

2019

# Winter hardiness of *Miscanthus* (III): Genome-wide association and genomic prediction for overwintering ability in *Miscanthus sinensis*

Hongxu Dong

University of Illinois at Urbana-Champaign

Lindsay V. Clark

University of Illinois at Urbana-Champaign

Alexander E. Lipka

University of Illinois at Urbana-Champaign

Joe E. Brummer

Colorado State University - Fort Collins

Katarzyna Glowacka

University of Nebraska - Lincoln, kglowacka2@unl.edu

*See next page for additional authors*

Follow this and additional works at: <https://digitalcommons.unl.edu/biochemfacpub>

 Part of the [Biochemistry Commons](#), [Biotechnology Commons](#), and the [Other Biochemistry, Biophysics, and Structural Biology Commons](#)

---

Dong, Hongxu; Clark, Lindsay V.; Lipka, Alexander E.; Brummer, Joe E.; Glowacka, Katarzyna; Hall, Megan C.; Heo, Kweon; Jin, Xiaoli; Peng, Junhua; Yamada, Toshihiko; Ghimire, Bimal Kumar; Yoo, Ji Hye; Yu, Chang Yeon; Zhao, Hua; Long, Stephen P.; and Sacks, Erik J., "Winter hardiness of *Miscanthus* (III): Genome-wide association and genomic prediction for overwintering ability in *Miscanthus sinensis*" (2019). *Biochemistry -- Faculty Publications*. 431.

<https://digitalcommons.unl.edu/biochemfacpub/431>





This Article is brought to you for free and open access by the Biochemistry, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Biochemistry -- Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

---

**Authors**

Hongxu Dong, Lindsay V. Clark, Alexander E. Lipka, Joe E. Brummer, Katarzyna Glowacka, Megan C. Hall, Kweon Heo, Xiaoli Jin, Junhua Peng, Toshihiko Yamada, Bimal Kumar Ghimire, Ji Hye Yoo, Chang Yeon Yu, Hua Zhao, Stephen P. Long, and Erik J. Sacks

# Winter hardiness of *Miscanthus* (III): Genome-wide association and genomic prediction for overwintering ability in *Miscanthus sinensis*

Hongxu Dong<sup>1</sup>  | Lindsay V. Clark<sup>1</sup>  | Alexander E. Lipka<sup>1</sup> | Joe E. Brummer<sup>2</sup> | Katarzyna Głowacka<sup>3</sup> | Megan C. Hall<sup>4</sup> | Kweon Heo<sup>5</sup> | Xiaoli Jin<sup>6</sup> | Junhua Peng<sup>7</sup> | Toshihiko Yamada<sup>8</sup>  | Bimal Kumar Ghimire<sup>9</sup> | Ji Hye Yoo<sup>5</sup> | Chang Yeon Yu<sup>5</sup> | Hua Zhao<sup>10</sup> | Stephen P. Long<sup>1</sup>  | Erik J. Sacks<sup>1</sup>

<sup>1</sup>Department of Crop Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois

<sup>2</sup>Department of Soil and Crop Sciences, Colorado State University, Fort Collins, Colorado

<sup>3</sup>Department of Biochemistry, University of Nebraska, Lincoln, Nebraska

<sup>4</sup>Bio Architecture Lab, Berkeley, California

<sup>5</sup>Department of Applied Plant Sciences, Kangwon National University, Chuncheon, Gangwon, South Korea

<sup>6</sup>Department of Agronomy, Zhejiang University, Hangzhou, Zhejiang, China

<sup>7</sup>Huazhi Biotech Co., Ltd, Changsha, Hunan, China

<sup>8</sup>Field Science Center for Northern Biosphere, Hokkaido University, Sapporo, Hokkaido, Japan

<sup>9</sup>Department of Applied Bioscience, Konkuk University, Seoul, South Korea

<sup>10</sup>College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, Hubei, China

## Correspondence

Erik J. Sacks, Department of Crop Sciences, University of Illinois at Urbana-Champaign, Urbana, IL.

Email: esacks@illinois.edu

## Funding information

Energy Biosciences Institute, Grant/Award Number: ILLU-802-311; Biological and Environmental Research, Grant/Award Number: DE-SC0006634 and DE-SC0018420; National Institute of Food and Agriculture, Grant/Award Number: 2018-68005-27937

## Abstract

Overwintering ability is an important selection criterion for *Miscanthus* breeding in temperate regions. Insufficient overwintering ability of the currently leading *Miscanthus* biomass cultivar, *M. ×giganteus* (M×g) ‘1993–1780’, in regions where average annual minimum temperatures are  $-26.1^{\circ}\text{C}$  (USDA hardiness zone 5) or lower poses a pressing need to develop new cultivars with superior cold tolerance. To facilitate breeding of *Miscanthus*, this study characterized phenotypic and genetic variation of overwintering ability in an *M. sinensis* germplasm panel consisting of 564 accessions, evaluated in field trials at three locations in North America and two in Asia. Genome-wide association (GWA) and genomic prediction analyses were performed. The Korea/N China *M. sinensis* genetic group is a valuable gene pool for cold tolerance. The Yangtze-Qinling, Southern Japan, and Northern Japan genetic groups were also potential sources of cold tolerance. A total of 73 marker–trait associations were detected for overwintering ability. Estimated breeding value for overwintering ability based on these 73 markers could explain 55% of the variation for first winter overwintering ability among *M. sinensis*. Average genomic prediction ability for overwintering ability

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *GCB Bioenergy* Published by John Wiley & Sons Ltd.

across 50 fivefold cross-validations was high (~0.73) after accounting for population structure. Common genomic regions for overwintering ability were detected by GWA analyses and a previous parallel QTL mapping study using three interconnected biparental F<sub>1</sub> populations. One QTL on *Miscanthus* LG 8 encompassed five GWA hits and a known cold-responsive gene, *COR47*. The other overwintering ability QTL on *Miscanthus* LG 11 contained two GWA hits and three known cold stress-related genes, carboxylesterase 13 (*CEX13*), *WRKY2* transcription factor, and cold shock domain (*CSDP1*). *Miscanthus* accessions collected from high latitude locations with cold winters had higher rates of overwintering, and more alleles for overwintering, than accessions collected from southern locations with mild winters.

#### KEYWORDS

breeding, genome-wide association analysis, genomic prediction, germplasm, *Miscanthus sinensis*, Overwintering ability

## 1 | INTRODUCTION

The perennial C<sub>4</sub> grass *Miscanthus* is a promising bioenergy crop (Clifton-Brown, Chiang, & Hodkinson, 2008; Clifton-Brown, Stampfl, & Jones, 2004; Głowacka et al., 2014, 2015; Heaton, Dohleman, & Long, 2008; Heaton, Voigt, & Long, 2004; Somerville, Youngs, Taylor, Davis, & Long, 2010). Key objectives for *Miscanthus* breeding are greater biomass yield and better adaptation to local production environments than the current leading biomass cultivar of *M. ×giganteus* (M×g). Currently, only a single triploid clone of M×g, which is derived from the interspecific hybridization between *Miscanthus sacchariflorus* and *M. sinensis*, is available to US farmers (Głowacka et al., 2015; Hodkinson & Renvoize, 2001). We refer to this genotype as M×g ‘1993–1780’ in reference to its accession number in the Kew Living Collection (Hodkinson & Renvoize, 2001), and it was first imported from Yokohama Japan to Denmark in the 1930s (Greef & Deuter, 1993; Linde-Laursen, 1993); in previous studies, we called this genotype ‘Illinois’, as we obtained our initial stock plant from the Chicago Botanic Garden (Maughan et al., 2011). Natural populations of *Miscanthus* are found from tropical and subtropical areas of East Asia and Oceania to ~50°N in eastern Russia (Clifton-Brown et al., 2008; Sacks, Juvik, Lin, Stewart, & Yamada, 2013). Such a wide distribution of *Miscanthus* provides a genetically diverse and valuable gene pool from which to breed new cultivars, including those of M×g.

Overwintering ability is an important selection criterion for perennial bioenergy crops in temperate environments (Burner, Tew, Harvey, & Belesky, 2009; Clifton-Brown & Lewandowski, 2000; Clifton-Brown et al., 2001). Insufficient overwintering ability is a consistent limitation of M×g ‘1993–1780’ in regions with cold winters, such as those in USDA hardiness zone 4 (average annual minimum air

temperature of –34.4°C to –28.9°C; U.S. Department of Agriculture, 2012) and lower, and an intermittent problem in hardiness zone 5 (average annual minimum air temperature of –28.9°C to –23.3°C) and zone 6 (average annual minimum air temperature of –23.3°C to –17.8°C) (Clifton-Brown & Lewandowski, 2000; Dong, Green, et al., 2019; Dong et al., 2018; Dong, Liu, et al., 2019; Heaton et al., 2008; Jørgensen, Mortensen, Kjeldsen, & Schwarz, 2003; Lewandowski, Clifton-Brown, Scurlock, & Huisman, 2000). Clifton-Brown and Lewandowski (2000) observed variation among two M×g, one *M. sacchariflorus*, and two *M. sinensis* grown at four field trial locations in Europe for ability to survive the first winter, and this was not associated with plant size or early senescence in the first autumn. Moreover, controlled environment freeze tests on dormant rhizomes (cold-acclimated) removed from the field during the winter indicated that the lethal temperature at which 50% of the rhizomes died (LT<sub>50</sub>) was –3.4°C for M×g ‘1993–1780’ and *M. sacchariflorus*, whereas the most winter-hardy *M. sinensis* genotype had an LT<sub>50</sub> of –6.5°C, which were consistent with prior observations from field trials (Clifton-Brown & Lewandowski, 2000). Similarly, Friesen, Peixoto, Lee, and Sage (2015) reported that the LT<sub>50</sub> for dormant rhizomes of M×g ‘1993–1780’ was –4°C and it had only 10% survival at –8°C based on controlled environment tests. Heaton et al. (2008) noted that M×g ‘1993–1780’ could survive in the field when air temperatures during the establishment winter were as low as –8°C, with only a 14% loss of plants observed at a northern Illinois site (Shabbona, IL) in the first year, and no losses observed at two more southern locations in year 1 or at any of the locations in the following years.

*Miscanthus* is typically most sensitive to winter damage during the establishment year (i.e., first winter). Christian and Haase (2001) hypothesized that the first-year plants of M×g ‘1993–1780’ did not go dormant early enough in the

season to avoid damage from cold temperatures. Boersma, Dohleman, Miguez, and Heaton (2015) found that when autumn temperature fell below 10°C in Ontario, Canada, CO<sub>2</sub> assimilation rate and photosystem II efficiency for first-year stands of M×g '1993–1780' were almost four times higher than for the third-year plants, and leaf [N] was about 2.4 times greater, suggesting that the first-year plants were still actively growing before a killing frost. From this photosynthesis data, Boersma et al. (2015) concluded that limited translocation of nutrients to rhizomes in the first-year plants of M×g '1993–1780' contributed to their poorer overwintering ability than the third-year plants in Ontario, though no measurements of rhizome nutrient contents were made. Genetic variation in *Miscanthus* for overwintering ability in the establishment year could be due to differences in freeze tolerance of dormant rhizomes, differences among genotypes in timing of dormancy and translocation of nutrients belowground, differences in depth of dormant buds in the soil to facilitate avoidance of freezing temperatures, or combinations of these traits.

