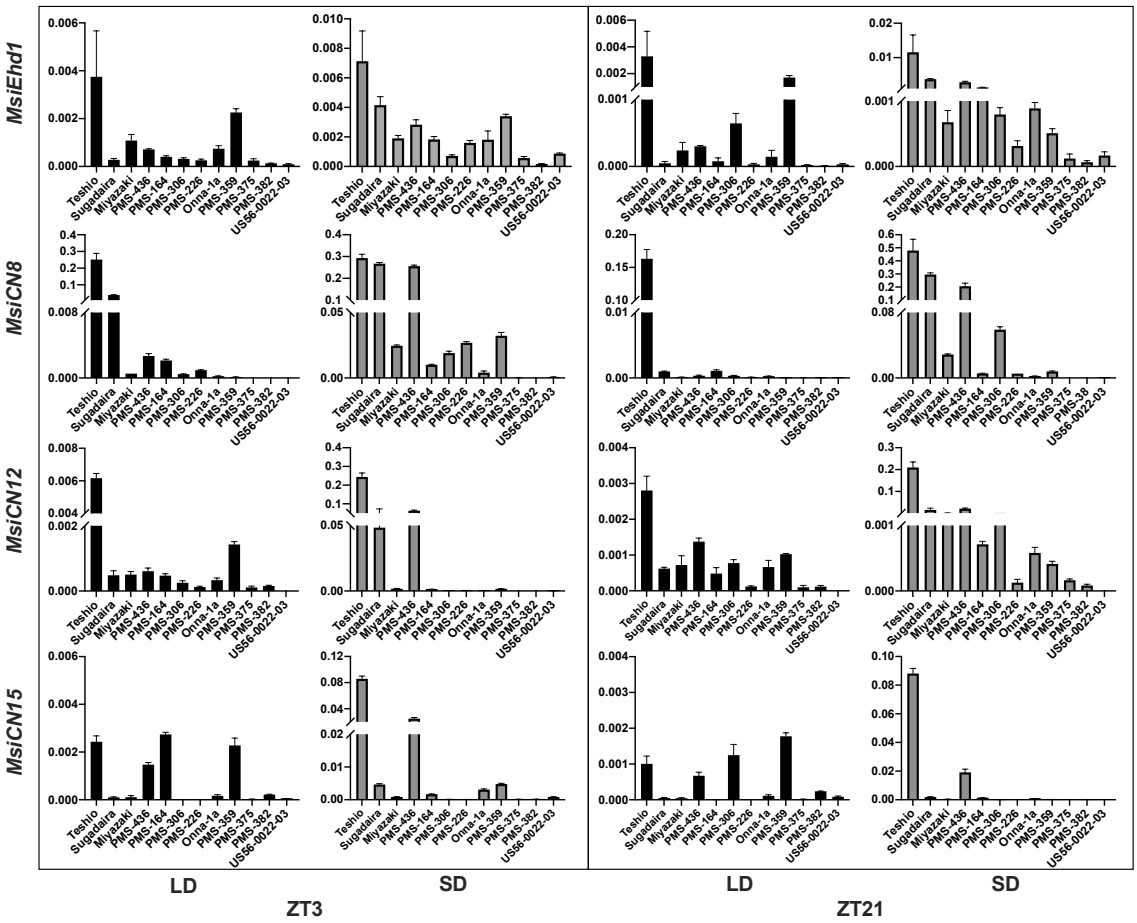


Figure 4.7 Expression levels of *MsiEhd1*, *MsiCN8*, *MsiCN12* and *MsiCN15* under long days (LD, dark bar chart) and short days (SD, grey bar chart) at Zeitgeber times (ZT) 3 and 21 of the day. The *x*-axis indicates accessions' name and *y*-axis represents relative mRNA levels normalized to *ACTIN* transcript levels. The white and black bar at the bottom of each graph indicate the light and dark period, respectively. Mean ± 1SE for three replications are given for each data point.

4.3.4 *FTs* activation under short days

Under SD condition, rice *Hd1*/*CO* could bind to the promoter region of *Hd3a* (and possibly also to that of *RFT1*) through Ghd8-OsHAP5-Hd1 or Hd1-OsHAL3 complex for activating the expression of *Hd3a* (possibly *RFT1*) (Xue *et al.*, 2008; Su *et al.*, 2016; Wang *et al.*, 2019*a*). Therefore, the gene expression patterns of *Ghd8*, *Hd1* and *FTs* in *M. sinensis* under SD condition were compared. The expressions of *MsiGhd8* were observed at night period (ZT15 and ZT21) in SD (Figures 2.4 and 4.8). And the mRNA transcript of *MsiHd1*/*CO* was also abundant in night period (ZT15 or ZT21) (Figures 4.1 and 4.8). In this case, there is a probability that Ghd8 and Hd1 form a complex to induce *FTs* (*CN8*, *CN12* and *CN15*) expressions in *M. sinensis*. Though three *FTs* expression patterns in SD varied among twelve accessions (Figures 4.4, 4.5, 4.6 and 4.8), the relative higher expressions of at least one *FT* were observed at morning ZT3 or ZT21 for some accessions, suggesting the potential ability that Ghd8-Hd1 complex activated *FTs* expressions in *M. sinensis*. While, this interaction in *M. sinensis* will require further verification. In sorghum, *Hd1*/*CO* also induced *SbCN8* and *SbCN12* to promote flowering under SD (Yang *et al.*, 2014*b*). In the dark period of SD, *FTs* also expressed in ZT15 for most accessions, suggesting it might be induced by *MsiHd1*, showed in *M. sinensis* showed the peaks in the Considering the above facts, the results of gene expression profile models in present study provide possible regulation models that *FTs* might be activated by *Hd1* or Ghd8-Hd1 complex in *M. sinensis* accessions under SD condition. Competition or cooperation of regulation models may occur; however, it should be verified in further experiments.

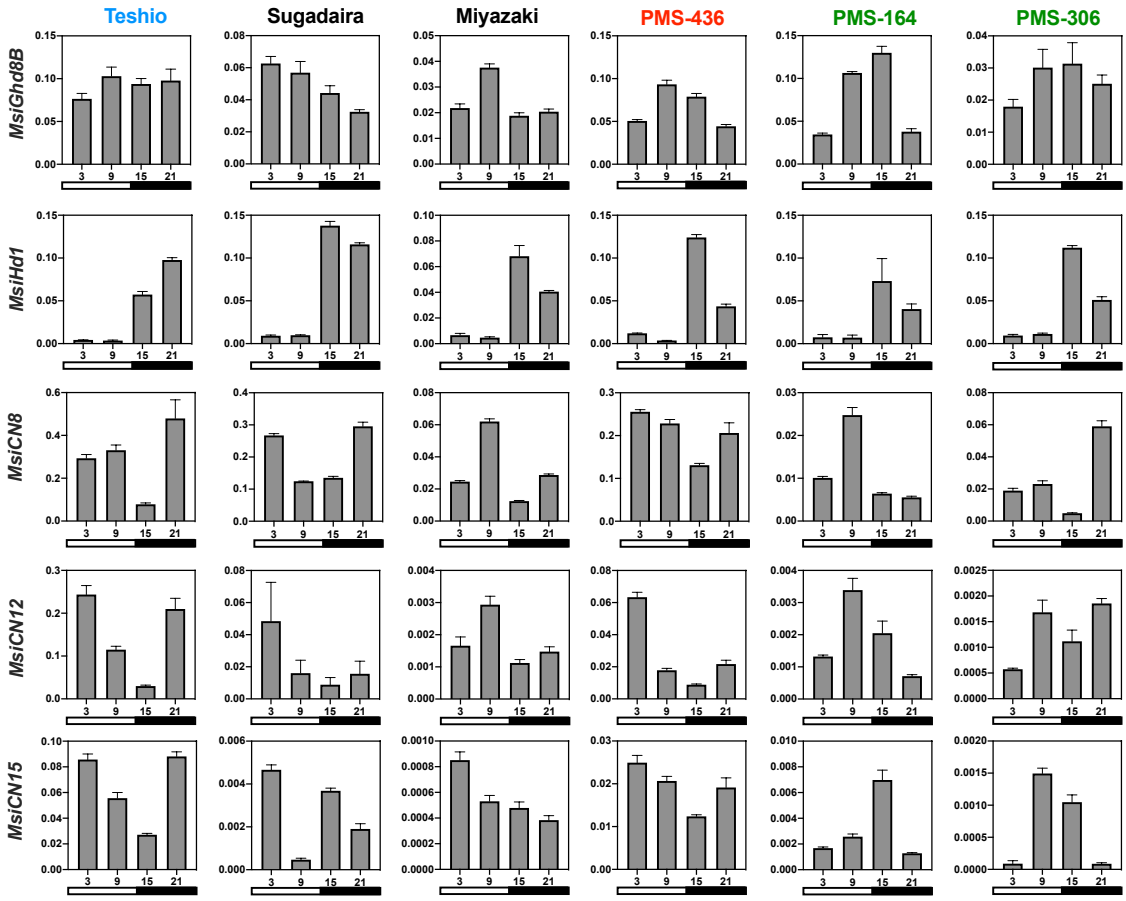


Figure 4.8 Expression patterns of *MsiGhd8B*, *MsiHd1*, *MsiCN8, MsiCN12, MsiCN15* under short days (12.5h, SD). The *x*-axis indicates Zeitgeber times (ZT) of the day and *y*-axis represents relative mRNA levels normalized to *ACTIN* transcript levels. The white and black bar at the bottom of each graph indicate the light and dark period, respectively. Accessions' names are colored to represent *Miscanthus sinensis* genetic groups previously described by Clark *et al*. (2014, 2015); Sugadaira and Miyazaki are changed from yellow to black for making it clear. **(Continued)**

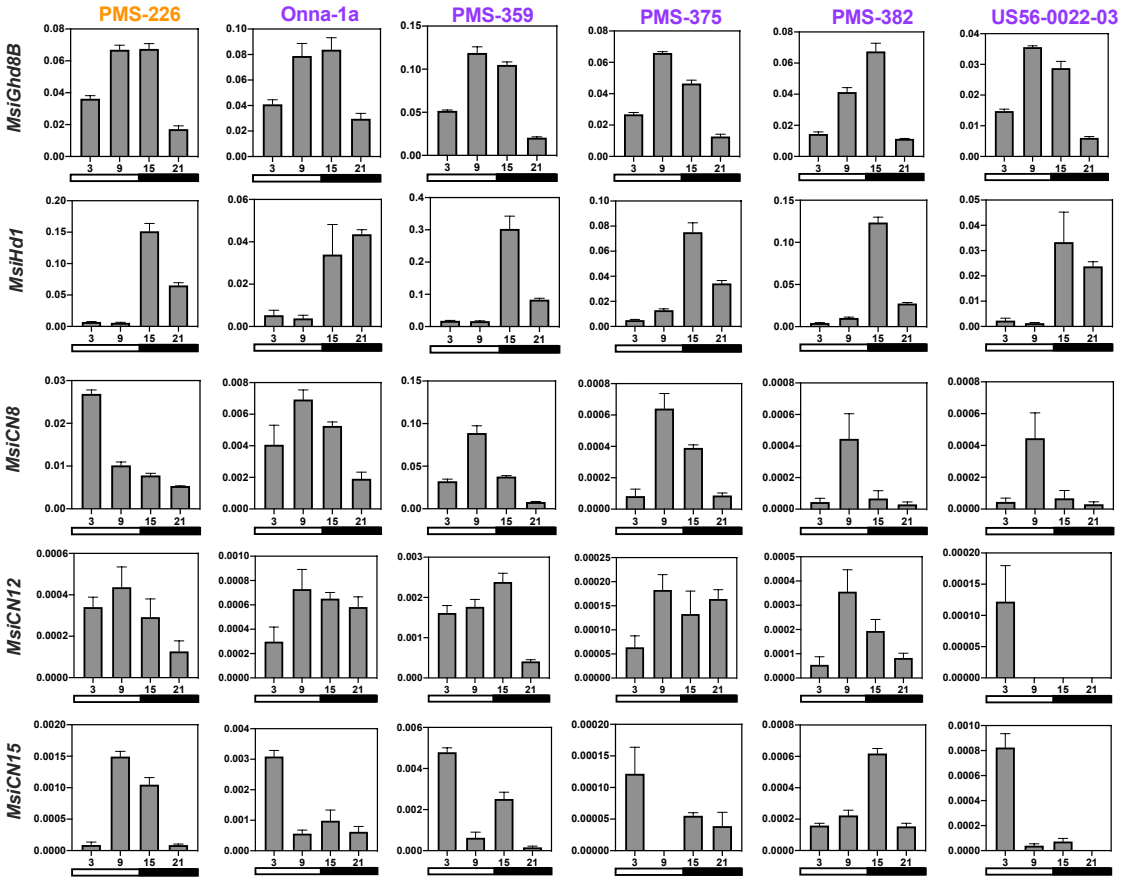


Figure 4.8 Expression patterns of *MsiGhd8B*, *MsiHd1*, *MsiCN8, MsiCN12, MsiCN15* under short days (12.5h, SD). The *x*-axis indicates Zeitgeber times (ZT) of the day and *y*-axis represents relative mRNA levels normalized to *ACTIN* transcript levels. The white and black bar at the bottom of each graph indicate the light and dark period, respectively. Accessions' names are colored to represent *Miscanthus sinensis* genetic groups previously described by Clark *et al*. (2014, 2015); Sugadaira and Miyazaki are changed from yellow to black for making it clear.