Prior studies of overwintering ability and freeze tolerance have been constrained by access to only a narrow subset of *Miscanthus* genetic diversity. In studies of population structure, Clark et al. (2014, 2015) identified seven distinct genetic groups of *M. sinensis* from Asia, including three from Japan, three from China, and one from Korea and northern China; additionally, two *M. sinensis* groups from the US that were derived from the southern and central Japan groups, including US ornamental cultivars and US naturalized populations, and a group of natural diploid *M. sacchariflorus* × *M. sinensis* hybrids (i.e., diploid M×g) found in China were identified. Given that the *M. sinensis* germplasm that has been available in Europe and North America was derived entirely from the southern and central Japan populations (Clark et al., 2014, 2015), studies of hardiness have been limited to just a subset of the southern Japanese germplasm, which itself is a subset of the entire species' diversity. For *M. sacchariflorus*, the situation is similar, with prior studies of hardiness conducted almost entirely on tetraploids from southern Japan, with the exception of just one diploid genotype (*M. sacchariflorus* 'Robustus') that originated from near the Amur River around the border of northeastern China and southeastern Russia (Clark et al., 2018). Moreover, prior studies of M×g have focused on only one or a few genotypes because different genotypes were not available prior to recent efforts by our research group and others to collect and breed new ones.

Thus, an important objective for *Miscanthus* breeders who target temperate growing environments, such as the northern US, Canada, and northern Eurasia, is to identify species, populations within species, and genotypes within populations that are best adapted to the severe winters in these locations yet are also able to produce high biomass yields. We hypothesize that *Miscanthus* accessions collected from high

latitude and/or high altitude locations with cold winters will have higher rates of overwintering in hardiness zone 5 environments such as central Illinois, and more alleles for overwintering, than accessions collected from southern locations with mild winters.

Genome-wide association (GWA) analysis is a powerful approach to determine the genetics of complex traits by exploiting linkage disequilibrium between a marker allele and the causative QTL allele (Lipka et al., 2015). Similarly, genomic prediction can be an effective method to facilitate marker-assisted selection but without the need to identify statistically significant associations between genomic regions and phenotypes. Given that *Miscanthus* is a long-lived perennial, marker-assisted selection is expected to enable plant breeders to efficiently select superior genotypes and greatly improve breeding efficiency. However, few GWA or genomic prediction studies on *Miscanthus* have been published and these have focused primarily on yield traits of *M. sinensis* (Nie et al., 2016; Slavov et al., 2014, 2018) and *M. sacchariflorus* (Clark et al., 2016); to the best of our knowledge, none GWA studies have been published on overwintering ability, and only one genetic mapping study reported QTLs for overwintering ability (Dong, Liu, et al., 2019). To facilitate the use of a previously characterized *M. sinensis* germplasm panel (Clark et al., 2014, 2015) and identify superior genotypes for breeding, we evaluated the overwintering ability of *M. sinensis* genotypes in this germplasm panel in North America and Asia. The objectives of this study were to (a) quantify phenotypic and genetic variation for overwintering ability in *M. sinensis*; (b) identify potential associations between overwintering ability and environmental factors such as latitude and hardiness zone at origin; (c) identify molecular markers associated with overwintering ability, dissect their allelic effects, and explore the potential of marker-assisted selection using these detected markers; and (d) assess the potential of genomic prediction for overwintering ability in *M. sinensis* using tens of thousands of markers.

## 2 | MATERIALS AND METHODS

### 2.1 | Plant materials and experimental design

We studied the overwintering ability of 565 *Miscanthus* genotypes, including 561 *M. sinensis*, 3 closely related *M. floridulus*, and the M×g '1993–1780' control. Hereafter, we refer all 561 *M. sinensis* and 3 closely related *M. floridulus* as *M. sinensis* since Clark et al. (2014) found that the *M. floridulus* were part of the southeastern China *M. sinensis* group. These 564 genotypes were previously assigned to the following eight genetic groups (Clark et al., 2014, 2015): US ornamental cultivars (76), US naturalized populations (38), Southern Japan (28), Northern Japan (83), Korea/N China (154), Sichuan Basin

**TABLE 1** Field trials of a *Miscanthus sinensis* diversity panel and controls planted at five locations and evaluated for overwintering ability. Field trials were established at each location in 2012, and three additional field trials were established at UI, CSU, and KNU in 2013

Location	Lat	Long	Hardiness zone <sup>a</sup>	Establishment year	Rep	N <sub>total</sub>	N <sub>sub</sub> 2013	N <sub>sub</sub> 2014	N <sub>sub</sub> 2015
UI	40.1	−88.2	5	2012	4	142	140	111	
				2013	4	164		145	
CSU	40.7	−105.0	5	2012	3	142	120	54	
				2013	4	164		140	
HU	43.1	141.4	7	2012	4	565	512	488	
KNU	37.9	127.8	6	2012	3	565	244	186	
				2013	3	565		424	366
NEF	42.1	−82.6	6	2012	4	565	561	471	

Note. UI = University of Illinois at Urbana-Champaign, Urbana, IL, USA; CSU = Colorado State University, Fort Collins, CO, USA; HU = Hokkaido University, Sapporo, Japan; KNU = Kangwon National University, Chuncheon, South Korea; NEF = New Energy Farms, Leamington, ON, Canada.

N<sub>total</sub>: Number of accessions planted, including the control Mxg '1993–1780'.

N<sub>sub</sub>: Number of *M. sinensis* genotypes that were alive in the autumn of the previous year and thus included in GWA analysis of overwintering ability.

<sup>a</sup><http://www.plantmaps.com>.

(25), Yangtze-Qinling (73), Southeastern China plus tropical (87). Six of the aforementioned genetic groups were previously determined based on discriminant analysis of principal components (DAPC; Jombart, Devillard, & Balloux, 2010) and by the software Structure (Falush, Stephens, & Pritchard, 2003); the US ornamental cultivars and US naturalized populations were found to be derived from southern Japan but were labeled independently to denote their provenance (Clark et al., 2014, 2015). Moreover, about half of the accessions in the US ornamental cultivars group have  $\leq 30\%$  ancestry from diploid *M. sacchariflorus*, presumably the result of past efforts by ornamental grass breeders in Germany to increase winter hardiness and obtain earlier flowering (Clark et al., 2014). Detailed information about these 564 *M. sinensis* accessions is listed in Data S1. All accessions were maintained as clonal stock plants in pots in a greenhouse at New Energy Farms in Ontario, Canada and vegetatively propagated. Ramets of each accession were distributed to each of five field trial locations during the winter of 2012.

In the spring of 2012, field trials were planted at two locations in East Asia (Table 1, Figure 1; HU = Hokkaido University, Sapporo, Japan; KNU = Kangwon National University, Chuncheon, South Korea) and three in North America (UI = University of Illinois at Urbana-Champaign, Urbana, IL, USA; CSU = Colorado State University, Fort Collins, CO, USA; NEF = New Energy Farms, Leamington, ON, Canada). Additional field trials were planted in 2013 at three locations (UI, CSU, and KNU; Table 1). Thus, a total of eight field trials were conducted. Nearly a full complement of the entire panel of 564 *M. sinensis* genotypes was planted in each of the trials at HU, KNU, and NEF. However, due to U.S. quarantine restrictions, only a subset of *M. sinensis* genotypes was planted in the field trials at UI and CSU (141 in 2012 trial, 163 in 2013 trial). Each field trial was arranged in a randomized complete block design with four replications except for KNU, where only one block included the entire panel and two additional small blocks included a same

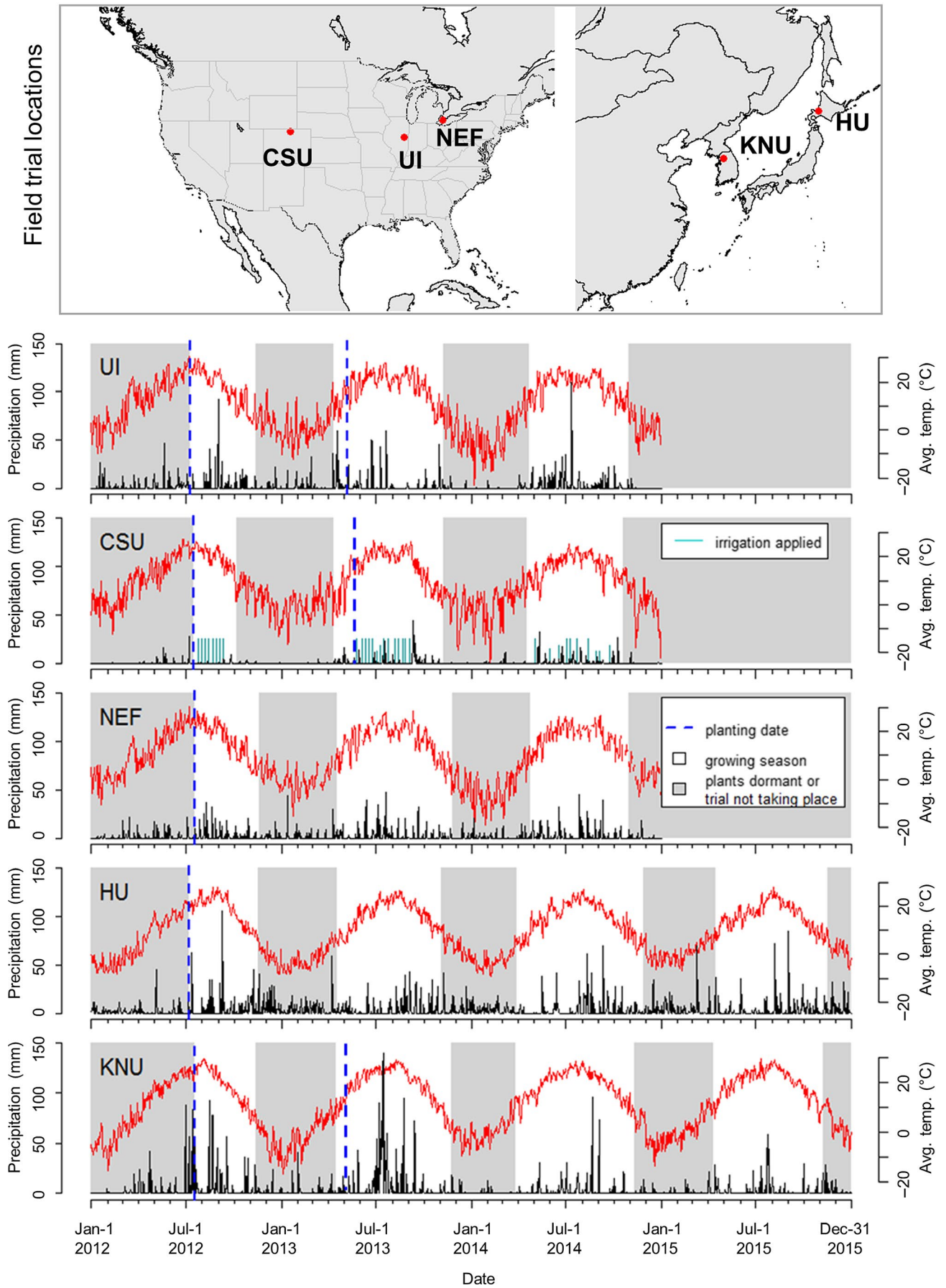
subset of 141 genotypes due to space constraints. Each plot contained one plant. Spacing between and within rows was 1.5 m. In each trial, irrigation was applied as needed during the first year to ensure good establishment and discontinued in subsequent years. The USDA hardiness zones of the field locations ranged from 5 to 7 (Table 1), representing a range of low temperature stresses during the winter (30-year average annual minimum temperature from  $-26.1^{\circ}\text{C}$  to  $-12.2^{\circ}\text{C}$ ).

## 2.2 | Phenotypic data collection

Overwintering ability was calculated as: 0, plant was alive in previous year's autumn but was dead in current year's spring; 1, plant was alive in previous year's autumn and also regrew in current year's spring. For each trial location, data on air temperature (daily mean, maximum, and minimum) and precipitation were compiled from nearby weather stations (Figure 1). For the IL location, daily records of soil temperature at 10 cm below the surface were also obtained. Overwintering ability of each plant in each of the eight field trials was evaluated for the first and second winters post-establishment. For the 2012 trials at all five locations, overwintering ability was evaluated after the 2012–2013 and 2013–2014 winters. For the 2013 trials at UI, CSU, and KNU, overwintering ability was evaluated for 2013–2014 winter only at UI and CSU, and for both the 2013–2014 and 2014–2015 winters at KNU. In summary, a total of 14 overwintering ability evaluations were performed for these eight field trials across five locations.

## 2.3 | Statistical analysis of phenotypic data

Analyses of variance (ANOVAs) were conducted in SAS 9.4 procedure MIXED (SAS Institute Inc., Cary, NC, USA) to determine how overwintering ability during the first and



**FIGURE 1** Environmental conditions during multiyear evaluations of a *Miscanthus sinensis* diversity panel at five field trial locations. The five locations included three in North America (University of Illinois at Urbana-Champaign = UI, Colorado State University = CSU, New Energy Farms in Leamington, Ontario = NEF) and two in Asia (Hokkaido University = HU and Kangwon National University = KNU). Planting date at each location was highlighted using dashed vertical blue lines (UI, CSU, and KNU had two trials planted in consecutive years). Red lines indicate daily average air temperature in °C, black lines represent daily precipitation in mm, gray-shaded areas indicate when plants were dormant, or field trial was not performed, and unshaded areas represent the growing season. For CSU trials, irrigation was applied as indicated by cyan solid bars

second winters was each affected by genetic group, genotype, location, year of establishment, number of growing seasons, and their interactions using the mixed linear model:

$$\begin{aligned}
 OWA_{ijklm} = & \mu + DAPC_i + G(DAPC)_{ij} + L_k + \\
 & Y_l + S_m + B(LY)_{klm} + \\
 & DAPC * L_{ik} + DAPC * Y_{il} + \\
 & DAPC * S_{im} + G(DAPC) * L_{ijk} + \\
 & G(DAPC) * Y_{ijl} + G(DAPC) * S_{ijm} + \epsilon_{ijklmn}
 \end{aligned}
 \quad (1)$$

where *OWA* represents overwintering ability,  $\mu$  is the grand mean, *DAPC* represents the genetic groups determined in Clark et al. (2014, 2015), *G(DAPC)* is genotype nested in genetic group, *L* is location, *Y* is the year the trial was planted, *S* is number of growing season, *B(LY)* is block nested in location by year (i.e., in field trial), *DAPC\*L* represents genetic group by location interaction, *DAPC\*Y* represents genetic group by year interaction, *DAPC\*S* represents genetic group by growing season interaction, *G(DAPC)\*L* represents genotype nested within genetic group by location interaction, *G(DAPC)\*Y* represents genotype nested within genetic group by year interaction, *G(DAPC)\*S* represents genotype nested within genetic group by growing season interaction, and  $\epsilon$  is a random error term. All model terms were set as fixed except for block. Traditionally, logistic regression is preferred over linear regression for dichotomous dependent variable modeling (i.e., overwintering ability in this study). However, it has been shown that, in most practical cases, linear regression can be as accurate as or even better than logistic regression, especially in high-dimensional data (Hellevik, 2009). To confirm the reliability of mixed linear model in analyzing binary overwintering ability data, we ran both logistic regression and a mixed linear model for significance tests. Results of the two tests were identical, and the correlation coefficient between the *p*-values of two testes was 0.99998 (Table S1). Therefore, this conformed to the conclusion of Hellevik (2009) that linear regression could be reliably used in analyzing dichotomous dependent variables. To estimate best linear unbiased predictors (BLUPs) of overwintering ability for each combination of winter, trial location, and year of establishment, ANOVAs were conducted using the following linear mixed model:

$$OWA_{ij} = \mu + G_i + B_j + \epsilon_{ij} \quad (2)$$

## 2.4 | BioClim analysis

In order to identify associations between overwintering ability and environmental factors, we first explored the relationship between overwintering ability and latitude and hardiness zone at origin. We further assessed the influence of climate variables on *Miscanthus* overwintering ability with BioClim data (Version 1.4; Hijmans, Cameron, Parra, Jones, & Jarvis, 2005; <http://www.worldclim.org/>). For BioClim analysis, we first ran a linear model by regressing first-winter overwintering ability BLUPs on the 19 climate variables (BIO1-19), which were all set as fixed effects, with the aim of identifying significant climate variables influencing overwintering ability. Then, we ran principal component analysis on the significant BioClim variables and regressed first-winter overwintering ability BLUPs on BioClim PC1.

## 2.5 | Marker development and missing data imputation

Restriction site-associated DNA sequencing (RAD-seq) was performed according to a previously described protocol (Clark et al., 2014). In brief, genomic DNA was digested with *MspI* and either *PstI*-HF or *NsiI*-HF (New England Biolabs). Digested DNA was then ligated to a barcoded adapter with a *PstI/NsiI* overhang and a universal adapter with an *MspI* overhang. Ninety-five barcoded samples were then pooled into one library, and 200–500 bp PCR products were selected on 2% agarose. Libraries were sequenced on an Illumina HiSeq 2500 at the Roy J. Carver Biotechnology Center at the University of Illinois. Nine *PstI* libraries from a previous study (Clark et al., 2014; data available at the NCBI Sequence Read Archive, accession SRP026347), as well as eight additional *PstI* libraries and 13 *NsiI* libraries (Clark et al., 2016; data available at NCBI, accession SRP063572), were included in the analysis. Because diploid *M. sinensis* has a large genome of ~5.3 Gb (Chae et al., 2014; Rayburn, Crawford, Rayburn, & Juvik, 2009), multiple sequencing runs can improve read depth and reduce missing data, which are especially important for correctly calling heterozygous loci in obligately outcrossing species such as this. Therefore, every individual in the study was included in at least two *PstI* libraries and two *NsiI* libraries.



Single-nucleotide polymorphisms (SNPs) were mined from all *M. sinensis* entries included in the field trial using the UNEAK pipeline (version 3.0 standalone; Lu et al., 2013) with an initial minimum call rate of 0.04 and a minimum minor allele frequency of 0.002. Three doubled haploid *M. sinensis* individuals (Głowacka, Kaczmarek, & Jeżowski, 2012; Swaminathan et al., 2012) were also included in the analysis to identify and remove any SNPs that appeared heterozygous and thus likely the result of paralogs. The dataset was then filtered to only include SNPs that had at least a 70% call rate in at least one of the genetic groups previously identified by discriminant analysis of principal components (Clark et al., 2014; Jombart et al., 2010) and a minor allele frequency of at least 0.05 in at least one of the nine groups, resulting in 70,327 SNPs retained, with an overall missing data rate of 39%. Imputation was performed with an estimation–maximization (EM) method based on relatedness (Poland, Brown, Sorrells, & Jannink, 2012) implemented in the R package rrBLUP (Endelman, 2011). To obtain genomic positions of SNPs for Manhattan plots and identify candidate genes, sequence tags were aligned to the *Sorghum bicolor* 3.0 reference genome (Paterson et al., 2009) using Bowtie2 (Langmead & Salzberg, 2012).

## 2.6 | Genome-wide association analysis

GWA analyses for overwintering ability were performed on the 564 *M. sinensis* genotypes (including the three *M. floridulus*) for each combination of growing season (i.e., winter) and field trial separately, for a total of 14 analyses. GWA analyses were performed using the multilocus mixed model (MLMM) implemented in R (Segura et al., 2012). MLMM adopts a forward–backward model selection procedure to potentially improve power (the ability to find true positives) and reduce false-positive rates. We allowed up to 10 forward selection and backward elimination steps for each GWA analysis. The optimal model was determined based on the extended Bayesian information criterion, defined as the largest model in which all cofactors have a *p*-value below the 5% Bonferroni-corrected threshold. To control for the confounding effects of cryptic relatedness and population structure, we incorporated into the MLMM model the additive relationship

matrix that was calculated using the EM imputation method in rrBLUP (Endelman, 2011; Poland et al., 2012) and the principal components of population structure calculated in R package adegenet (Jombart et al., 2010). Significant markers associated with overwintering ability were dissected for additive and dominance effects. In brief, additive effect was calculated as the phenotypic difference between two homozygous genotypic classes, and dominance effect was calculated as the phenotypic difference between the average of two homozygous genotypes and the heterozygous genotype. Then, for each *M. sinensis* accession, using all markers that were significantly associated with overwintering ability detected across 14 GWA analyses, the genotypic value of overwintering ability was calculated based on genotypic effects (additive and dominance), and breeding value was calculated based on only additive effects.