4.3.5 The relationship between florigens and flowering time in *M. sinensis*

The critical signal for the transition to flowering was the formation of *FTs*, which were regulated by multiple upstream genes (Fornara *et al.*, 2010), therefore, the relationship between *FTs* expressions and days to first flower in response to the photoperiod was evaluated in *M. sinensis* (Figure 4.9). Four *M. sinensis* accessions (Miyazaki, PMS-306, PMS-226 and PMS-359) had a relative higher expression of *MsiCN8* under SD condition, and some of them also showed a higher mRNA transcript level of *MsiCN12* or *MsiCN15* in SD, consistent with their earlier flowering in SD relative to LD (Figure 4.9). Three accessions (PMS-375, PMS-382 and US56-0022-03) originated from the tropics could flower in SD, later than other entries (Figure 4.11), but displayed relatively higher expression of three *FTs* compared to that when they grew in LD (Figure 4.9). These results indicated that these florigens promote flowering in *M. sinensis*. Interestingly, when plants grown under LD condition, significant differences in expressions of *MsiCN8*, *MsiCN12* and *MsiCN15* were observed in Teshio and Onna-1a. Teshio (44.9 °N) and Onna-1a (26.5 °N) exhibited opposite flowering time for earliest flowering (66 days) and late flowering (274 days), respectively, as expected, Teshio showed significantly higher expression of three *FTs* relative to Onna-1a, especially *MsiCN8* and *MsiCN12* (Figure 4.9). Three accessions (PMS-375, PMS-382 and US56-0022-03) originated from the tropics (23.5 °S to 23.5 °N) failed to flower in LD, and presented similar lower transcript levels of *MsiCN8*, *MsiCN12* and *MsiCN15* compared to other entries (Figure 4.10). These results indicated that *FT*s were positively associate with days to first flower in *M. sinensis*.

Under SD, five *M. sinensis* genotypes (Teshio, Sugadaira, PMS-436, PMS-164, Onna-1a) were unable to flower, all but Onna-1a had a high expression level of *MsiCN8*, *MsiCN12* and *MsiCN15*, opposite with flowering date (Figures 4.9 and 4.11). While these four accessions originated from high latitudes (≥ 36 °N), Dong *et al*. (2021) reported that SD was also a signal for *M. sinensis* from high latitude to induce a short-internode dormancy response, which is an adaptation to protect apical meristems from damaging low temperatures during winter in high latitudes, and this dormancy response was epistatic to flowering. Similar dormancy responses to SD have been found in several quantitative SD, perennial, C_4 grasses, including *M. sacchariflorus* (Jensen *et al.*, 2013), switchgrass (Castro *et al.*, 2011) and big bluestem (McMillan, 1959). Besides, DNA methylation is an epigenetic modification and important to many biological processes in plants (Zhang *et al.*, 2018). Saad *et al*. (2019) found photoperiod could induce genotype-specific shift in DNA methylation in Tartary buckwheat by genome-wide DNA methylation analysis. Therefore, day length might affect changes of DNA methylation or some unknown substance for the above four accessions. In general, *MsiCN8*, *MsiCN12* and *MsiCN15* were correlated to flowering date of *M. sinensis* for most accessions and could promote flowering. Given that the expression profile of *MsiCN8*, *MsiCN12* and *MsiCN15* was diverse among each accession, it might because the studied materials are wild collections and the genetic background is complex.

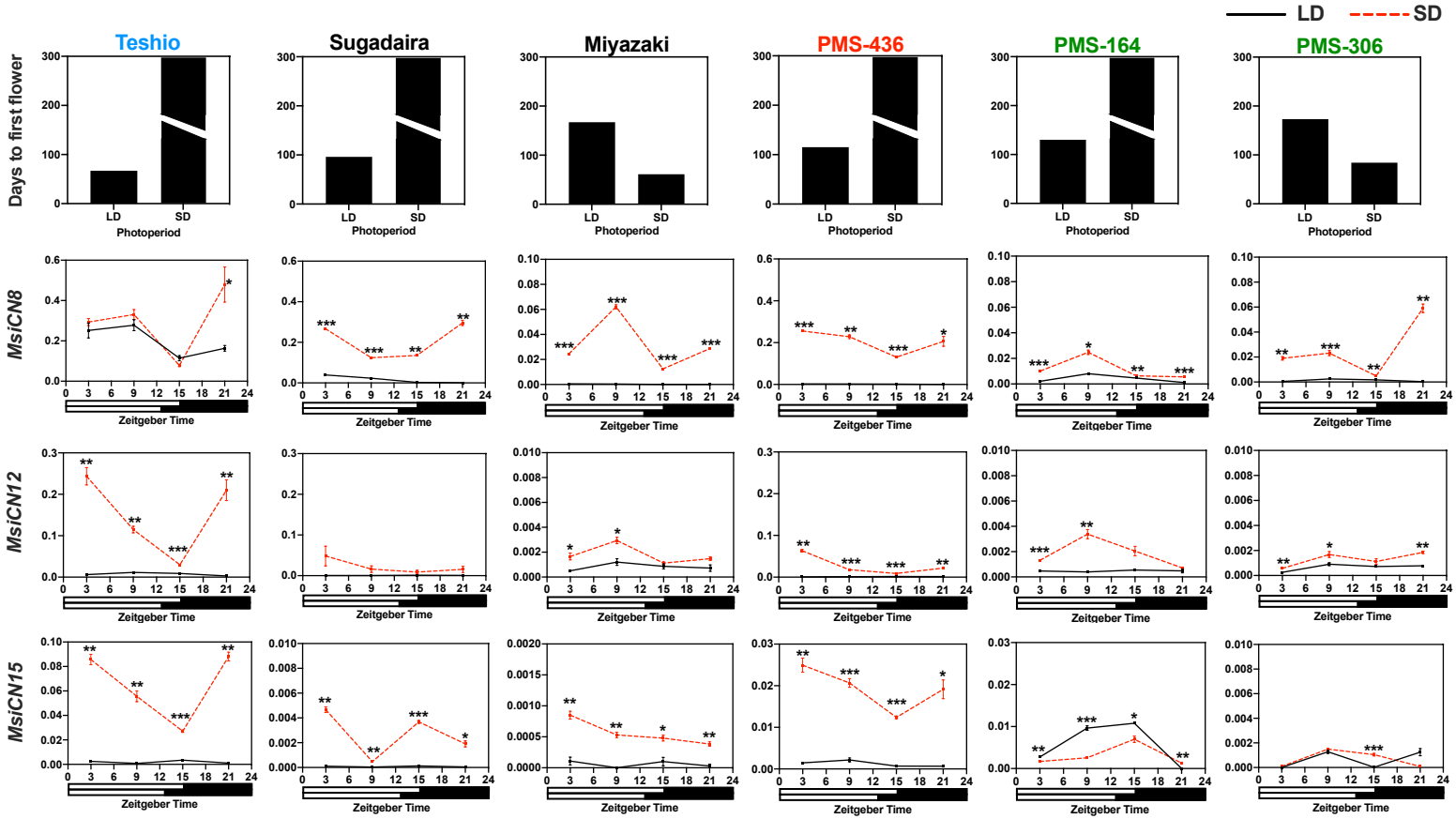


Figure 4.9 Days to first flower, and expression of *MsiCN8*, *MsiCN12* and *MsiCN15* in 12 *Miscanthus sinensis* accessions under long days (LD) and short days (SD). Break bar box represents non-flowering under LD or SD. The *x*-axis indicates Zeitgeber times (ZT) of the day and *y*-axis represents relative mRNA levels normalized to *ACTIN* transcript levels. Mean ± 1SE for three replications are given for each data point. Asterisks indicate significant difference between the two means under LD and SD at the same ZT of the day (Student's *t*-test, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). No asterisk indicates the difference between the two means is not statistically significant ($P > 0.05$). Accessions' names are colored to represent *M. sinensis* genetic groups previously described by Clark *et al.* (2014, 2015); Sugadaira and Miyazaki are changed from yellow to black for making it clear. **(Continued)**

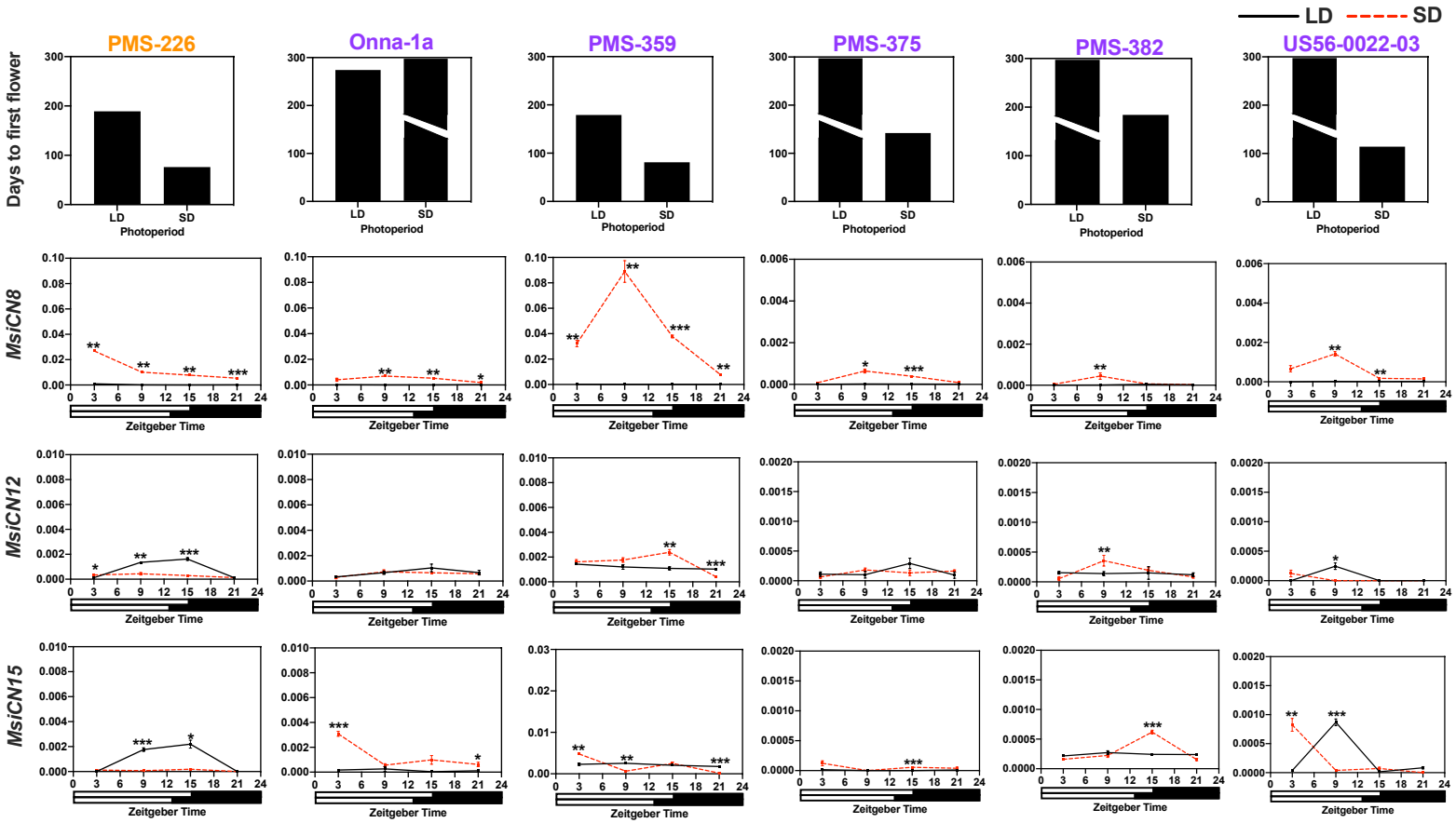


Figure 4.9 Days to first flower, and expression of *MsiCN8*, *MsiCN12* and *MsiCN15* in 12 *Miscanthus sinensis* accessions under long days (LD) and short days (SD). Break bar box represents non-flowering under LD or SD. The *x*-axis indicates Zeitgeber times (ZT) of the day and *y*-axis represents relative mRNA levels normalized to *ACTIN* transcript levels. Mean ± 1SE for three replications are given for each data point. Asterisks indicate significant difference between the two means under LD and SD at the same ZT of the day (Student's *t*-test, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). No asterisk indicates the difference between the two means is not statistically significant ($P > 0.05$). Accessions' names are colored to represent *M. sinensis* genetic groups previously described by Clark *et al.* (2014, 2015); Sugadaira and Miyazaki are changed from yellow to black for making it clear.