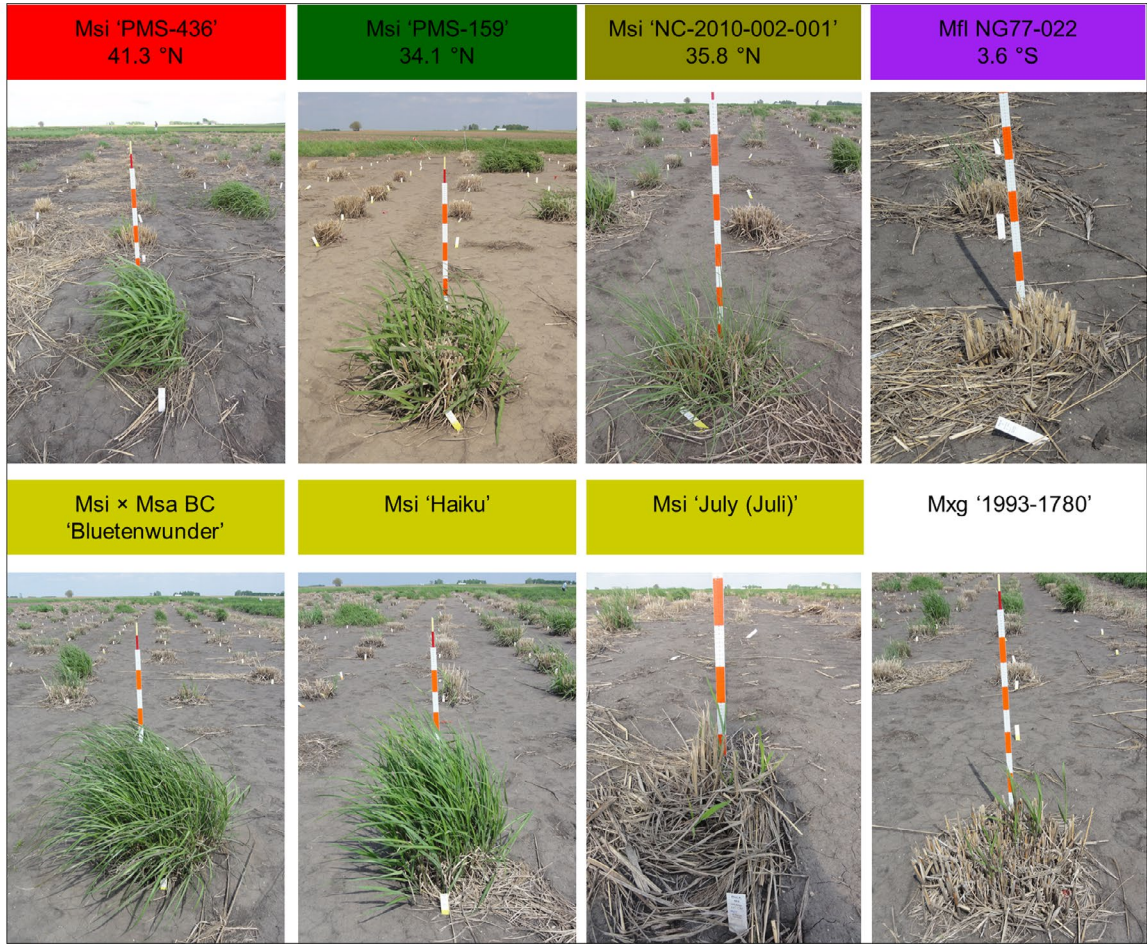
## 2.7 | Genomic prediction

Genomic prediction for first-winter overwintering ability was conducted in the R package rrBLUP (Endelman, 2011) based on the 70,327 SNPs. Predictive ability was assessed by a fivefold cross-validation strategy. Briefly, *M. sinensis* accessions were randomly split into five sets, of which four sets were used as training population to estimate marker effects, and the remaining set was used as validation population to calculate the genomic estimated breeding values (GEBVs). This process was repeated for each of the five sets. Thus, GEBVs were obtained for all *M. sinensis* accessions. Predictive ability (*r*) was defined as the Pearson's correlation between GEBVs and trait BLUPs (Daetwyler, Calus, Pong-Wong, de Los Campos, & Hickey, 2013). This fivefold cross-validation process was iterated 50 times. Mean and standard deviation of these 50 correlation coefficients were reported as predictive ability and its standard deviation.

As the presence of closely related individuals in both training and validation populations could artificially inflate prediction accuracies, it is recommended to control for population structure in designing the cross-validation scheme (Ly et al., 2013; Riedelsheimer et al., 2013; Spindel et al., 2015). To do so, three statistical methods were investigated

**FIGURE 2** Spring 2014 regrowth of *Miscanthus* accessions tested at five field trial locations, with three in North America (University of Illinois at Urbana-Champaign = UI, Colorado State University = CSU, New Energy Farms in Leamington, Ontario = NEF) and two in Asia (Hokkaido University = HU and Kangwon National University = KNU). Field trials were established at each location in 2012, and three additional field trials were established at UI, CSU, and KNU in 2013. (a) Photos of individual plants showing range of first-winter survival and hardiness at UI. *M. sinensis* (Msi) genotype names are colored based on DAPC groups determined in Clark et al. (2014, 2015). Msi 'PMS-438' is from Korea/N China group (red); Msi 'PMS-159' is from Yangtze-Qinling group (dark green); Msi 'NC-2010-002-001' is from US naturalized populations group (yellow4); *M. floridulus* (Mfl) 'NG77-022' is from Southeastern China plus tropical group (purple); 'Bluetenwunder' is an Msi × *M. sacchariflorus* (Msa) backcross hybrid and is from US ornamental cultivar (yellow3), and Msi 'Haiku' and 'July (Juli)' are pure *M. sinensis* ornamental cultivars (yellow3). Latitude of origin of genotype is shown as available. The control genotype *M. ×giganteus* (M×g) '1993-1780' is also shown. (b) Overview photos of UI, NEF, HU, KNU fields. Note the severe overwintering losses in the UI field trial

(a)



(b)



in genomic prediction for first-winter overwintering ability across the whole panel ( $N = 564$ ). These three statistical methods differed in controlling the previously determined genetic structure in this *M. sinensis* diversity panel. The *M. sinensis* accessions were previously assigned to one of the following eight groups: Southern Japan (28), Northern Japan (83), Korea/N China (154), Yangtze-Qinling (73), Sichuan Basin (25), Southeastern China plus tropical (87), US ornamental cultivars (76), and US naturalized populations (38). For Method 1, as a control comparison, population structure was not accounted for, and best linear unbiased predictor (BLUP) values were calculated using the following equation:

$$y = \mu + G + L + Y + B(LY) + G * L + G * Y + \varepsilon \quad (3)$$

where  $y$  is the first-winter *OWA* raw data,  $\mu$  is grand mean,  $G$  is genotype,  $L$  is location,  $Y$  is year of field trial establishment,  $B(LY)$  is block nested within location by year (i.e., in field trial),  $G * L$  is genotype by location interaction,  $G * Y$  is genotype by year interaction, and  $\varepsilon$  is random error. All model terms were set as random. BLUPs of the  $G$  term were used in genomic prediction.

For Method 2, the BLUPs obtained from Equation 3 were used as the independent variables, and then, we fitted a mixed model by including DAPC group as fixed effect:

$$y = \mu + DAPC + \varepsilon \quad (4)$$

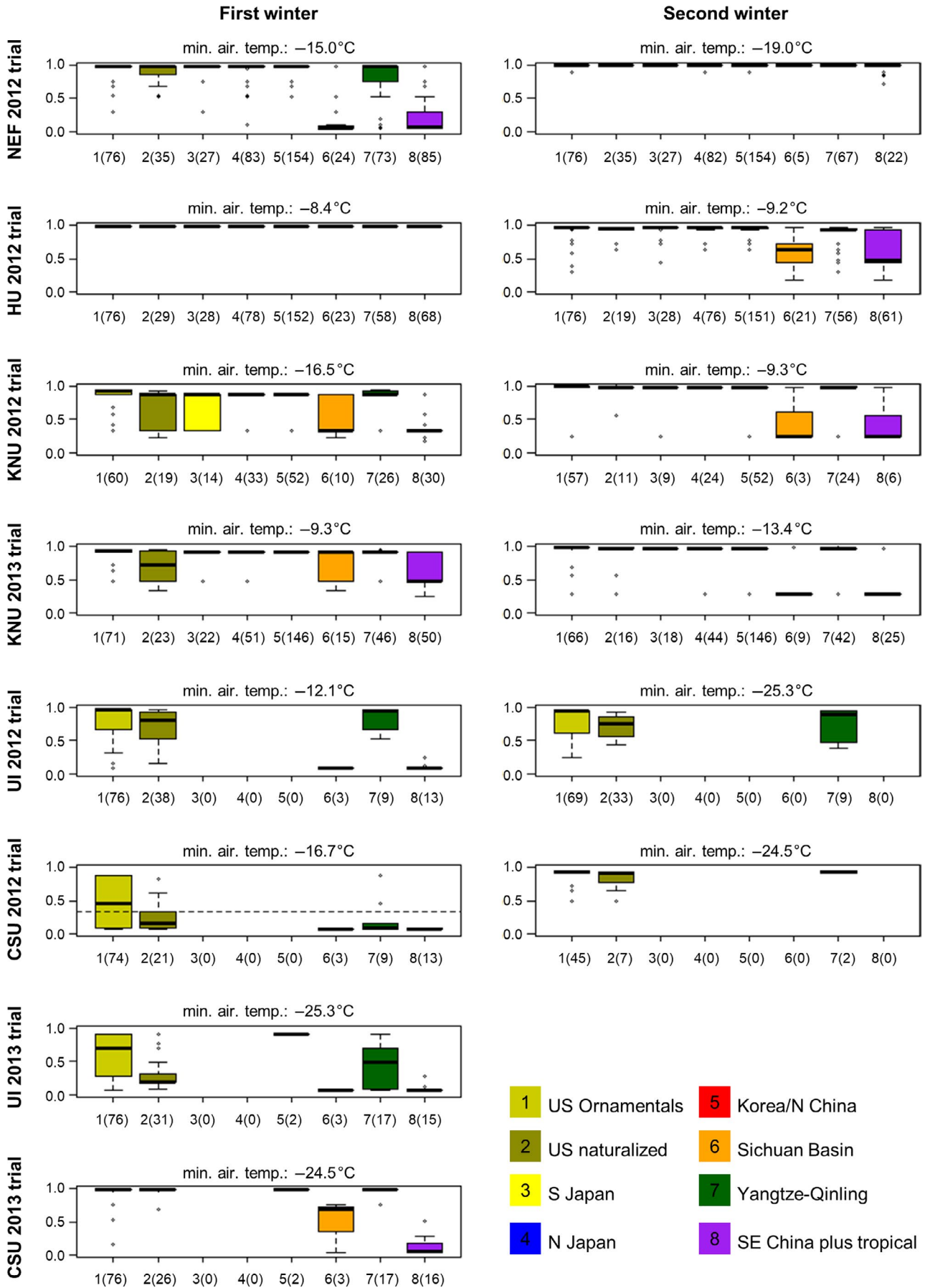
where  $y$  is the genotype BLUPs calculated from Equation 3,  $\mu$  is grand mean,  $DAPC$  represents the previously determined

Term <sup>a</sup>	df	Mean squares	F value	Pr (>F)
Block (location × year)	23	1.85	40.65	<0.0001
DAPC	7	14.88	326.31	<0.0001
Genotype (DAPC)	556	0.22	4.74	<0.0001
Location	4	16.81	368.55	<0.0001
Year	1	0.37	8.05	0.0046
Growing season	1	11.21	245.71	<0.0001
DAPC × location	22	3.39	74.40	<0.0001
DAPC × year	7	0.09	1.87	0.0698
DAPC × growing season	7	2.92	64.05	<0.0001
Genotype (DAPC) × location	1,008	0.16	3.45	<0.0001
Genotype (DAPC) × year	258	0.10	2.30	<0.0001
Genotype (DAPC) × growing season	536	0.09	1.87	<0.0001
Error	7,707	0.05		

<sup>a</sup> $OWA_{ijklm} = \mu + DAPC_i + G(DAPC)_{ij} + L_k + Y_l + S_m + B(LY)_{klm} + DAPC * L_{ik} + DAPC * Y_{il} + DAPC * S_{im} + G(DAPC) * L_{ijk} + G(DAPC) * Y_{ijl} + G(DAPC) * S_{ijm} + \varepsilon_{ijklmn}$ , where *OWA* represents overwintering ability,  $\mu$  is the grand mean, *DAPC* is genetic group, *G(DAPC)* is genotype nested in genetic group, *L* is location, *Y* is year, *S* is growing season, *B(LY)* is block nested in location by year (field trial), *DAPC \* L* represents genetic group by location interaction, *DAPC \* Y* represents genetic group by year interaction, *DAPC \* S* represents genetic group by growing season interaction, *G(DAPC) \* L* represents genotype nested within genetic group by location interaction, *G(DAPC) \* Y* represents genotype nested within genetic group by year interaction, *G(DAPC) \* S* represents genotype nested within genetic group by growing season interaction, and  $\varepsilon$  is random error. All model terms were set as fixed except for block.