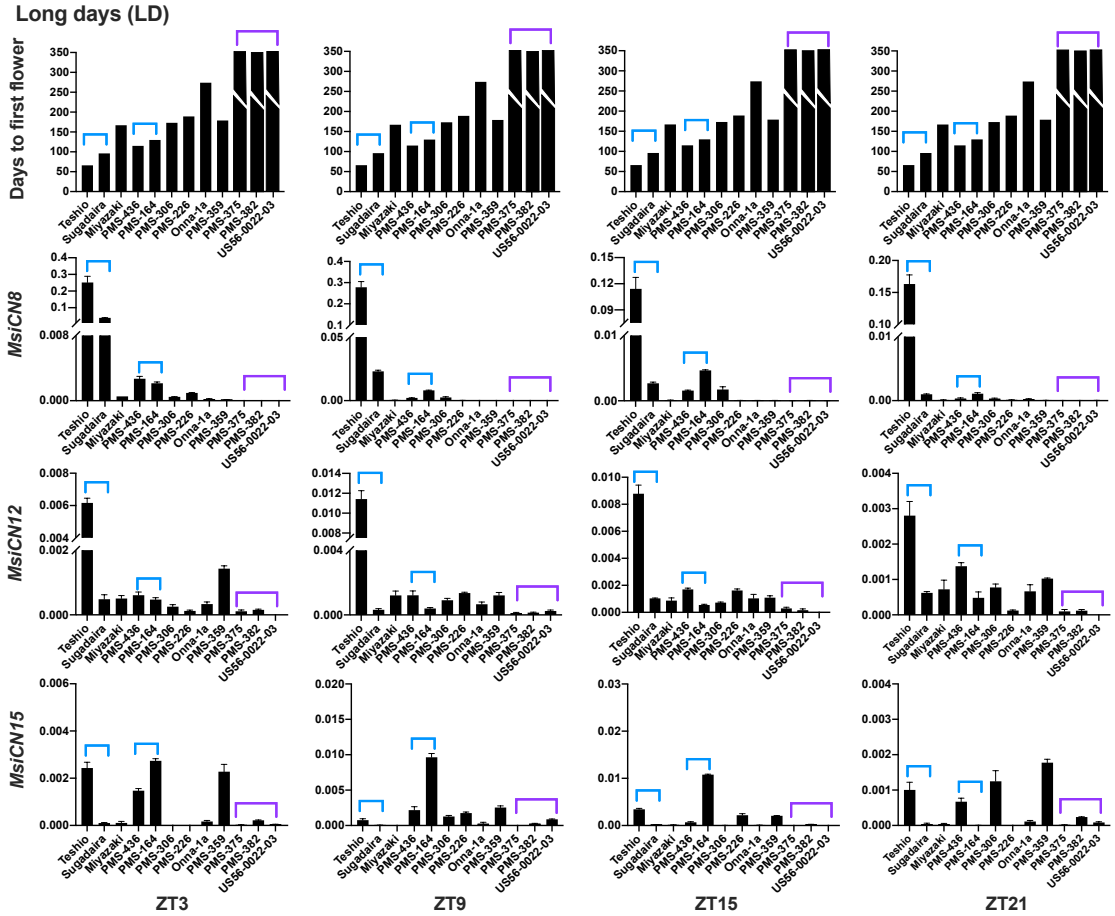


Figure 4.10 Comparison between days to first flower and expression profiles of *MsiCN8*, *MsiCN12* and *MsiCN15* in 12 *Miscanthus sinensis* accessions at Zeitgeber times (ZT) of the day under long days (LD). Break bar box represents non-flowering under LD. The *y*-axis represents relative mRNA levels normalized to *ACTIN* transcript levels. Mean ± 1SE for three replications are given for each data point.

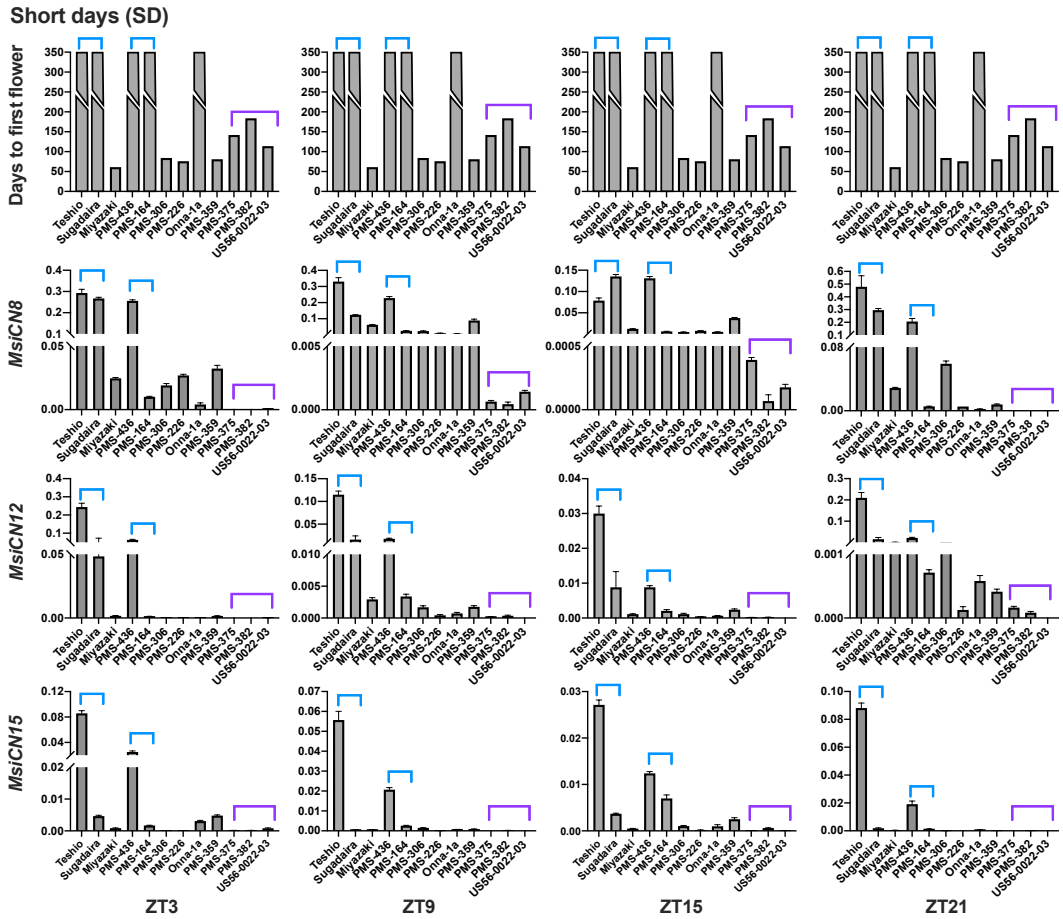


Figure 4.11 Comparison between days to first flower and expression profiles of *MsiCN8*, *MsiCN12* and *MsiCN15* in 12 *Miscanthus sinensis* accessions at Zeitgeber times (ZT) of the day under short days (SD). Break bar box represents non-flowering under SD. The *y*-axis represents relative mRNA levels normalized to *ACTIN* transcript levels. Mean ± 1SE for three replications are given for each data point.

In conclusion, generally, for most accessions, the expression patterns of candidate flowering genes were consistent with rice and sorghum. Through comparisons one by one, the flowering pathway *Ghd8*/*Hd1*-*Ghd7*-*Ehd1*-*CN8*/*CN12*/*CN15* might exist in some accessions of *M. sinensis* under LD condition. Some accessions failed to flower in SD, indicating more complicated regulation model of *M. sinensis* in SD. Under SD condition, the flowering regulation model *Ghd8*/*Hd1*-*Ehd1*-*CN8*/*CN12*/*CN15* might occur in some *M. sinensis* accessions, and *CN8*/*CN12*/*CN15* might also be induced by *Hd1* directly for partial accessions. Though protein-protein interaction hasn't been confirmed yet, the gene expression profile of the present research provided available functional information to support the roles of these genes, and will underpin future work on the control of flowering in *M. sinensis* by guiding experiments on responses to induction and modification of critical genes.

Supplementary materials for Chapter 4:

Supplementary Table S4.1: Seven main flowering time genes *Miscanthus sinensis* identified in *Sorghum bicolor*, *Oryza sativa* and *Zea mays*.

Supplementary Table S4.2: List of primers for quantitative real time (qRT)-PCR used in the present study.

Table S4.1 Seven main flowering time genes *Miscanthus sinensis* identified in *Sorghum bicolor*, *Oryza sativa* and *Zea mays*

Gene	Locus ID in *Miscanthus sinensis**	Location	Locus ID in *Sorghum bicolor*	Location	Locus ID Gene in *Oryza sativa*	Location	Locus ID Gene in Zea mays	Location
Ghd8/*Hd5*/ *DHT8*	Misin13G040800/ Misin07G127500	Chr13: 10535222..10536050/ Chr07: 26826120..26827213	Sobic.007G059500	Chr07: 6219129.. 6220244	LOC_Os08g07740	Chr8: 4333716.. 4335434	Zm00001d049485	Chr04: 31972820.. 31973600
Ghd7/*Hd4*/ *CCT*	Misin11G003800/not assembled on Chr12 in MsB subgenome	Chr11: 1998112..2000968/ Chr12: 1953905.. 1952462; 1804245.. 1803910; scaffold01511 (HSPs)	Sobic.006G004400	Chr06: 697458.. 700101	LOC_Os07g15770	Chr07: 9152401.. 9155185	Zm00001d024909	Chr10: 94430849.. 94433495
Hd1/*CO*	Misin18G123200/ Misin19G121300	Chr18: 32024942..32026959/ Chr19: 31343989..31346736	Sobic.010G115800	Chr10: 12353900.. 12355900	LOC_Os06g16370	Chr06: 9336358.. 9338643	Zm00001d045735	Chr09: 36009334.. 36013889
Ehd1	Misin01G211100/ Misin02G206700	Chr01: 47390154..47397173/ Chr02: 48487164..48488408	Sobic.001G227900	Chr01: 21860029.. 21867056	LOC_Os06g08440	Chr06: 4137247.. 4142663	Zm00001d032784	Chr01: 238041003..238043016
CN8/ *FTL10*	Misin17G201500/ Misin16G202300	Chr17: 71278774..71280708/ Chr16: 69633845..69636144	Sobic.009G199900	Chr09: 54961461.. 54963786	LOC_Os05g44180	Chr05: 25667454.. 25669369	Zm00001d010752	Chr08: 126880530..126882389
CN12/ *FTL9*	Not assembled on Chr05 in MsA subgenome/ Misin06G261800	Chr06: 90956954..90958833	Sobic.003G295300	Chr03: 62747945.. 62749919	LOC_Os01g54490	Chr01: 31343363.. 31345522	Zm00001d043461	Chr03: 200703154..200704705
CN15/*FTL3*/*Hd3a*	Misin18G047800/ Misin19G047400	Chr18: 10052573..10055909/ Chr19: 10788531..10791755	Sobic.010G045100	Chr10: 3499964.. 3502278	LOC_Os06g06300	Chr06: 2926822.. 2928474	Zm00001d036242	Chr06: 79177906.. 79179890

*Gene Locus IDs in *Miscanthus sinensis*, *Sorghum bicolor*, *Oryza sativa* and *Zea mays* are from Phytozome v13 (https://phytozome-next.jgi.doe.gov). Some homoeologous loci in *M. sinensis* subgenome are still not assembled, only the location of high-scoring sequencing pairs (HSPs) is showed through the sequence blast.

Table S4.2 List of primers for quantitative real time (qRT)-PCR used in the present study

Gene name	Forward primers (5' → 3')	Reverse primers (5 ' → 3')	Amplified products (bp)
MsiGhd7A	TCAAAGAGACAACCCTGACCGACGA	GGTTACCTTAGCAAAGCGGCCTC	194
MsiGhd7B	TCAAGGAGCCAACCCTGACCGATGG	TCGGTTACCTTGGCAAAGCGGCCTT	196
MsiGhd8A	CTCAACCGCTACCGCGAGGTC	TCATCCGCCGCGCCATCT	89
MsiGhd8B	ACGTCGGGCTCATGATGGGAGCA	ATACGACTTCCGTGCTGCCGT	95
Hd1/CO	ATCAGCCTCTTCTCGTCAGGT	TGCTTCTGCATATGTCTTCCTC	156
Ehd1	GCTCAACACCTCGATCAGGTTTC	GTATATCTGTGACCGTCCCGTTAG	151
CN8	GTGTCAACTTTGGCCAAGAGCTAG	CGATCCACCTTCCCTTTGACAGTT	133
CN12	CACCGCATGGTATTTGTGCTGTTC	GAAATATGTGGCAGCCACAATG	132
CN15	CAGAACTTCAACACCAGGGACTT	TCACGCTGGCAGTTGAAGTAGAC	83
Actin	AGGGCTGTTTTCCCTAGCATCGT	GGGTACTTGAGCGTGAGAATACCTC	128

Chapter 5

General discussion

The optimization of flowering time is likely to improve biomass quantity and quality and therefore is an important factor in the domestication of *Miscanthus* into a bioenergy grass crop. Controlling flowering will facilitate the hybridization of genotypes from diverse geographical locations, but also assist the intergeneric crosses, such as *Miscanthus* and *Saccharum*. Synchronizing flowering time will also be crucial for the development of a seed-propagated crop. *M. sinensis* is identified as a warm-season grass, typically facultative SD plant (Dong *et al.*, 2021). The major genes of the pathways that contribute to floral induction in *A. thaliana* and rice are well known (Fornara *et al.*, 2010; Kim, 2020; Wei *et al.*, 2020). Studies in *A. thaliana*, rice, maize, sorghum and other grasses have provided a framework by which to compare *Miscanthus* flowering genes. As a potential bioenergy crop, it is important to determine how *M. sinensis* commits itself to the reproductive phase in order to optimize breeding efforts and obtain high biomass accumulation, as well as provide more information about flowering in facultative SD perennial plants.

5.1 Homoeologous loci of *Ghd8* and *Ghd7* at each *M. sinensis*' subgenome

The results of the current study demonstrate that *Ghd8* and *Ghd7* are present in *M. sinensis*, and likely contribute to a regulatory function for flowering time in this species in a similar manner to that in rice. Firstly, collinearity analysis revealed that two homoeologous *Ghd8* loci (Misin13G040800 and Misin07G127500) (Figure 2.1), one each in the two *M. sinensis* subgenomes (MsA and MsB), corresponding

to the same genomic region on rice Chr.08 (LOC_Os08g07740) and sorghum Chr.07 (Sobic.007g059500) (Figure S2.1 and Table S2.5), which was consistent with the known paleo-duplications (rice Chr.08- sorghum Chr.07, sorghum Chr.07-*Miscanthus* Chr.13 and Chr.07) from the ancestral grass chromosomal groups (Salse *et al.*, 2009; Dong *et al.*, 2018; Mitros *et al.*, 2020). Similar to *MsiGhd8*, *Ghd7* homoeologous loci were identified in *M. sinensis*, with one on each of this paleo-allotetraploid species' subgenomes (Figure 3.1). As multiple alleles were detected from *MsiGhd7B* in the B subgenome, and the complete sequence of *MsiGhd7B* has not yet been assembled in the database of *Miscanthus sinensis* v13 (Mitros *et al.*, 2020), it indicated that small-scale gene duplications might occur the B subgenome during the evolution. Small-scale gene duplications have been confirmed for complementary functions for whole genome duplication, and provides robust mutations as a selective advantage for adapting to a large range of environments (Fares *et al.*, 2013; Glover *et al.*, 2015).

Additionally, at each of the two homoeologous *Ghd8* loci in *M. sinensis*, each accession in this study had at least one putatively functional full-length allelic copy containing a highly conserved HAP3/NF-YB DNA-binding domain that is required for the transcription factor function of *Ghd8* in rice (Wei *et al.*, 2010) and *A. thaliana* (Kumimoto *et al.*, 2008), and *MsiGhd8* may have a conserved function to regulate flowering time. Moreover, the two homoeologous *Ghd8* loci in *M. sinensis* expressed under LD and SD conditions, but a significant differentiation of mRNA transcript level was observed with higher expression of *MsiGhd8B* for each accession (Figures 2.4 and 2.5), which was consistent with a previously observation with ~10% more pairs of genes having higher expression in the *M. sinensis* B

subgenome (Mitros *et al.*, 2020). *MsiGhd7* homoeologous loci contained the conserved CCT protein domain, which have been reported to have crucial roles in regulating processes such as photoperiodic flowering (Putterill *et al.*, 1995; Yano *et al.*, 2000; Turner *et al.*, 2005; Hung *et al.*, 2012; Murphy *et al.*, 2014; Yang *et al.*, 2014*b*), vernalization (Yan *et al.*, 2004), circadian rhythms (Strayer *et al.*, 2000; Salomé *et al.*, 2006), and light signaling (Kaczorowski and Quail, 2003; Murphy *et al.*, 2011), indicating conservative function on the flowering regulation in *M. sinensis*. Furthermore, expression of *MsiGhd7* largely decreased (Figure 3.6) in SD, in similar to rice, sorghum and other grasses (Xue *et al.*, 2008; Murphy *et al.*, 2014; Glassop and Rae, 2019; Weng *et al.*, 2019).

5.2 Allele and amino acid sequence diversity of *MsiGhd8* and *MsiGhd7*

Allelic variation in genes for flowering time is a major driver of environmental adaptability. Considering that the nucleotide diversity in the coding region cannot precisely represent the protein diversity owing to SNVs in exons, allelic and deduced protein sequence diversity were analyzed for *Ghd8* and *Ghd7* in the present study. Accounting for nsSNVs, several alleles and predicted amino acid sequence variants of *MsiGhd8* and *MsiGhd7* showed a geographic and latitudinal distribution (Figures 2.3, S2.3 and 3.5). Though no high frequency of non-functional alleles was been observed among studied accessions, four nsSNVs identified in the HAP3/NF-YB domain of *MsiGhd*8 from five accessions (Tables S2.1 and S2.2) may have an important effect on protein stability and function. One putative non-function allele identified in *MsiGhd7B* was found in PMS-164 from Northern China and was characterized by an eight-base insertion in the first exon, resulting in a frameshift

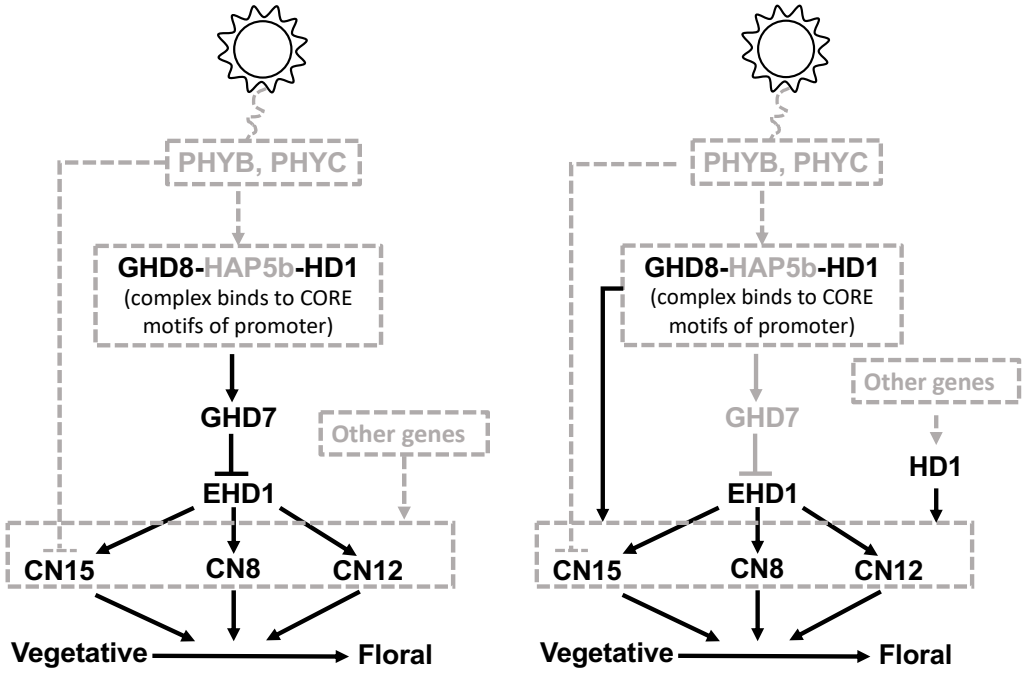
and eventual premature termination of the protein, totally lacking CCT domain (Figure 3.1). Mutations in the SD plant rice and sorghum flowering genes were critical for adaptability in LD ecosystems during early human migration or trade, and production of grain and energy hybrid crops (Rooney *et al.*, 2007; Xue *et al.*, 2008; Murphy *et al.*, 2014; Klein *et al.*, 2015; Zhang *et al.*, 2015). In rice, *Ghd8* with 19 bp deletion causing a loss-of-function, has been selected and used widely for breeding early heading varieties in Hokkaido, the northernmost region of Japan and one of the northern limits of rice cultivation in the world (Fujino *et al.*, 2013). Recently, a novel non-functional *Ghd7* allele, resulting from 12 bp insertion in the upstream from the transcription start site, was distinctly distributed in rice varieties from Northern Japan (Fujino and Yamanouchi, 2020) but not be detected in the cultivated rice from other regions (Xue *et al.*, 2008; Lu *et al.*, 2012). As the sample size of the present studies was limited, allelic variants of heading date genes were identified in the unique accession, but this mutation might also contribute to domestication and expansion of *M. sinensis* growing region. Further research is needed to quantify the effects of individual putative functional *MsiGhd8* and *MsiGhd7* alleles with nsSNVs and/or sSNVs, and nonfunctional *MsiGhd7B* alleles on flowering time in response to day length in a large *Miscanthus* population. These studies can evaluate segregating populations derived from controlled biparental crosses, or be achieved by gene editing. The current study provides information on which alleles are present in different accessions that can be used to conduct genetics studies of segregating biparental populations. Additionally, the sequence data obtained by the current study for many different natural *MsiGhd8* and *MsiGhd7*

alleles can be used to plan gene-editing studies in *Miscanthus*, rice or other species to dissect function while controlling for genetic background.