**TABLE 2** Analysis of variance, testing the effects of genetic group, genotype within genetic group, location, year of establishment, number of growing seasons, and their interactions on overwintering ability during the first and second winter of field trials for a *Miscanthus sinensis* diversity panel evaluated at three locations in North America (UI = University of Illinois at Urbana-Champaign, Urbana, IL, USA; CSU = Colorado State University, Fort Collins, CO, USA; NEF = New Energy Farms, Leamington, ON, Canada) and two locations in Asia (HU = Hokkaido University, Sapporo, Japan; KNU = Kangwon National University, Chuncheon, South Korea). Field trials were established at each location in 2012, and three additional field trials were established at UI, CSU, and KNU in 2013

**FIGURE 3** Box plots of overwintering ability during the first winter (establishment year) and second winter for a *Miscanthus sinensis* diversity panel at each of five field trial locations. Each box plot represents one field trial location, with three in North America (University of Illinois at Urbana-Champaign = UI, Colorado State University = CSU, New Energy Farms in Leamington, Ontario = NEF) and two in Asia (Hokkaido University = HU and Kangwon National University = KNU). Field trials were established at each location in 2012, and three additional field trials were established at UI, CSU, and KNU in 2013. Genotypes are shown by DAPC group (X-axis; Clark et al., 2014, 2015). The eight genetic groups included US ornamental cultivars (1), US naturalized populations (2), Southern Japan (3), Northern Japan (4), Korea/N China (5), Sichuan Basin (6), Yangtze-Qinling (7), Southeastern China plus tropical (8). On the X-axis, the number of genotypes with data for each DAPC group in each location is indicated in parentheses after the group number. The Y-axis represents overwintering ability ranging from 0 to 1, where 0 indicates plants failed to survive the winter and 1 means plants survived the winter. Overwintering ability of *M. x giganteus* '1993–1780' is shown as horizontal dashed line (only shown where <100%). Boxes span from the first to third quartile for each group. Whiskers extend to the minimum and maximum values or to the first and third quartile  $\pm 1.5$  times the box length, respectively, whichever is shorter. Points indicate genotypes with values outside the range spanned by the whiskers. Internal bar shows the median value



genetic groups in this *M. sinensis* diversity panel (Clark et al., 2014,2015). The residuals from Equation 4 were used as phenotypic data in genomic prediction.

For Method 3, we fitted a model by nesting genotype within DAPC group using the following equation:

$$y = \mu + DAPC + G(DAPC) + L + Y + B(LY) + G(DAPC) * L + G(DAPC) * Y + \epsilon \quad (5)$$

where model terms were defined as in Equations 3 and 4, except that  $G(DAPC)$  represents genotype nested within  $DAPC$  group. All model terms were set as random effect except for  $DAPC$ , which was set as fixed effect. BLUPs of the  $G(DAPC)$  term were used in genomic prediction.

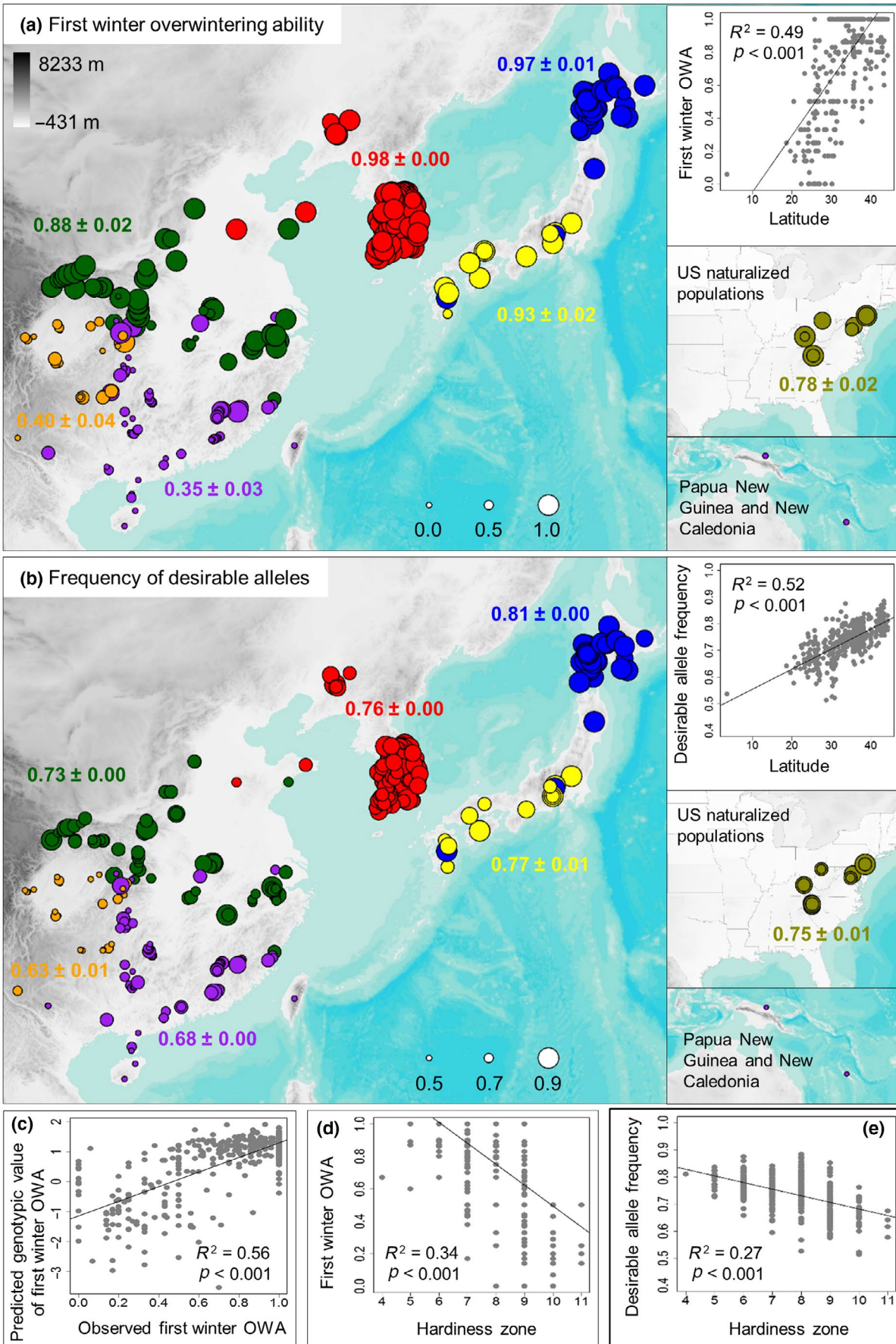
Additionally, we explored the effectiveness of genomic prediction for first-winter OWA within each genetic group. Within-genetic-group genomic prediction was performed for each of the following *M. sinensis* genetic groups: Southern Japan ( $N = 142$ ), Northern Japan ( $N = 83$ ), Korea/N China ( $N = 154$ ), Yangtze-Qinling ( $N = 73$ ), and Southeastern China plus tropical ( $N = 87$ ). Within each of these five genetic groups, BLUPs were calculated using Equation 3, and genotype BLUPs were used in genomic prediction. The Sichuan Basin group ( $N = 25$ ) was not included in the within-genetic-group genomic prediction analysis due to its small size. For the within-genetic-group genomic prediction analysis, the Southern Japan group studied was a composite of its subset groups, US naturalized population ( $N = 38$ ), US ornamental cultivars ( $N = 76$ ), and Southern Japan group ( $N = 28$ ; Clark et al., 2014,2015); DAPC analysis had included these genotypes in a single group and keeping them together allowed for a larger population size, which is typically advantageous for genomic prediction.

### 3 | RESULTS

In the *M. sinensis* diversity panel, highly significant differences in overwintering ability were observed among genetic groups, genotypes within genetic groups, locations, year of

establishment, number of growing seasons, and their interactions except for genetic group by year interaction (Table 2, Figures 2, 3, 4a). The control genotype M×g ‘1993–1780’ survived in all field trials except for the first winter of the CSU 2012 trial (Figure 3). Among the six *M. sinensis* genetic groups from Asia, first-winter overwintering ability, averaged over all of the field trial locations, was highest for the Korea/N China group (0.98; Figure 4a, Table 3), and relatively high for the Northern Japan group (0.96), the Southern Japan group (0.93), and the Yangtze-Qinling group (0.88). The Korea/N China group showed consistently superior overwintering ability at each of the field trial locations (0.86–0.99). The lowest first-winter overwintering ability was observed for the Southeastern China plus tropical group (0.36), followed by the Sichuan Basin group (0.40), which was consistent with their natural adaptation to low-latitude environments that have mild winters (Figure 4a, Table 3). Such differences in overwintering ability among different genetic groups were further delineated by a strong association between first-winter overwintering ability and latitude of origin ( $R^2 = 0.49$ , Figure 4a). Moreover, a similar association was also observed between first-winter overwintering ability and hardiness zones from which *M. sinensis* accessions were collected ( $R^2 = 0.34$ , Figure 4d), indicating *M. sinensis* accessions collected from high latitude locations with cold winters had higher rates of overwintering than accessions from southern locations with mild winters. In addition, BioClim analysis indicated that BIO1 (Annual mean temperature), BIO6 (Minimum temperature of coldest month), BIO10 (Mean temperature of warmest quarter), and BIO11 (Mean temperature of coldest quarter) were significant at  $\alpha = 0.05$  (Table S2). BioClim PC1 explained 92.4% of total variance among these four significant variables. Intriguingly, a similarly strong association between first-winter overwintering ability and BioClim PC1 was observed ( $R^2 = 0.50$ ; Table S2). Nevertheless, some exceptional individuals from the Southeastern China plus tropical group and the Sichuan Basin group survived first and second winters at HU, KNU, and/or NEF (Figure 3, Table 3).