5.3 Flowering regulation model in *M. sinensis*

The flowering is controlled by a complex genetic regulatory network rather than one gene. In *A. thaliana*, HAP3b subunits can directly interact with Hd1/CO through its CCT-domain, forming CCAAT-binding CBF-complexes that bind to *FT* promoters and activate its transcription to promote flowering in LD (Ben-Naim *et al.*, 2006; Wenkel *et al.*, 2006; Lv *et al.*, 2021). Recent studies revealed the interaction among *Ghd7*, *Ghd8* and *Hd1* in rice (Wang *et al.*, 2019*a*; Zong *et al.*, 2021). The grass-specific gene *Ghd7* is upregulated by a Ghd8-OsHAP5b-Hd1 complex in LD, enabling *Ghd7* to suppress *Ehd1* and delay flowering (Li *et al.*, 2015; Zhang *et al.*, 2015, 2019; Nemoto *et al.*, 2016; Du *et al.*, 2017; Liu *et al.*, 2020; Zong *et al.*, 2021). The current study of gene expression pattern suggested that MsiGhd8 might interact with MsiHd1/CO to form the complex and then activated the transcription of *MsiGhd7* as described in rice (Wang *et al.*, 2019*a*; Zong *et al.*, 2021). Moreover, expression profiles of five candidate flowering genes (*Hd1*, *Ehd1*, *CN8*, *CN12* and *CN15*) in *M. sinensis* were evaluated (Figures 4.1, 4.3 to 4.6) and revealed high similarity to the observation in rice and sorghum (Tamaki *et al.*, 2007; Xue *et al.*, 2008; Murphy *et al.*, 2014; Yang *et al.*, 2014*b*). Under LD condition, gene expression patterns of most accessions proposed the possibility that Ghd8-Hd1 complex activated *Ghd7* expression (Figure 4.2), subsequently largely inhibited *Ehd1* and then downregulated *CN8*/*CN12*/*CN15* expression, and delayed flowering in *M. sinensis* (Figures 4.3, 4.4 and 5.1). In SD condition, the expression

of *Ghd7* dramatically decreased (Figures 3.6 and 4.3). The latest studies showed that Hd1 might form a complex with NF-YB/YC to recognize the core elements of *Hd3a*/*FT* promoter (Goretti *et al.*, 2017; Lv *et al.*, 2021). In sorghum, *Hd1*/*CO* also could induce the expression of *SbEhd1* and *SbCN8* and *SbCN12* (Yang *et al.*, 2014*b*). Both *MsiGhd8* and *MsiHd1* expressed under LD and SD, raising a possibility that *FTs* were regulated by *Hd1* directly or Ghd8-Hd1 complex in SD (Figures 4.8 and 5.1). Generally, present results on the mRNA transcript level suggested that these main flowering genetic regulatory models in rice are likely existed in *M. sinensis*, further studies on the interaction between the genes (protein-DNA and protein-protein) are needed to verify possible models and fully elucidate the flowering regulatory network in *M. sinensis*.



Long days (LD) Short days (SD)

Figure 5.1 Simplified model of flowering-time regulation in *Miscanthus sinensis* under long days (LD) and short days (SD). In LD, the heterotrimer GHD8-OsHAP5b-HD1 complex targets the CORE motif of the *Ghd7* promoter to activate its expression, leading to suppression of *Ehd1* and downregulation of *CN8*, *CN12* and *CN15* expression thereby inhibit flowering. In SD, the GHD8-OsHAP5b-HD1 complex may directly bind to the promoter of *FTs* (*CN8*, *CN12* and *CN15*) to floral induction. *Hd1* may activates expression of *FTs* (*CN8*, *CN12* and *CN15*). Solid black lines indicate the transcriptional activation/ repression based on the present research. Genes' names and dashed arrow or lines colored grey indicates it has not been decided in *M. sinensis*.

In addition, under SD condition, many *M. sinensis* accessions from high latitude could not flower, the mechanism should be addressed. One explanation is that SD is a signal for dormancy response, which was epistatic to flowering. Dong *et al*. (2021) observed that relatively short days could accelerate floral induction of *M. sinensis*, but below a critical threshold, especially for genotypes adapted to high latitudes, SD can signal that plants should prepare for winter, and importantly this response is epistatic to flowering. Similar dormancy responses to SD have been found in several quantitative SD, perennial, C_4 grasses, including *M. sacchariflorus* (Jensen *et al.*, 2013), switchgrass (Castro *et al.*, 2011) and big bluestem (McMillan, 1959). Furthermore, day length might affect accumulation of DNA methylation for these accessions originated from high latitude. The feature that photoperiod could induce genotype-specific shift in DNA methylation was identified in Tartary buckwheat (Saad *et al.*, 2019). The other possible explanation is that some flowering genes' interactions unidentified in *M. sinensis* help counteract severely delayed flowering that arose by *Ghd8*, *Ghd7* and *Hd1*.

In summary, the current study identified the *Ghd8* and *Ghd7* homoeologous loci in *M. sinensis*, with at least one on each of this paleo-allotetraploid species' subgenomes. At least two *MsiGhd7B* loci were identified in the B subgenome. The small-scale duplication may occur in the B subgenome. Gene expression analysis provides a better understanding of how *Ghd8*, *Ghd7* and up- and downstream genes regulate flowering time response to photoperiod in *M. sinensis*. Possible flowering regulation pathway in *M. sinensis* were raised (Figure 5.1). The present study adds functional information to elucidate the potential roles of these genes in regulating flowering time and will make a foundation for future work on the control of

flowering in *Miscanthus*. Predicting the flowering time of a given genotype in a given environment would allow breeders to select promising varieties from a very early stage, allowing for much more rapid improvement of *Miscanthus* traits, which is beneficial for energy security and sustainability.

5.4 Future directions

The genetic control of photoperiod-induced flowering has been well documented in other species like *A. thaliana*, rice and sorghum, illustrating the high degree of conservation within this pathway and providing an excellent starting place for elucidating the *Miscanthus* flowering pathway. Due to the interaction among Ghd8, Hd1 and Ghd7, it would be worthwhile to investigate this function on the protein level in *M. sinensis* through biochemistry assay. Furthermore, recent studies found that Ghd7 protein stability and function were affected by phytochromes and rice GI (Zheng *et al.*, 2019), and genetic linkage maps in *M. sinensis* revealed two flowering time QTLs (Gifford *et al.*, 2015; Dong *et al.*, 2018) that corresponded to *SbPhyB* (Yang *et al.*, 2014*a*) and *PhyC*, respectively, characterization of these genes would be valuable to complement networks on the flowering regulation of *M. sinensis*.

M. sinensis is considered as a quantitative SD plant (Dong *et al.*, 2021), and flowering time in *M. sinensis* were also affected by multiple factors, including degree days, temperature, photoperiod and precipitation (Jensen *et al.*, 2011). The questions are raised whether the impact that degree days or dormancy signal, in conjunction with photoperiod, may have effects on regulating flowering time in *M. sinensis*, and whether the expression of *Ghd8*, *Ghd7* or other flowering genes might

be modified by these factors, resulting in the differentiation of flowering time? Moreover, the genus *Miscanthus* reveals great photosynthetic efficiency, high biomass yield capacity, low input demands and good tolerance of diverse climate. Except for the control of heading date and morphogenesis in rice, *Ghd7* and *Ghd8* also involved in response to abiotic stress (Du *et al.*, 2018; Wang *et al.*, 2019*b*; Alam *et al.*, 2020). Abiotic stresses also affect the transcription of florigen genes *Hd3a* and *RFT1* (Cho *et al.*, 2017). However, the integration of responses to abiotic stresses and heading date has not been yet analyzed in *Miscanthus*. The elucidation of the intricate mechanism of heading date control in the presence of abiotic stresses, especially salinity and seasonal drought stress, would significantly benefit *Miscanthus* molecular design breeding in the future.

References

Abdul-Awal SM, Chen J, Xin Z, Harmon FG. 2020. A sorghum *gigantea* mutant attenuates florigen gene expression and delays flowering time. Plant Direct **4**, 1–12.

Abraha M, Chen J, Hamilton SK, Robertson GP. 2020. Long-term evapotranspiration rates for rainfed corn versus perennial bioenergy crops in a mesic landscape. Hydrological Processes **34**, 810–822.

Adati S, Shiotani I. 1962. The cytotaxonomy of the genus *Miscanthus* and its phylogenic status. Bulletin of the Faculty of Agriculture, Mie University **25**, 1–24.

Alam M, Chen Y, Li P, *et al.* 2020. *Ghd7* is a negative regulator of zinc concentration in brown rice. Molecular Breeding **40**, 110.

Amasino RM, Michaels SD. 2010. The timing of flowering. Plant Physiology **154**, 516–520.

Andrés F, Coupland G. 2012. The genetic basis of flowering responses to seasonal cues. Nature Reviews Genetics **13**, 627–639.

Anzoua KG, Yamada T, Henry RJ. 2011. *Miscanthus*. In: Kole C, ed. Wild Crop Relatives: Genomic and Breeding Resources. Springer Berlin Heidelberg, 157–164.

Atienza SG, Ramirez MC, Martin A. 2003. Mapping QTLs controlling flowering date in *Miscanthus sinensis* Anderss. Cereal Research Communications **31**,

265–271.

Beale CV, Long SP. 1995. Can perennial C_4 grasses attain high efficiencies of radiant energy conversion in cool climates? Plant, Cell & Environment **18**, 641–650.

Ben-Naim O, Eshed R, Parnis A, Teper-Bamnolker P, Shalit A, Coupland G, Samach A, Lifschitz E. 2006. The CCAAT binding factor can mediate interactions between CONSTANS-like proteins and DNA. Plant Journal **46**, 462–476.

Benedict HM. 1941. Effect of day length and temperature on the flowering and growth of four species of grasses. Journal of Agriculture Research **61**, 661–672.

Bookout AL, Mangelsdorf DJ. 2003. Quantitative real-time PCR protocol for analysis of nuclear receptor signaling pathways. Nuclear Receptor Signaling **1**, nrs.01012.

Borchert R, Renner SS, Calle Z, Havarrete D, Tye A, Gautier L, Spichiger R, Von Hildebrand P. 2005. Photoperiodic induction of synchronous flowering near the Equator. Nature **433**, 627–629.

Cai X, Ballif J, Endo S, Davis E, Liang M, Chen D, Dewald D, Kreps J, Zhu T, Wu Y. 2007. A putative CCAAT-binding transcription factor is a regulator of flowering timing in arabidopsis. Plant Physiology **145**, 98–105.

Castelletti S, Coupel-Ledru A, Granato I, Palaffre C, Cabrera-Bosquet L,

Tonelli C, Nicolas SD, Tardieu F, Welcker C, Conti L. 2020. Maize adaptation across temperate climates was obtained via expression of two florigen genes. PLoS Genetics **16**, e1008882.

Casto AL, Mattison AJ, Olson SN, Thakran M, Rooney WL, Mullet JE. 2019. *Maturity2*, a novel regulator of flowering time in *Sorghum bicolor*, increases expression of *SbPRR37* and *SbCO* in long days delaying flowering. PLoS ONE **14**, e0212154.

Castro JC, Boe A, Lee DK. 2011. A simple system for promoting flowering of upland switchgrass in the greenhouse. Crop Science **51**, 2607–2614.

Chalfun-Junior A, Franken J, Mes JJ, Marsch-Martinez N, Pereira A, Angenent GC. 2005. *ASYMMETRIC LEAVES2-LIKE1* gene, a member of the AS2/LOB family, controls proximal-distal patterning in *Arabidopsis* petals. Plant Molecular Biology **57**, 559–575.

Cho LH, Yoon J, An G. 2017. The control of flowering time by environmental factors. Plant Journal **90**, 708–719.

Chramiec-Głabik A, Grabowska-Joachimiak A, Sliwinska E, Legutko J, Kula A. 2012. Cytogenetic analysis of *Miscanthus* × *giganteus* and its parent forms. Caryologia **65**, 234–242.

Chu S, Majumdar A. 2012. Opportunities and challenges for a sustainable energy future. Nature **488**, 294–303.

Clark LV, Brummer JE, Głowacka K, *et al.* 2014. A footprint of past climate

change on the diversity and population structure of *Miscanthus sinensis*. Annals of Botany **114**, 97–107.

Clark LV, Stewart JR, Nishiwaki A, *et al.* 2015. Genetic structure of *Miscanthus sinensis* and *Miscanthus sacchariflorus* in Japan indicates a gradient of bidirectional but asymmetric introgression. Journal of Experimental Botany **66**, 4213–4225.

Clifford HT, Clayton WD, Renvoize SA. 1986. Genera Graminum. Grasses of the World. Kew, London: Royal Botanic Gardens.**13**, 1-389.

Clifton-Brown J, Harfouche A, Casler MD, *et al.* 2019. Breeding progress and preparedness for mass-scale deployment of perennial lignocellulosic biomass crops switchgrass, *Miscanthus*, willow and poplar. GCB Bioenergy **11**, 118–151.

Clifton-Brown JC, Jones MB. 1997. The thermal response of leaf extension rate in genotypes of the C_4-grass *Miscanthus*: an important factor in determining the potential productivity of different genotypes. Journal of Experimental Botany **48**, 1573–1581.

Clifton-Brown JC, Lewandowski I, Andersson B, *et al.* 2001. Performance of 15 *Miscanthus* genotypes at five sites in Europe. Agronomy Journal **93**, 1013–1019.

Dai X, Ding Y, Tan L, *et al.* 2012. *LHD1*, an allele of *DTH8*/*Ghd8*, controls late heading date in common wild rice (*Oryza rufipogon*). Journal of Integrative Plant Biology **54**, 790–799.

Danilevskaya ON, Meng X, Hou Z, Ananiev EV, Simmons CR. 2008. A genomic and expression compendium of the expanded PEBP gene family from maize. Plant Physiology **146**, 250–264.

Deuter M. 2000. Breeding approaches to improvement of yield and quality in *Miscanthus* grown in Europe. European *Miscanthus* improvement-final Report. Stuttgart, Germany, 28–52.

Doi K, Izawa T, Fuse T, Yamanouchi U, Kubo T, Shimatani Z, Yano M, Yoshimura A. 2004. *Ehd1*, a B-type response regulator in rice, confers short-day promotion of flowering and controls *FT-like* gene expression independently of *Hd1*. Genes and Development **18**, 926–936.

Dong H, Clark LV, Jin X, *et al.* 2021. Managing flowering time in *Miscanthus* and sugarcane to facilitate intra- and intergeneric crosses. PLoS ONE **16**, e0240390.

Dong H, Clark LV, Lipka AE, *et al.* 2019*a*. Winter hardiness of *Miscanthus* (III): Genome-wide association and genomic prediction for overwintering ability in *Miscanthus sinensis*. GCB Bioenergy **11**, 930–955.

Dong H, Green SV, Nishiwaki A, Yamada T, Stewart JR, Deuter M, Sacks EJ. 2019*b*. Winter hardiness of *Miscanthus* (I): Overwintering ability and yield of new *Miscanthus* × *giganteus* genotypes in Illinois and Arkansas. GCB Bioenergy **11**, 691–705.

Dong H, Liu S, Clark LV, Sharma S, Gifford JM, Juvik JA, Lipka AE, Sacks EJ. 2018. Genetic mapping of biomass yield in three interconnected

Miscanthus populations. GCB Bioenergy **10**, 165–185.

Du H, Huang F, Wu N, Li X, Hu H, Xiong L. 2018. Integrative regulation of drought escape through ABA-dependent and -independent pathways in rice. Molecular Plant **11**, 584–597.

Du A, Tian W, Wei M, Yan W, He H, Zhou D, Huang X, Li S, Ouyang X. 2017. The *DTH8-Hd1* module mediates day-length-dependent regulation of rice flowering. Molecular Plant **10**, 948–961.

Dubis B, Jankowski KJ, Załuski D, Bórawski P, Szempliński W. 2019. Biomass production and energy balance of *Miscanthus* over a period of 11 years: A case study in a large-scale farm in Poland. GCB Bioenergy **11**, 1187–1201.

Dwiyanti MS, Yamada T, Sato M, Abe J, Kitamura K. 2011. Genetic variation of γ-tocopherol methyltransferase gene contributes to elevated α-tocopherol content in soybean seeds. BMC Plant Biology **11**, 152.

Ehleringer JR, Cerling TE. 2002. C_3 and C_4 photosynthesis. In: Mooney HA, Canadell JG, eds. Encyclopedia of Global Environmental Change. Chichester: John Wiley & Sons, Ltd., 186–190.

Fares MA, Keane OM, Toft C, Carretero-Paulet L, Jones GW. 2013. The roles of whole-genome and small-scale duplications in the functional specialization of *Saccharomyces cerevisiae* genes. PLoS Genetics **9**, e1003176.

Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39**, 783.

Fornara F, de Montaigu A, Coupland G. 2010. SnapShot: Control of flowering in arabidopsis. Cell **141**, 550-550.e2.

Fornara F, Panigrahi KCS, Gissot L, Sauerbrunn N, Rühl M, Jarillo JA, Coupland G. 2009. *Arabidopsis* DOF transcription factors act redundantly to reduce *CONSTANS* expression and are essential for a photoperiodic flowering response. Developmental Cell **17**, 75–86.

Frank DC, Esper J, Raible CC, Büntgen U, Trouet V, Stocker B, Joos F. 2010. Ensemble reconstruction constraints on the global carbon cycle sensitivity to climate. Nature **463**, 527–530.

Fujino K. 2020. Days to heading, controlled by the heading date genes, *HD1* and *DTH8*, limits rice yield-related traits in Hokkaido, Japan. Breeding Science **70**, 277–282.

Fujino K, Obara M, Ikegaya T. 2019. Establishment of adaptability to the northern-limit of rice production. Molecular Genetics and Genomics **294**, 729–737.

Fujino K, Yamanouchi U. 2020. Genetic effect of a new allele for the flowering time locus *Ghd7* in rice. Breeding Science **70**, 342–346.

Fujino K, Yamanouchi U, Yano M. 2013. Roles of the *Hd5* gene controlling heading date for adaptation to the northern limits of rice cultivation. Theoretical and Applied Genetics **126**, 611–618.

Gale MD, Devos KM. 1998. Comparative genetics in the grasses. Proceedings of

the National Academy of Sciences of the United States of America **95**, 1971–1974.

Gao H, Jin M, Zheng X, *et al.* 2014. *Days to heading 7*, a major quantitative locus determining photoperiod sensitivity and regional adaptation in rice. Proceedings of the National Academy of Sciences of the United States of America **111**, 16337–16342.

Gifford JM, Chae WB, Swaminathan K, Moose SP, Juvik JA. 2015. Mapping the genome of *Miscanthus sinensis* for QTL associated with biomass productivity. GCB Bioenergy **7**, 797–810.

Gil KE, Park MJ, Lee HJ, Park YJ, Han SH, Kwon YJ, Seo PJ, Jung JH, Park CM. 2017. Alternative splicing provides a proactive mechanism for the diurnal *CONSTANS* dynamics in *Arabidopsis* photoperiodic flowering. Plant Journal **89**, 128–140.

Glassop D, Rae AL. 2019. Expression of sugarcane genes associated with perception of photoperiod and floral induction reveals cycling over a 24-hour period. Functional Plant Biology **46**, 314–327.

Glémin S, Bataillon T. 2009. A comparative view of the evolution of grasses under domestication. New Phytologist **183**, 273–290.

Glover NM, Daron J, Pingault L, Vandepoele K, Paux E, Feuillet C, Choulet F. 2015. Small-scale gene duplications played a major role in the recent evolution of wheat chromosome 3B. Genome Biology **16**, 188.

Głowacka K, Adhikari S, Peng J, Gifford J, Juvik JA, Long SP, Sacks EJ. 2014. Variation in chilling tolerance for photosynthesis and leaf extension growth among genotypes related to the C_4 grass *Miscanthus* × *giganteus*. Journal of Experimental Botany **65**, 5267–5278.

Godiska R, Mead D, Dhodda V, Wu C, Hochstein R, Karsi A, Usdin K, Entezam A, Ravin N. 2009. Linear plasmid vector for cloning of repetitive or unstable sequences in Escherichia coli. Nucleic Acids Research **38**, e88.

Gómez-Ariza J, Galbiati F, Goretti D, Brambilla V, Shrestha R, Pappolla A, Courtois B, Fornara F. 2015. Loss of floral repressor function adapts rice to higher latitudes in Europe. Journal of Experimental Botany **66**, 2027–2039.

Goretti D, Martignago D, Landini M, *et al.* 2017. Transcriptional and post-transcriptional mechanisms limit *Heading Date 1* (*Hd1*) function to adapt rice to high latitudes. PLoS Genetics **13**, e1006530.

Greef J, Deuter M. 1993. Syntaxonomy of *Miscanthus* × *giganteus* Greef et Deu. Angewandte Botanik **67**, 87–90.

Griffiths S, Dunford RP, Coupland G, Laurie DA. 2003. The evolution of *CONSTANS-like* gene families in barley, rice, and Arabidopsis. Plant Physiology **131**, 1855–1867.

Hamilton SK, Hussain MZ, Bhardwaj AK, Basso B, Robertson GP. 2015. Comparative water use by maize, perennial crops, restored prairie, and poplar trees in the US Midwest. Environmental Research Letters **10**, 64015.

Hastings A, Mos M, Yesufu JA, *et al.* 2017. Economic and environmental assessment of seed and rhizome propagated *Miscanthus* in the UK. Frontiers in Plant Science **8**, 1058.

Hayama R, Yokoi S, Tamaki S, Yano M, Shimamoto K. 2003. Adaptation of photoperiodic control pathways produces short-day flowering in rice. Nature **422**, 719–722.

Heaton EA, Dohleman FG, Long SP. 2008. Meeting US biofuel goals with less land: The potential of *Miscanthus*. Global Change Biology **14**, 2000–2014.

Hodkinson TR, Chase MW, Takahashi C, Leitch IJ, Bennett MD, Renvoize SA. 2002. The use of DNA sequencing (ITS and TRNL-F), AFLP, and fluorescent in situ hybridization to study allopolyploid *Miscanthus* (Poaceae). American Journal of Botany **89**, 279–286.

Hodkinson TR, Klaas M, Jones MB, Prickett R, Barth S. 2015. *Miscanthus*: a case study for the utilization of natural genetic variation. Plant Genetic Resources **13**, 219–237.

Holder IT, Wagner S, Xiong P, Sinn M, Frickey T, Meyer A, Hartig JS. 2015. Intrastrand triplex DNA repeats in bacteria: a source of genomic instability. Nucleic Acids Research **43**, 10126–10142.

Hu FY, Tao DY, Sacks E, *et al.* 2003. Convergent evolution of perenniality in rice and sorghum. Proceedings of the National Academy of Sciences of the United States of America **100**, 4050–4054.

Huang CL, Liao WC, Lai YC. 2011. Cultivation studies of Taiwanese native *Miscanthus floridulus* lines. Biomass and Bioenergy **35**, 1873–1877.

Hung HY, Shannon LM, Tian F, *et al.* 2012. *ZmCCT* and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. Proceedings of the National Academy of Sciences of the United States of America **109**, E1913–E1921.

Hwang DY, Park S, Lee S, Lee SS, Imaizumi T, Song YH. 2019. *GIGANTEA* regulates the timing stabilization of CONSTANS by altering the interaction between FKF1 and ZEITLUPE. Molecules and Cells **42**, 693–701.

International Energy Agency. 2017. Technology Roadmap: Delivering Sustainable Bioenergy.

International Energy Agency. 2020. Key World Energy Statistics.

Iqbal Y, Gauder M, Claupein W, Graeff-Hönninger S, Lewandowski I. 2015. Yield and quality development comparison between *Miscanthus* and switchgrass over a period of 10 years. Energy **89**, 268–276.

Iqbal Y, Kiesel A, Wagner M, Nunn C, Kalinina O, Hastings AFSJ, Clifton-Brown JC, Lewandowski I. 2017. Harvest time optimization for combustion quality of different miscanthus genotypes across europe. Frontiers in Plant Science **8**, 727.

Ito S, Song YH, Josephson-Day AR, Miller RJ, Breton G, Olmstead RG, Imaizumi T. 2012. FLOWERING BHLH transcriptional activators control

expression of the photoperiodic flowering regulator *CONSTANS* in *Arabidopsis*. Proceedings of the National Academy of Sciences of the United States of America **109**, 3582–3587.

Itoh H, Nonoue Y, Yano M, Izawa T. 2010. A pair of floral regulators sets critical day length for *Hd3a* florigen expression in rice. Nature Genetics **42**, 635–638.

Izawa T, Mihara M, Suzuki Y, *et al.* 2011. *Os-GIGANTEA* confers robust diurnal rhythms on the global transcriptome of rice in the field. Plant Cell **23**, 1741–1755.

Izawa T, Oikawa T, Sugiyama N, Tanisaka T, Yano M, Shimamoto K. 2002. Phytochrome mediates the external light signal to repress *FT* orthologs in photoperiodic flowering of rice. Genes and Development **16**, 2006–2020.

Jackson SD. 2009. Plant responses to photoperiod. New Phytologist **181**, 517–531.

Jensen E, Farrar K, Thomas-Jones S, Hastings A, Donnison I, Clifton-Brown J. 2011. Characterization of flowering time diversity in *Miscanthus* species. GCB Bioenergy **3**, 387–400.

Jensen E, Robson P, Norris J, Cookson A, Farrar K, Donnison I, Clifton-Brown J. 2013. Flowering induction in the bioenergy grass *Miscanthus sacchariflorus* is a quantitative short-day response, whilst delayed flowering under long days increases biomass accumulation. Journal of Experimental Botany **64**, 541–552.