**FIGURE 4** First-winter overwintering ability (OWA) and frequency of desirable alleles of each *Miscanthus sinensis* accession. Genotypes (filled circles) are shown by DAPC groups (color of fill; Clark et al., 2014, 2015). Seven genetic groups included Northern Japan (blue), Southern Japan (yellow), Korea/N China (red), Yangtze-Qinling (dark green), Sichuan Basin (orange), Southeastern China plus tropical (purple), and US naturalized populations (yellow4) are shown. The US ornamental cultivars genetic group is not shown due to lack of geographical source information. (a) First-winter OWA for each *M. sinensis* accession across all field trials. Note that high latitude genotypes generally had higher overwintering ability than those from low latitude. The inset regression plot shows a strong positive association between first-winter OWA and latitude of origin ( $R^2 = 0.49$ ,  $p < 0.001$ ). (b) Frequency of desirable alleles across the 73 markers detected for OWA in each *M. sinensis* genotype. Note that high latitude genotypes generally had a higher frequency of advantageous alleles than those from low latitude. The inset regression plot shows a strong association between frequency of desirable alleles and latitude of origin ( $R^2 = 0.52$ ,  $p < 0.001$ ). (c) The predicted genotypic values of first-winter OWA (Y-axis) based on the 73 markers, regressed on the observed first-winter OWA (X-axis) of 564 *M. sinensis* accessions. A strong association between these two variables was observed ( $R^2 = 0.56$ ,  $p < 0.001$ ). (d) *M. sinensis* first-winter OWA was associated with the hardiness zone of its origin ( $R^2 = 0.34$ ,  $p < 0.001$ ). (e) The frequency of desirable alleles in each *M. sinensis* accession was associated with the hardiness zone of its origin ( $R^2 = 0.27$ ,  $p < 0.001$ )



Among the five field trial locations, first winter overwintering ability was typically lowest at the two US locations, UI (0.40–0.49, two trials) and CSU (0.22–0.76, two trials), though these field trial locations primarily tested entries from the US ornamental and US naturalized groups, and included few of the entries from Asia due to quarantine regulations that delayed their importation beyond the timeframe of this study (Table 3). In contrast to the US locations, first-winter overwintering was typically higher at HU (0.99), KNU (0.69–0.82, two trials), and NEF (0.73), even when the comparison was limited to just the US ornamental and US naturalized groups (Table 3). The lower overwintering rates observed at the US field trial locations relative to the other sites was consistent with a combination of lower minimum air temperatures and/or less snow cover for the US trials (Figure 1, Table 3). The winter of 2013–2014 was especially cold at the US sites, with minimum air temperatures of  $-25.3^{\circ}\text{C}$  at UI and  $-24.5^{\circ}\text{C}$  at CSU, whereas the NEF site was moderated by Lake Erie and reached only  $-19^{\circ}\text{C}$ , and the two locations in Asia were warmer still at about  $-9^{\circ}\text{C}$  (Figure 1, Table 3).

At a given location, overwintering ability (i.e., the proportion of plants that were alive the previous growing season and also survived the subsequent winter) was typically lower after the first winter than after the second winter (Table 3), because the least hardy plants typically died during the first winter, leaving the more adapted genotypes and better established plants to be challenged by the second winter. For example, at NEF in Ontario, overwintering ability was 0.99 after the second winter (Table 3; min. air temp.:  $-19.0^{\circ}\text{C}$ ) but was only 0.73 after the first winter (min. air temp.:  $-15.0^{\circ}\text{C}$ ), indicating nearly all the plants that survived the first winter also survived the second colder winter. However, a notable exception to greater overwintering after the second winter was observed for the 2012 trial at HU in Japan, in which nearly all plants from each of the eight genetic groups survived the first winter (Table 3; overwintering ability: 0.99; min. air temp.:  $-8.4^{\circ}\text{C}$ ), whereas after the second winter, the overall overwintering ability was only 0.80 (min. air temp.:  $-9.2^{\circ}\text{C}$ ). In particular, two genetic groups, including Sichuan Basin and Southeastern China plus tropical, had substantial losses (0.60–0.61 OWA) during the second winter at HU. This atypical outcome at HU was likely the result of high and early snow cover during the first winter, which insulated the plants from low air temperatures (Figure 1).

Genome-wide association analyses across the five field trial locations detected a total of 73 significant marker–trait associations for overwintering ability (Figure S1), of which 42 could be aligned to physical positions on *S. bicolor* genome version 3.0 (Table 4). For the first-winter overwintering ability, additive effects among these 73 SNPs ranged from 0.01 to 0.42, and dominance effects ranged from  $-0.59$  to 0.44 (Data S2). Among the 564 *M. sinensis* accessions, the

frequency of desirable alleles (up to two desirable alleles per locus per individual) over all 73 overwintering ability loci ranged from 0.51 (accession PMS-044, Sichuan Basin genetic group) to 0.88 (accession JM0047.002, Northern Japan genetic group) (Data S2, Figure 4b). The predicted genotypic values for overwintering ability based on the 73 SNPs accounted for 56% of the variation observed for first-winter survival among *M. sinensis* accessions averaged over the five field trial locations (Figure 4c). Similarly, the breeding values for overwintering ability explained 55% of the variation for observed first-winter overwintering ability (Figure S2). Moreover, a strong association was observed between frequency of desirable alleles and latitude ( $R^2 = 0.52$ , Figure 4b), and a moderate association was observed between frequency of desirable alleles and hardiness zones ( $R^2 = 0.27$ , Figure 4e), indicating that *M. sinensis* accessions collected from high latitude locations with cold winters had more alleles for overwintering than accessions from southern locations with mild winters. Notably, observed first-winter overwintering ability was also significantly correlated with the frequency of desirable alleles ( $R^2 = 0.35$ , Figure S3).

Genomic prediction analyses for first-winter overwintering ability resulted in high and stable estimates, as indicated by the small standard error and range of genomic prediction abilities (Table 5). For Method 1, which did not control for population structure, genomic prediction accuracy for the entire panel averaged 0.81. In contrast, for Methods 2 and 3, which controlled population structure (i.e., DAPC groups), average genomic predictive ability was 0.73 and 0.71, respectively. However, genomic prediction within DAPC group varied greatly by group (Table 5). The Southeastern China plus tropical group, and the Southern Japan group had moderate genomic predictive abilities, with average values of 0.34 and 0.48, respectively. For the Yangtze-Qinling group, genomic predictive ability was low, with an average of 0.04. For the Northern Japan and Korea/N China groups, genomic predictions could not be obtained because nearly all accessions in these two groups survived the first winter (Figure 4a, Table 3), resulting in zero variance among genotypes for this trait.

## 4 | DISCUSSION

### 4.1 | High genetic diversity within *M. sinensis* for overwintering ability: Implications for breeding

Like most prior studies of overwintering ability in *Miscanthus* (Clifton-Brown & Lewandowski, 2000; Clifton-Brown et al., 2001; Yan et al., 2012), the current study focused on survival over the first winter after planting. However, this study extended our understanding of overwintering ability in *M. sinensis* by having phenotyped and genotyped a large germplasm

**TABLE 3** Average overwintering ability (OWA; proportion survival) and associated standard error during the first winter (establishment year) and second winter for genetic groups (identified previously by Clark et al., 2014, 2015) in a *Miscanthus sinensis* diversity panel, evaluated at three locations in North America (UI = University of Illinois at Urbana-Champaign, Urbana, IL, USA; CSU = Colorado State University, Fort Collins, CO, USA; NEF = New Energy Farms, Leamington, ON, Canada) and two locations in Asia (HU = Hokkaido University, Sapporo, Japan; KNU = Kangwon National University, Chuncheon, South Korea). Field trials were established at each location in 2012, and three additional field trials were established at UI, CSU, and KNU in 2013

Field trial	Winter <sup>a</sup>	Min. air temp. during winter (°C)	Genetic groups										Overall average overwintering ability by field trial
			US ornamental cultivars	US naturalized populations	Southern Japan	Northern Japan	Korea/N China	Sichuan Basin	Yangtze-Qinling	Southeastern China plus tropical			
UI 2012 trial	1	-12.1	0.76 ± 0.03	0.69 ± 0.04				0.09 ± 0.00	0.80 ± 0.06	0.10 ± 0.01			0.49 ± 0.03
	2	-25.3	0.77 ± 0.03	0.70 ± 0.03					0.72 ± 0.08				0.73 ± 0.02
UI 2013 trial	1	-25.3	0.56 ± 0.04	0.31 ± 0.04			0.91 ± 0.00	0.07 ± 0.00	0.46 ± 0.08	0.08 ± 0.01			0.40 ± 0.03
	1	-16.7	0.48 ± 0.04	0.26 ± 0.05				0.07 ± 0.00	0.22 ± 0.09	0.08 ± 0.00			0.22 ± 0.03
CSU 2013 trial	2	-24.5	0.86 ± 0.02	0.83 ± 0.06					0.94 ± 0.01				0.87 ± 0.02
	1	-24.5	0.97 ± 0.01	0.98 ± 0.01			0.99 ± 0.00	0.49 ± 0.18	0.98 ± 0.01	0.13 ± 0.03			0.76 ± 0.02
NEF 2012 trial	1	-15.0	0.96 ± 0.01	0.90 ± 0.02	0.92 ± 0.04	0.95 ± 0.01	0.96 ± 0.01	0.18 ± 0.05	0.83 ± 0.03	0.21 ± 0.03			0.73 ± 0.01
	2	-19.0	0.99 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	0.99 ± 0.00	0.99 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	0.96 ± 0.02			0.99 ± 0.00
HU 2012 trial	1	-8.4	0.99 ± 0.00	0.99 ± 0.00	0.99 ± 0.00	0.99 ± 0.00	0.99 ± 0.00	0.99 ± 0.00	0.99 ± 0.00	0.99 ± 0.00			0.99 ± 0.00
	2	-9.2	0.92 ± 0.01	0.92 ± 0.02	0.94 ± 0.02	0.95 ± 0.01	0.95 ± 0.01	0.60 ± 0.05	0.89 ± 0.02	0.61 ± 0.03			0.80 ± 0.01
KNU 2012 trial	1	-16.5	0.84 ± 0.02	0.64 ± 0.06	0.69 ± 0.07	0.77 ± 0.04	0.86 ± 0.01	0.49 ± 0.08	0.85 ± 0.03	0.37 ± 0.03			0.69 ± 0.02
	2	-9.3	0.98 ± 0.01	0.95 ± 0.04	0.90 ± 0.08	0.98 ± 0.00	0.94 ± 0.02	0.48 ± 0.20	0.95 ± 0.03	0.41 ± 0.11			0.82 ± 0.01
KNU 2013 trial	1	-9.3	0.87 ± 0.02	0.73 ± 0.04	0.83 ± 0.04	0.85 ± 0.02	0.91 ± 0.00	0.73 ± 0.06	0.87 ± 0.02	0.66 ± 0.03			0.82 ± 0.01
	2	-13.4	0.93 ± 0.02	0.86 ± 0.06	0.97 ± 0.00	0.95 ± 0.02	0.96 ± 0.01	0.36 ± 0.07	0.87 ± 0.04	0.44 ± 0.06			0.81 ± 0.01
Over all locations	1		0.83 ± 0.02	0.73 ± 0.03	0.93 ± 0.03	0.96 ± 0.01	0.98 ± 0.00	0.40 ± 0.04	0.88 ± 0.02	0.36 ± 0.03			0.80 ± 0.01
	2		0.92 ± 0.01	0.90 ± 0.02	0.95 ± 0.04	0.98 ± 0.01	0.98 ± 0.00	0.44 ± 0.07	0.92 ± 0.02	0.46 ± 0.05			0.87 ± 0.01
Over all locations	1&2		0.77 ± 0.02	0.64 ± 0.03	0.88 ± 0.04	0.94 ± 0.01	0.97 ± 0.01	0.20 ± 0.05	0.80 ± 0.03	0.18 ± 0.03			0.73 ± 0.01

<sup>a</sup>Year 1 OWA indicates the proportion of genotypes that were alive in the autumn of the first year that regrew in the spring of the second year (i.e., survived the first winter). Year 2 OWA indicates the proportion of genotypes that were alive in the autumn of the second year that regrew in the spring of the third year (i.e., survived the second winter).