Jensen E, Shafiei R, Ma X, *et al.* 2021. Linkage mapping evidence for a syntenic

QTL associated with flowering time in perennial C_4 rhizomatous grasses *Miscanthus* and switchgrass. GCB Bioenergy **13**, 98–111.

Jiao X, Kørup K, Andersen MN, Sacks EJ, Zhu X, Lærke PE, Jørgensen U. 2017. Can *Miscanthus* C_4 photosynthesis compete with festulolium C_3 photosynthesis in a temperate climate? GCB Bioenergy **9**, 18–30.

Johansson M, Staiger D. 2015. Time to flower: Interplay between photoperiod and the circadian clock. Journal of Experimental Botany **66**, 719–730.

Jørgensen U. 2011. Benefits versus risks of growing biofuel crops: The case of *Miscanthus*. Current Opinion in Environmental Sustainability **3**, 24–30.

Jorgensen U, Muhs HJ. 2001. *Miscanthus* breeding and improvement. In: Jones MB, Walsh M, eds. *Miscanthus* for energy and fibre. London, UK: James & James Ltd., 68–85.

Kaczorowski KA, Quail PH. 2003. *Arabidopsis Pseudo-Response Regulator 7* is a signaling intermediate in phytochrome-regulated seedling deetiolation and phasing of the circadian clock. Plant Cell **15**, 2654–2665.

Kim DH. 2020. Current understanding of flowering pathways in plants: focusing on the vernalization pathway in *Arabidopsis* and several vegetable crop plants. Horticulture Environment and Biotechnology **61**, 209–227.

Kim WY, Fujiwara S, Suh SS, Kim J, Kim Y, Han L, David K, Putterill J, Nam HG, Somers DE. 2007. ZEITLUPE is a circadian photoreceptor stabilized by GIGANTEA in blue light. Nature **449**, 356–360.

Kim J, Geng R, Gallenstein RA, Somers DE. 2013. The F-box protein ZEITLUPE controls stability and nucleocytoplasmic partitioning of GIGANTEA. Development **140**, 4060–4069.

Kim C, Zhang D, Auckland SA, Rainville LK, Jakob K, Kronmiller B, Sacks EJ, Deuter M, Paterson AH. 2012. SSR-based genetic maps of *Miscanthus sinensis* and *M. sacchariflorus*, and their comparison to sorghum. Theoretical and Applied Genetics **124**, 1325–1338.

Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution **16**, 111–120.

Klein RR, Miller FR, Dugas DV, Brown PJ, Burrell AM, Klein PE. 2015. Allelic variants in the *PRR37* gene and the human-mediated dispersal and diversification of sorghum. Theoretical and Applied Genetics **128**, 1669–1683.

Kørup K, Lærke PE, Baadsgaard H, Andersen MN, Kristensen K, Münnich C, Didion T, Jensen ES, Mårtensson LM, Jørgensen U. 2018. Biomass production and water use efficiency in perennial grasses during and after drought stress. GCB Bioenergy **10**, 12–27.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Molecular Biology and Evolution **35**, 1547–1549.

Kumimoto RW, Adam L, Hymus GJ, Repetti PP, Reuber TL, Marion CM, Hempel FD, Ratcliffe OJ. 2008. The Nuclear Factor Y subunits NF-YB2 and

NF-YB3 play additive roles in the promotion of flowering by inductive long-day photoperiods in *Arabidopsis*. Planta **228**, 709–723.

Lallemand T, Leduc M, Landès C, Rizzon C, Lerat E. 2020. An overview of duplicated gene detection methods: Why the duplication mechanism has to be accounted for in their choice. Genes **11**, 1–40.

Lewandowski I. 1998. Propagation method as an important factor in the growth and development of *Miscanthus* × *giganteus*. Industrial Crops and Products **8**, 229–245.

Lewandowski I, Clifton-Brown JC, Andersson B, *et al.* 2003. Environment and harvest time affects the combustion qualities of *Miscanthus* genotypes. Agronomy Journal **95**, 1274–1280.

Li X, Liu H, Wang M, *et al.* 2015. Combinations of *Hd2* and *Hd4* genes determine rice adaptability to Heilongjiang Province, northern limit of China. Journal of Integrative Plant Biology **57**, 698–707.

Liao P, Lee KH. 2010. From SNPs to functional polymorphism: The insight into biotechnology applications. Biochemical Engineering Journal **49**, 149–158.

Liu J, Gong J, Wei X, Yang S, Huang X, Li C, Zhou X. 2020. Dominance complementation of *Hd1* and *Ghd8* contributes to extremely late flowering in two rice hybrids. Molecular Breeding **40**, 1–10.

Liu H, Liu H, Zhou L, Zhang Z, Zhang X, Wang M, Li H, Lin Z. 2015. Parallel domestication of the *heading date 1* gene in cereals. Molecular Biology and

Evolution **32**, 2726–2737.

Lu L, Yan W, Xue W, Shao D, Xing Y. 2012. Evolution and association analysis of *Ghd7* in rice. PLoS ONE **7**, e34021.

Luccioni L, Krzymuski M, Sánchez-Lamas M, Karayekov E, Cerdán PD, Casal JJ. 2019. *CONSTANS* delays *Arabidopsis* flowering under short days. Plant Journal **97**, 923–932.

Lv X, Zeng X, Hu H, *et al.* 2021. Structural insights into the multivalent binding of the *Arabidopsis FLOWERING LOCUS T* promoter by the CO-NF-Y master transcription factor complex. The Plant Cell **0**, 1–14.

Mandley SJ, Daioglou V, Junginger HM, van Vuuren DP, Wicke B. 2020. EU bioenergy development to 2050. Renewable and Sustainable Energy Reviews **127**, 109858.

Matlaga DP, Quinn LD, Davis AS, Stewart JR. 2012. Light response of native and introduced *Miscanthus sinensis* seedlings. Invasive Plant Science and Management **5**, 363–374.

McCormick RF, Truong SK, Sreedasyam A, *et al.* 2018. The *Sorghum bicolor* reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. Plant Journal **93**, 338–354.

McMillan C. 1959. The role of ecotypic variation in the distribution of the central grassland of North America. Ecological Monographs **29**, 285–308.

McWatters HG, Devlin PF. 2011. Timing in plants - a rhythmic arrangement.

FEBS Letters **585**, 1474–1484.

Meng X, Muszynski MG, Danilevskaya ON. 2011. The *FT-like ZCN8* gene functions as a floral activator and is involved in photoperiod sensitivity in maize. Plant Cell **23**, 942–960.

Meyer MH, Paul J, Anderson NO. 2010. Competive ability of invasive *Miscanthus* biotypes with aggressive switchgrass. Biological Invasions **12**, 3809–3816.

Ming R, Del Monte TA, Hernandez E, Moore PH, Irvine JE, Paterson AH. 2002. Comparative analysis of QTLs affecting plant height and flowering among closely-related diploid and polyploid genomes. Genome **45**, 794–803.

Mitros T, Session AM, James BT, *et al.* 2020. Genome biology of the paleotetraploid perennial biomass crop *Miscanthus*. Nature Communications **11**, 1–11.

Murphy RL, Klein RR, Morishige DT, Brady JA, Rooney WL, Miller FR, Dugas DV, Klein PE, Mullet JE. 2011. Coincident light and clock regulation of *pseudoresponse regulator protein 37* (*PRR37*) controls photoperiodic flowering in sorghum. Proceedings of the National Academy of Sciences of the United States of America **108**, 16469–16474.

Murphy RL, Morishige DT, Brady JA, Rooney WL, Yang S, Klein PE, Mullet JE. 2014. *Ghd7* (*Ma6*) represses sorghum flowering in long days: *Ghd7* alleles enhance biomass accumulation and grain production. The Plant Genome **7**, 1–10.

Nagano H, Clark LV, Zhao H, *et al.* 2015. Contrasting allelic distribution of *CO*/*Hd1* homologues in *Miscanthus sinensis* from the East Asian mainland and the Japanese archipelago. Journal of Experimental Botany **66**, 4227–4237.

Nemoto Y, Nonoue Y, Yano M, Izawa T. 2016. *Hd1*, a *CONSTANS* ortholog in rice, functions as an *Ehd1* repressor through interaction with monocot-specific CCT-domain protein Ghd7. The Plant journal **86**, 221–233.

Nolan T, Hands RE, Bustin SA. 2006. Quantification of mRNA using real-time RT-PCR. Nature Protocols **1**, 1559–1582.

Oliveira JA, West CP, Afif E, Palencia P. 2017. Comparison of *Miscanthus* and switchgrass cultivars for biomass yield, soil nutrients, and nutrient removal in northwest Spain. Agronomy Journal **109**, 122–130.

Park DH, Somers DE, Kim YS, Choy YH, Lim HK, Soh MS, Kim HJ, Kay SA, Nam HG. 1999. Control of circadian rhythms and photoperiodic flowering by the *Arabidopsis GIGANTEA* gene. Science **285**, 1579–1582.

Pignon CP, Spitz I, Sacks EJ, Jørgensen U, Kørup K, Long SP. 2019. Siberian *Miscanthus sacchariflorus* accessions surpass the exceptional chilling tolerance of the most widely cultivated clone of *Miscanthus* × *giganteus*. GCB Bioenergy **11**, 883–894.

Preston JC, Fjellheim S. 2020. Understanding past, and predicting future, niche transitions based on grass flowering time variation. Plant Physiology **183**, 822–839.

Putterill J, Robson F, Lee K, Simon R, Coupland G. 1995. The *CONSTANS* gene of arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. Cell **80**, 847–857.

Quinn LD, Allen DJ, Stewart JR. 2010. Invasiveness potential of *Miscanthus sinensis*: implications for bioenergy production in the United States. GCB Bioenergy **2**, 310–320.

Quinn LD, Culley TM, Stewart JR. 2012. Genetic comparison of introduced and native populations of *Miscanthus sinensis* (Poaceae), a potential bioenergy crop. Grassland Science **58**, 101–111.

Ream TS, Woods DP, Amasino RM. 2012. The molecular basis of vernalization in different plant groups. Cold Spring Harbor Symposia on Quantitative Biology **77**, 105–115.

Robertson GP, Hamilton SK, Barham BL, Dale BE, Izaurralde RC, Jackson RD, Landis DA, Swinton SM, Thelen KD, Tiedje JM. 2017. Cellulosic biofuel contributions to a sustainable energy future: Choices and outcomes. Science **356**, eaal2324.

Robson P, Jensen E, Hawkins S, White SR, Kenobi K, Clifton-Brown J, Donnison I, Farrar K. 2013. Accelerating the domestication of a bioenergy crop: Identifying and modelling morphological targets for sustainable yield increase in *Miscanthus*. Journal of Experimental Botany **64**, 4143–4155.

Rooney WL, Blumenthal J, Bean B, Mullet JE. 2007. Designing sorghum as a dedicated bioenergy feedstock. Biofuels, Bioproducts and Biorefining **1**, 147–

157.

Saad M, Mary H, Amjid U, Shabir G, Aslam K, Shah SM, Khan AR. 2019. Photoperiodic stress induces genotype-specific shift in DNA methylation in Tartary buckwheat. Biologia Futura **70**, 278–285.

Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution **4**, 406–425.

Salomé PA, To JPC, Kieber JJ, McClung CR. 2006. *Arabidopsis* response regulators *ARR3* and *ARR4* play cytokinin-independent roles in the control of circadian period. Plant Cell **18**, 55–69.

Salse J, Abrouk M, Bolot S, Guilhot N, Courcelle E, Faraut T, Waugh R, Close TJ, Messing J, Feuillet C. 2009. Reconstruction of monocotelydoneous proto-chromosomes reveals faster evolution in plants than in animals. Proceedings of the National Academy of Sciences of the United States of America **106**, 14908–14913.

Sarid-Krebs L, Panigrahi KCS, Fornara F, Takahashi Y, Hayama R, Jang S, Tilmes V, Valverde F, Coupland G. 2015. Phosphorylation of CONSTANS and its COP1-dependent degradation during photoperiodic flowering of *Arabidopsis*. Plant Journal **84**, 451–463.

Sawa M, Kay SA. 2011. *GIGANTEA* directly activates *Flowering Locus T* in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences of the United States of America **108**, 11698–11703.

Sawa M, Nusinow DA, Kay SA, Imaizumi T. 2007. FKF1 and GIGANTEA complex formation is required for day-length measurement in *Arabidopsis*. Science **318**, 261–265.