**TABLE 4** Markers with significant associations for overwintering ability (proportion survival) in a *Miscanthus sinensis* diversity panel, evaluated at three locations in North America (UI = University of Illinois at Urbana-Champaign, Urbana, IL, USA; CSU = Colorado State University, Fort Collins, CO, USA; NEF = New Energy Farms, Leamington, ON, Canada) and two locations in Asia (HU = Hokkaido University, Sapporo, Japan; KNU = Kangwon National University, Chuncheon, South Korea). Field trials were established at each location in 2012, and three additional field trials were established at UI, CSU, and KNU in 2013

Chr. <sup>a</sup>	Position <sup>b</sup>	Trial <sup>c</sup>	Winter <sup>d</sup>	<i>Miscanthus</i> SNP <sup>e</sup>	<i>P</i> value <sup>f</sup>	Nearest <i>S. bicolor</i> gene	Position relative to the gene <sup>g</sup>	Arabidopsis gene	Gene ontology	Gene symbol
1	9 158 264	KNU 2013	1	UJIMiscanthus096890	2.48E-11	Sobic.001G117600	CDS	AT4G22760	Tetratricopeptide repeat (TPR)-like superfamily protein	
1	73 182 406	UI 2013	1	UJIMiscanthus017076	1.85E-07	Sobic.001G455800	3' UTR -> 1,171 bp	AT4G13780	Methionine-tRNA ligase	
1	73 966 201	KNU 2013	2	UJIMiscanthus104319	2.07E-10	Sobic.001G465900	3' UTR -> 6,387 bp	AT3G03900	Adenosine-5'-phosphosulfate (APS) kinase 3	APK3
1	77 425 506	NEF 2012	2	UJIMiscanthus111533	2.84E-116	Sobic.001G506100	3,342 bp -> 5' UTR	AT4G28650	Leucine-rich repeat transmembrane protein kinase	
2	3 043 560	CSU 2012	2	UJIMiscanthus015113	1.44E-22	Sobic.002G033200	CDS	AT1G30540	Actin-like ATPase	
2	3 279 281	NEF 2012	2	UJIMiscanthus118365	7.63E-09	Sobic.002G035100	Upstream			
2	3 406 536	KNU 2012	2	UJIMiscanthus118366	2.10E-07	Sobic.002G036150	Downstream	AT1G30135	Jasmonate-zim-domain protein 8	JAZ8
2	6 522 791	NEF 2012	2	UJIMiscanthus118367	3.13E-24	Sobic.002G065900	CDS	AT5G59190	Subtilase family protein	
2	22 943 245	NEF 2012	1	UJIMiscanthus118368	4.24E-09	Sobic.002G141300	Downstream	AT5G67300	myb domain protein r1	MYBRI
2	77 204 376	KNU 2012	2	UJIMiscanthus020585	7.49E-92	Sobic.002G425600	3' UTR	AT2G41070	Basic-leucine zipper (bZIP) transcription factor	ATBZIP12, DPBF4, EEL
3	1 603 696	KNU 2013	1	UJIMiscanthus118369	1.88E-09	Sobic.003G018000	Downstream			
3	7 227 095	KNU 2012	1	UJIMiscanthus008958	6.80E-07	Sobic.003G084100	CDS			
3	51 321 773	NEF 2012	2	UJIMiscanthus118370	1.02E-37	Sobic.003G191750	0.01 Mb downstream	AT4G36020	Cold shock domain protein 1	CSDP1
3	64 085 542	HU 2012	2	UJIMiscanthus118371	3.26E-07	Sobic.003G312550	Downstream			
3	66 638 290	NEF 2012	2	UJIMiscanthus118372	1.31E-31	Sobic.003G345100	5' UTR	AT3G11180	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase	
3	67 125 342	HU 2012	2	UJIMiscanthus017647	7.85E-16	Sobic.003G352200	CDS	AT5G02320	myb domain protein 3r-5	ATMYB3R5, MYB3R-5
3	72 078 520	KNU 2012	2	UJIMiscanthus019389	5.05E-45	Sobic.003G413600	Upstream	AT4G13420	High affinity K <sup>+</sup> transporter 5	ATHAK5, HAK5
4	446 677	UI 2013	1	UJIMiscanthus096439	2.23E-13	Sobic.004G005200	CDS	AT1G33230	TMPIT-like protein	

(Continues)

TABLE 4 (Continued)

Chr. <sup>a</sup>	Position <sup>b</sup>	Trial <sup>c</sup>	Winter <sup>d</sup>	Miscanthus SNP <sup>e</sup>	P value <sup>f</sup>	Nearest <i>S. bicolor</i> gene	Position relative to the gene <sup>g</sup>	Arabidopsis gene	Gene ontology	Gene symbol
4	3 252 258	CSU 2013	1	UJIMiscanthus118373	5.22E-50	Sobic.004G039500	Upstream	AT4G01070	UDP-Glycosyltransferase	<i>GT72B1</i> , <i>UGT72B1</i>
4	3 460 776	UI 2012	1	UJIMiscanthus091689	1.45E-07	Sobic.004G042100	Downstream	AT5G24870	RING/U-box superfamily protein	
4	5 948 390	CSU 2013	1	UJIMiscanthus100262	1.29E-20	Sobic.004G073500	CDS	AT3G28345	ABC transporter family protein	
4	7 436 056	CSU 2013	1	UJIMiscanthus118374	8.83E-30	Sobic.004G087800	CDS	AT1G15520	Pleiotropic drug resistance 12	<i>ABCG40</i> , <i>ATABCG40</i> , <i>ATPDR12</i> , <i>PDR12</i>
4	59 577 953	CSU 2012	2	UJIMiscanthus005328	8.42E-06	Sobic.004G248700	3' UTR	AT5G16600	myb domain protein 43	<i>AMYB43</i> , <i>MYB43</i>
4	62 207 466	NEF 2012	1	UJIMiscanthus022071	9.25E-07	Sobic.004G279300	CDS	AT5G37260	Homeodomain-like protein	<i>CIR1</i> , <i>RVE2</i>
4	62 602 070	KNU 2012	2	UJIMiscanthus021451	1.15E-49	Sobic.004G283800	CDS	AT4G00750	S-adenosyl-L-methionine-dependent methyltransferases	
4	62 928 690	KNU 2012	2	UJIMiscanthus016552	1.13E-95	Sobic.004G286600	0.02 Mb downstream	AT1G20440	Cold-regulated 47, dehydrin	<i>COR47</i>
4	65 644 782	KNU 2012	1	UJIMiscanthus107624	1.74E-06	Sobic.004G321400	Upstream	AT3G18730	Tetrapeptide repeat (TPR)-containing protein	<i>BRU1</i> , <i>MGO3</i> , <i>TSK</i>
6	3 007 286	HU 2012	2	UJIMiscanthus107102	8.16E-17	Sobic.006G018100	In the gene, near 5' UTR	AT3G05950	RmlC-like cupins superfamily protein	
6	46 253 326	NEF 2012	1	UJIMiscanthus014960	1.37E-09	Sobic.006G092300	CDS	AT5G10970	C2H2 and C2HC zinc fingers	
6	49 249 029	HU 2012	2	UJIMiscanthus118375	5.04E-09	Sobic.006G127500	Upstream	AT3G22320	Eukaryotic rp55 RNA polymerase subunit	<i>ATRPABC24.3</i> , <i>NRPB5</i> , <i>NRPD5</i> , <i>RPB5A</i>
6	58 727 756	HU 2012	2	UJIMiscanthus116154	8.25E-14	Sobic.006G247700	Downstream	AT1G02065	Squamosa promoter binding protein-like 8	<i>SPL8</i>
6	59 583 985	CSU 2012	1	UJIMiscanthus018713	8.99E-08	Sobic.006G260800	CDS	AT5G04490	vitamin E pathway gene 5	<i>VTE5</i>
7	1 678 335	HU 2012	2	UJIMiscanthus110730	1.87E-08	Sobic.007G018100	Downstream	AT2G46770	NAC (No Apical Meristem) domain transcriptional regulator	<i>AMAC043</i> , <i>EMB2301</i> , <i>NST1</i>

(Continues)

TABLE 4 (Continued)

Chr. <sup>a</sup>	Position <sup>b</sup>	Trial <sup>c</sup>	Winter <sup>d</sup>	<i>Miscanthus</i> SNP <sup>e</sup>	<i>P</i> value <sup>f</sup>	Nearest <i>S. bicolor</i> gene	Position relative to the gene <sup>g</sup>	Arabidopsis gene	Gene ontology	Gene symbol
7	61 432 860	KNU 2012	2	UJMiscanthus118376	2.08E-77	Sobic.007G181100	Upstream	AT5G19700	MATE efflux family protein	
8	255 984	UI 2013	1	UJMiscanthus010426	4.79E-06	Sobic.008G002950	CDS			
8	2 549 936	CSU 2012	2	UJMiscanthus097991	6.98E-12	Sobic.008G028600	Downstream	AT2G40750	WRKY DNA-binding protein 54	ATWRKY54, WRKY54
9	44 647 057	NEF 2012	2	UJMiscanthus118377	2.42E-37	Sobic.009G111000	5' UTR	AT4G02330	Plant invertase/pectin methyltransferase inhibitor	ATPMEPCRB
9	57 313 789	CSU 2013	1	UJMiscanthus012595	4.99E-09	Sobic.009G223450	0.07 Mb upstream	AT1G29390	Cold-regulated 314 thylakoid membrane 2	COR314
10	1 280 469	CSU 2012	2	UJMiscanthus012279	1.32E-24	Sobic.010G016100	CDS	AT2G21480	Malectin/receptor-like protein kinase	
10	6 194 258	KNU 2013	1	UJMiscanthus118378	8.99E-09	Sobic.010G075500	Upstream			
10	44 512 735	CSU 2013	1	UJMiscanthus118379	5.96E-72	Sobic.010G152900	CDS	AT5G64570	Beta-D-xylosidase 4	ATBXLA, XYL4
10	49 849 126	CSU 2012	1	UJMiscanthus118380	2.17E-06	Sobic.010G168900	Downstream	AT2G19600	K+ efflux antiporter 4	ATKEA4, KEA4
U		NEF 2012	2	UJMiscanthus007416	1.88E-17					
U		NEF 2012	2	UJMiscanthus025782	6.40E-30					
U		CSU 2013	1	UJMiscanthus005853	5.91E-52					
U		HU 2012	2	UJMiscanthus009034	3.90E-08					
U		KNU 2013	1	UJMiscanthus025992	4.18E-12					
U		KNU2012	1	UJMiscanthus001453	1.27E-10					
U		NEF 2012	1	UJMiscanthus012849	8.17E-09					
U		KNU 2012	2	UJMiscanthus016292	6.16E-127					
U		UI 2013	1	UJMiscanthus020970	4.30E-14					
U		UI 2013	1	UJMiscanthus101153	2.89E-08					
U		KNU 2013	1	UJMiscanthus101527	6.02E-12					
U		KNU 2012	2	UJMiscanthus102803	2.25E-09					
U		HU 2012	2	UJMiscanthus111354	3.57E-08					
U		CSU2013	1	UJMiscanthus111993	2.89E-17					
U		CSU2012	2	UJMiscanthus116202	7.80E-08					
U		CSU 2013	1	UJMiscanthus118381	1.97E-10					
U		KNU 2013	2	UJMiscanthus118382	8.34E-08					
U		KNU 2012	2	UJMiscanthus118383	1.36E-99					