Schaffer R, Landgraf J, Accerbi M, Simon V, Larson M, Wisman E. 2001. Microarray analysis of diurnal and circadian-regulated genes in *Arabidopsis*. Plant Cell **13**, 113–123.

Scholin CA, Herzog M, Sogin M, Anderson DM. 1994. Identification of group- and strain-specific genetic markers for globally distributed *Alexandrium* (Dinophyceae). ii. sequence analysis of a fragment of the LSU rRNA gene. Journal of Phycology **30**, 999–1011.

Scordia D, Scalici G, Clifton-Brown J, Robson P, Patanè C, Cosentino SL. 2020. Wild *Miscanthus* germplasm in a drought-affected area: Physiology and agronomy appraisals. Agronomy **10**, 679.

Scordia D, Testa G, Cosentino SL. 2014. Perennial grasses as lignocellulosic feedstock for second-generation bioethanol production in mediterranean environment. Italian Journal of Agronomy **9**, 84–92.

Shaw LM, Li C, Woods DP, Alvarez MA, Lin H, Lau MY, Chen A, Dubcovsky J. 2020. Epistatic interactions between *PHOTOPERIOD1*, *CONSTANS1* and *CONSTANS2* modulate the photoperiodic response in wheat. PLoS Genetics **16**, e1008812.

Shim JS, Kubota A, Imaizumi T. 2017. Circadian clock and photoperiodic flowering in *Arabidopsis*: *CONSTANS* is a hub for signal integration. Plant

Physiology **173**, 5–15.

Shrestha R, Gómez-Ariza J, Brambilla V, Fornara F. 2014. Molecular control of seasonal flowering in rice, *Arabidopsis* and temperate cereals. Annals of Botany **114**, 1445–1458.

Siri-Prieto G, Bustamante M, Picasso V, Ernst O. 2020. Impact of nitrogen and phosphorous on biomass yield, nitrogen efficiency, and nutrient removal of perennial grasses for bioenergy. Biomass and Bioenergy **136**, 105526.

Song YH, Estrada DA, Johnson RS, Kim SK, Lee SY, MacCoss MJ, Imaizumi T. 2014. Distinct roles of FKF1, GIGANTEA, and ZEITLUPE proteins in the regulation of constans stability in *Arabidopsis* photoperiodic flowering. Proceedings of the National Academy of Sciences of the United States of America **111**, 17672–17677.

Song YH, Ito S, Imaizumi T. 2010. Similarities in the circadian clock and photoperiodism in plants. Current Opinion in Plant Biology **13**, 594–603.

Song YH, Shim JS, Kinmonth-Schultz HA, Imaizumi T. 2015. Photoperiodic flowering: time measurement mechanisms in leaves. Annual Review of Plant Biology **66**, 441–464.

Stecher G, Tamura K, Kumar S. 2020. Molecular evolutionary genetics analysis (MEGA) for macOS. Molecular Biology and Evolution **37**, 1237–1239.

Strayer C, Oyama T, Schultz TF, Raman R, Somers DE, Mas P, Panda S, Kreps JA, Kay SA. 2000. Cloning of the *Arabidopsis* clock gene *TOC1*, an

autoregulatory response regulator homolog. Science **289**, 768–771.

Su L, Shan J, Gao J, Lin H. 2016. OsHAL3, a blue light-responsive protein, interacts with the floral regulator *Hd1* to activate flowering in rice. Molecular Plant **9**, 233–244.

Suárez-López P, Wheatley K, Robson F, Onouchi H, Valverde F, Coupland G. 2001. *CONSTANS* mediates between the circadian clock and the control of flowering in *Arabidopsis*. Nature **410**, 1116–1120.

Swaminathan K, Chae WB, Mitros T, *et al.* 2012. A framework genetic map for *Miscanthus sinensis* from RNAseq-based markers shows recent tetraploidy. BMC Genomics **13**, 142.

Tamaki S, Matsuo S, Hann LW, Yokoi S, Shimamoto K. 2007. Hd3a protein is a mobile flowering signal in rice. Science **316**, 1033–1036.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research **22**, 4673–4680.

Tilman D, Hill J, Lehman C. 2006. Carbon-negative biofuels from low-input high-diversity grassland biomass. Science **314**, 1598–1600.

Tubeileh A, Rennie TJ, Goss MJ. 2016. A review on biomass production from C_4 grasses: yield and quality for end-use. Current Opinion in Plant Biology **31**, 172–180.

Turner A, Beales J, Faure S, Dunford RP, Laurie DA. 2005. The pseudo-response regulator *Ppd-H1* provides adaptation to photoperiod in barley. Science **310**, 1031–1034.

Valverde F, Mouradov A, Soppe W, Ravenscroft D, Samach A, Coupland G. 2004. Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. Science **303**, 1003–1006.

Vermerris W. 2008. *Miscanthus*: Genetic resources and breeding potential to enhance bioenergy production. In: Vermerris W, ed. Genetic Improvement of Bioenergy Crops. Springer, New York, NY, 295–308.

Wang P, Gong R, Yang Y, Yu S. 2019*a*. *Ghd8* controls rice photoperiod sensitivity by forming a complex that interacts with *Ghd7*. BMC Plant Biology **19**, 1–14.

Wang C, Kong Y, Hu R, Zhou G. 2021. *Miscanthus*: a fast-growing crop for environmental remediation and biofuel production. GCB Bioenergy **13**, 58–69.

Wang P, Xiong Y, Gong R, Yang Y, Fan K, Yu S. 2019*b*. A key variant in the cis-regulatory element of flowering gene *Ghd8* associated with cold tolerance in rice. Scientific Reports **9**, 1–14.

Webb AAR. 2003. The physiology of circadian rhythms in plants. New Phytologist **160**, 281–303.

Wei H, Wang X, Xu H, Wang L. 2020. Molecular basis of heading date control in rice. aBIOTECH **1**, 219–232.

Wei X, Xu J, Guo H, Jiang L, Chen S, Yu C, Zhou Z, Hu P, Zhai H, Wan J. 2010. *DTH8* suppresses flowering in rice, influencing plant height and yield potential simultaneously. Plant Physiology **153**, 1747–1758.

van der Weijde T, Huxley LM, Hawkins S, Sembiring EH, Farrar K, Dolstra O, Visser RGF, Trindade LM. 2017*a*. Impact of drought stress on growth and quality of *Miscanthus* for biofuel production. GCB Bioenergy **9**, 770–782.

van der Weijde T, Kamei CLA, Severing EI, Torres AF, Gomez LD, Dolstra O, Maliepaard CA, McQueen-Mason SJ, Visser RGF, Trindade LM. 2017*b*. Genetic complexity of *Miscanthus* cell wall composition and biomass quality for biofuels. BMC Genomics **18**, 406.

Weng X, Lovell JT, Schwartz SL, Cheng C, Haque T, Zhang L, Razzaque S, Juenger TE. 2019. Complex interactions between day length and diurnal patterns of gene expression drive photoperiodic responses in a perennial C_4 grass. Plant, Cell & Environment **42**, 2165–2182.

Wenkel S, Turck F, Singer K, Gissot L, Le Gourrierec J, Samach A, Coupland G. 2006. CONSTANS and the CCAAT box binding complex share a functionally important domain and interact to regulate flowering of *Arabidopsis*. Plant Cell **18**, 2971–2984.

Wolabu TW, Zhang F, Niu L, Kalve S, Bhatnagar-Mathur P, Muszynski MG, Tadege M. 2016. Three *FLOWERING LOCUS T-like* genes function as potential florigens and mediate photoperiod response in sorghum. New Phytologist **210**, 946–959.

Woods DP, McKeown MA, Dong Y, Preston JC, Amasino RM. 2016. Evolution of *VRN2*/*Ghd7-like* genes in vernalization-mediated repression of grass flowering. Plant Physiology **170**, 2124–2135.

Wu W, Hasegawa T, Fujimori S, Takahashi K, Oshiro K. 2020. Assessment of bioenergy potential and associated costs in Japan for the 21st century. Renewable Energy **162**, 308–321.

Xue S, Kalinina O, Lewandowski I. 2015. Present and future options for *Miscanthus* propagation and establishment. Renewable and Sustainable Energy Reviews **49**, 1233–1246.

Xue W, Xing Y, Weng X, *et al.* 2008. Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. Nature Genetics **40**, 761–767.

Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, SanMiguel P, Bennetzen JL, Echenique V, Dubcovsky J. 2004. The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. Science **303**, 1640–1644.

Yan WH, Wang P, Chen HX, *et al.* 2011. A major QTL, *Ghd8*, plays pleiotropic roles in regulating grain productivity, plant height, and heading date in rice. Molecular Plant **4**, 319–330.

Yang Q, Li Z, Li W, *et al.* 2013. CACTA-like transposable element in *ZmCCT* attenuated photoperiod sensitivity and accelerated the postdomestication spread of maize. Proceedings of the National Academy of Sciences of the United States of America **110**, 16969–16974.

Yang L, Liu T, Li B, Sui Y, Chen J, Shi J, Wing RA, Chen M. 2012. Comparative sequence analysis of the *Ghd7* orthologous regions revealed movement of *Ghd7* in the grass genomes. PLoS ONE **7**, e50236.

Yang S, Murphy RL, Morishige DT, Klein PE, Rooney WL, Mullet JE. 2014*a*. Sorghum *phytochrome B* inhibits flowering in long days by activating expression of *SbPRR37* and *SbGHD7*, repressors of *SbEHD1*, *SbCN8* and *SbCN12*. PLoS ONE **9**, e105352.

Yang S, Weers BD, Morishige DT, Mullet JE. 2014*b*. *CONSTANS* is a photoperiod regulated activator of flowering in sorghum. BMC Plant Biology **14**, 1–15.

Yano M, Katayose Y, Ashikari M, *et al.* 2000. *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. Plant Cell **12**, 2473–2483.

Ye C, Hall SJ. 2020. Mechanisms underlying limited soil carbon gains in perennial and cover-cropped bioenergy systems revealed by stable isotopes. GCB Bioenergy **12**, 101–117.

Yeang HY. 2015. Cycling of clock genes entrained to the solar rhythm enables plants to tell time: data from *Arabidopsis*. Annals of Botany **116**, 15–22.

Yeh RH, Lin YS, Wang TH, Kuan WC, Lee WC. 2016. Bioethanol production from pretreated *Miscanthus floridulus* biomass by simultaneous saccharification and fermentation. Biomass and Bioenergy **94**, 110–116.

Zhang H, Lang Z, Zhu JK. 2018. Dynamics and function of DNA methylation in plants. Nature Reviews Molecular Cell Biology **19**, 489–506.

Zhang B, Liu H, Qi F, Zhang Z, Li Q, Han Z, Xing Y. 2019. Genetic interactions among *Ghd7*, *Ghd8*, *OsPRR37* and *Hd1* contribute to large variation in heading date in rice. Rice **12**, 48.

Zhang J, Zhou X, Yan W, *et al.* 2015. Combinations of the *Ghd7*, *Ghd8* and *Hd1* genes largely define the ecogeographical adaptation and yield potential of cultivated rice. New Phytologist **208**, 1056–1066.

Zhao J, Chen H, Ren D, *et al.* 2015. Genetic interactions between diverged alleles of *Early heading date 1* (*Ehd1*) and *Heading date 3a* (*Hd3a*)/*RICE FLOWERING LOCUS T1* (*RFT1*) control differential heading and contribute to regional adaptation in rice (*Oryza sativa*). New Phytologist **208**, 936–948.

Zheng T, Sun J, Zhou S, *et al.* 2019. Post-transcriptional regulation of Ghd7 protein stability by phytochrome and OsGI in photoperiodic control of flowering in rice. New Phytologist **224**, 306–320.

Zhu X, Long SP, Ort DR. 2010. Improving photosynthetic efficiency for greater yield. Annual Review of Plant Biology **61**, 235–261.

Zhuang Q, Qin Z, Chen M. 2013. Biofuel, land and water: maize, switchgrass or *Miscanthus*? Environmental Research Letters **8**, 015020.

Zong W, Ren D, Huang M, *et al.* 2021. Strong photoperiod sensitivity is controlled by cooperation and competition among *Hd1*, *Ghd7* and *DTH8* in rice heading.

New Phytologist **229**, 1635–1649.

Zuckerkandl E, Pauling Li. 1965. Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ, eds. Evolving Genes and Proteins. New York: Academic Press, 97–166.