(Continues)

TABLE 4 (Continued)

Chr. <sup>a</sup>	Position <sup>b</sup>	Trial <sup>c</sup>	Winter <sup>d</sup>	<i>Miscanthus</i> SNP <sup>e</sup>	<i>P</i> value <sup>f</sup>	Nearest <i>S. bicolor</i> gene	Position relative to the gene <sup>g</sup>	Arabidopsis gene	Gene ontology	Gene symbol
U		NEF 2012	1	UJMiscanthus118384	1.48E-08					
U		NEF 2012	1	UJMiscanthus118385	9.21E-08					
U		NEF 2012	1	UJMiscanthus118386	9.72E-20					
U		KNU 2012	1	UJMiscanthus118387	6.30E-07					
U		KNU 2013	1	UJMiscanthus118387	3.37E-09					
U		KNU 2013	2	UJMiscanthus118388	1.01E-08					
U		NEF 2012	1	UJMiscanthus118389	2.09E-11					
U		NEF 2012	1	UJMiscanthus118390	8.59E-12					
U		KNU 2013	2	UJMiscanthus118391	4.69E-13					
U		KNU2013	1	UJMiscanthus118392	7.47E-09					
U		KNU2013	2	UJMiscanthus118393	7.22E-15					
U		CSU2013	1	UJMiscanthus118394	1.60E-07					
U		HU2012.	1	UJMiscanthus118395	7.67E-14					
U		CSU2012	2	UJMiscanthus118396	1.95E-18					

<sup>a</sup>*Sorghum bicolor* 3.0 chromosome to which the marker was aligned. <sup>b</sup>U indicates marker was not aligned against *S. bicolor* genome. <sup>c</sup>S, *bicolor* 3.0 chromosome position to which the marker was aligned. <sup>d</sup>Trial, field trial location and associated establishment year. <sup>e</sup>Winter, first- and second-winter post field trial establishment. <sup>f</sup>*Miscanthus* SNP, significant marker-trait associations detected for overwintering ability. <sup>g</sup>*P* value, significance of marker-trait associations detected using multilocus mixed model (MLMM). <sup>h</sup>CDS: coding domain sequence; UTR: untranslated region.

Prediction methods	N	Prediction ability <sup>a</sup>			Proportion. of variance explained by genotype <sup>c</sup>
		Mean	SE	Range	
Across DAPC groups					
Method 1	564	0.81	0.00	0.80–0.82	0.16
Method 2	564	0.73	0.00	0.72–0.75	
Method 3	564	0.71	0.00	0.70–0.73	0.08
Within DAPC group					
Yangtze-Qinling	73	0.04	0.01	–0.08 to 0.28	0.04
SE China plus tropical	87	0.34	0.01	0.10–0.44	0.06
Southern Japan	142	0.48	0.00	0.41–0.57	0.05
Northern Japan <sup>d</sup>	83	NA	NA	NA	0.00
Korea/N China <sup>d</sup>	154	NA	NA	NA	0.00

<sup>a</sup>Method 1 was based on Equation 3, which had no control for population structure; Method 2 was based on Equation 4, which fitted DAPC group as a fixed effect and the derived residuals were used in genomic prediction; Method 3 was based on Equation 5, which controlled population structure by nesting genotype within DAPC group. For within-DAPC-group analysis, the Southern Japan group ( $N = 142$ ) included genotypes directly collected in southern Japan ( $N = 28$ ), US ornamentals ( $N = 76$ ), and US naturalized populations ( $N = 38$ ), given that the latter two were derived from southern Japan (Clark et al., 2014,2015) and larger sample sizes are advantageous for genomic prediction. <sup>b</sup>Genomic prediction statistics were calculated based on 50 fivefold cross-validations for each analysis. <sup>c</sup>Detailed ANOVA results provided in Data S3. <sup>d</sup>Genomic prediction within Northern Japan and Korea/N China groups were not estimable because nearly all accessions survived (Figure 4a, Table 3), which resulted in zero-value estimates for the percentage of variance explained by genotype for overwintering ability.

panel ( $N = 564$ ) that represented most of the species' natural geographic range, and evaluated the results in the context of previously ascertained population structure (Clark et al., 2014,2015). Consistent with prior studies of overwintering ability in *Miscanthus* (Clifton-Brown et al., 2001), we found that *M. sinensis* plants were typically more at risk of dying during the first winter after planting than during the second winter. However, the one exception we observed to this progression was associated with high snowfall early season at HU in year 1 that resulted in nearly all plants surviving the first winter (0.99; Table 3), but a lower overwintering rate was observed in year 2 (0.80), suggesting that avoidance of killing temperatures (in this case by insulating snow cover) can be an important contributor to overwintering in *Miscanthus*. First-year plants of *Miscanthus*, which typically produce much less aboveground biomass than second-year plants (Dong, Green, et al., 2019; Gifford, Chae, Swaminathan, Moose, & Juvik, 2015), likely also produce less belowground biomass and more shallow rhizomes than more mature plants, and thus may have few or no buds sufficiently deep in the soil to avoid killing temperatures. Especially as *Miscanthus* plants mature, those genotypes that are able to produce deep rhizomes will be able to avoid temperatures sufficiently cold to damage or kill dormant buds.

Great variation among and within *M. sinensis* genetic groups was observed for overwintering ability at five field trial

**TABLE 5** Genomic prediction ability of first-winter overwintering ability in an *M. sinensis* diversity panel, evaluated at three locations in North America (UI = University of Illinois at Urbana-Champaign, Urbana, IL, USA; CSU = Colorado State University, Fort Collins, CO, USA; NEF = New Energy Farms, Leamington, ON, Canada) and two locations in Asia (HU = Hokkaido University, Sapporo, Japan; KNU = Kangwon National University, Chuncheon, South Korea). Field trials were established at each location in 2012, and three additional field trials were established at UI, CSU, and KNU in 2013

locations, with 73 significantly associated SNPs identified via GWA analyses. Genetic groups and genotypes that originated from temperate environments at relatively high latitudes typically had greater overwintering ability than those from subtropical and tropical environments, as expected (Table 3, Figures 2, 3, 4a). We observed positive association between latitude of origin and overwintering ability ( $R^2 = 0.49$ , Figure 4a). Similarly, Yan et al. (2012) observed that first-year overwintering ability of 31 *M. sinensis* accessions collected in China was positively associated with latitude of origin when grown at two northern field trial sites, and the correlation was greatest at the most northern site with the coldest winter (Xilinhot). In addition, BioClim analysis revealed that BIO1 (annual mean temperature), BIO6 (minimum temperature of coldest month), BIO10 (mean temperature of warmest quarter), and BIO11 (mean temperature of coldest quarter) had significant effects on first-winter overwintering ability (Table S2), indicating that adaptation to low temperatures in winters as well as to a certain amount of heat units in the growing season (as determined by location of origin) both influenced actual overwintering at the trial sites. For example, if a plant is adapted to a long growing season but is then grown in a place with a short growing season, we would expect it to be less likely to survive the winter because it will not be at the right physiological state (e.g., insufficient storage of carbohydrates

underground) for winter dormancy. During the especially cold winter of 2013–2014 at UI, first-year plants of many ornamental *M. sinensis* cultivars that typically overwinter well at Urbana, IL, in spite of being derived from the Southern Japan group, either died or were severely damaged, yet genotypes from northern China such as ‘PMS-436’, and rare *M. sinensis* ornamental cultivars such as ‘Haiku’ survived and grew vigorously in spring 2014 (Figure 2a). This event highlighted the importance of selecting for the coldest winter expected in a target environment, especially given that *Miscanthus* is a long-lived perennial crop and establishment of new commercial plantings is expensive.

Thus, one strategy to breed *M. sinensis* for temperate environments is to select concurrently for overwintering ability and high biomass yield within the genetic groups that naturally have the greatest adaptation to cold winters (e.g., Korea/N China, Northern Japan group, Southern Japan group, and the Yangtze-Qinling group). Additionally, the rare relatively hardy individuals identified within the subtropical- and tropical-adapted Sichuan Basin group (e.g., PMS-005; Data S1) and Southeastern China plus tropical group (e.g., PMS-008; Data S1) represent a potentially valuable breeding opportunity because *M. sinensis* accessions from lower latitudes typically have higher biomass yield potential than those from higher latitudes (Clark et al., 2019). Moreover, the presence of advantageous overwintering alleles even within the subtropical- and tropical-adapted *M. sinensis* genetic groups, albeit at lower frequencies than in the temperate-adapted groups (Figure 4b, Data S2), should facilitate rapid genetic gains via marker-assisted selection. The opportunity to conduct marker-assisted selection for advantageous overwintering alleles at a large number of loci within any of the *M. sinensis* genetic groups is expected to greatly improve breeding efficiency, especially if the breeder must otherwise rely on unusually cold winters to screen populations.

#### 4.2 | Introgression of genes for overwintering ability from northern-adapted *M. sacchariflorus* into *M. sinensis* is another potentially useful breeding strategy

Introgression of winter hardiness alleles from cold temperate-adapted *M. sacchariflorus* may be another viable strategy for improving overwintering ability in *M. sinensis*. We have shown previously that about half of the ornamental cultivars sold as *M. sinensis* in the US and Europe are in fact hybrids between *M. sinensis* and diploid *M. sacchariflorus* that have been backcrossed to *M. sinensis* one or more times. Notably, the *M. sacchariflorus* ancestry in these hybrids is from northern China or eastern Russia (Clark et al., 2014, 2018), which are cold temperate environments (hardiness zones 3–5). In this study, we included 76 accessions of predominantly *M. sinensis* ancestry from

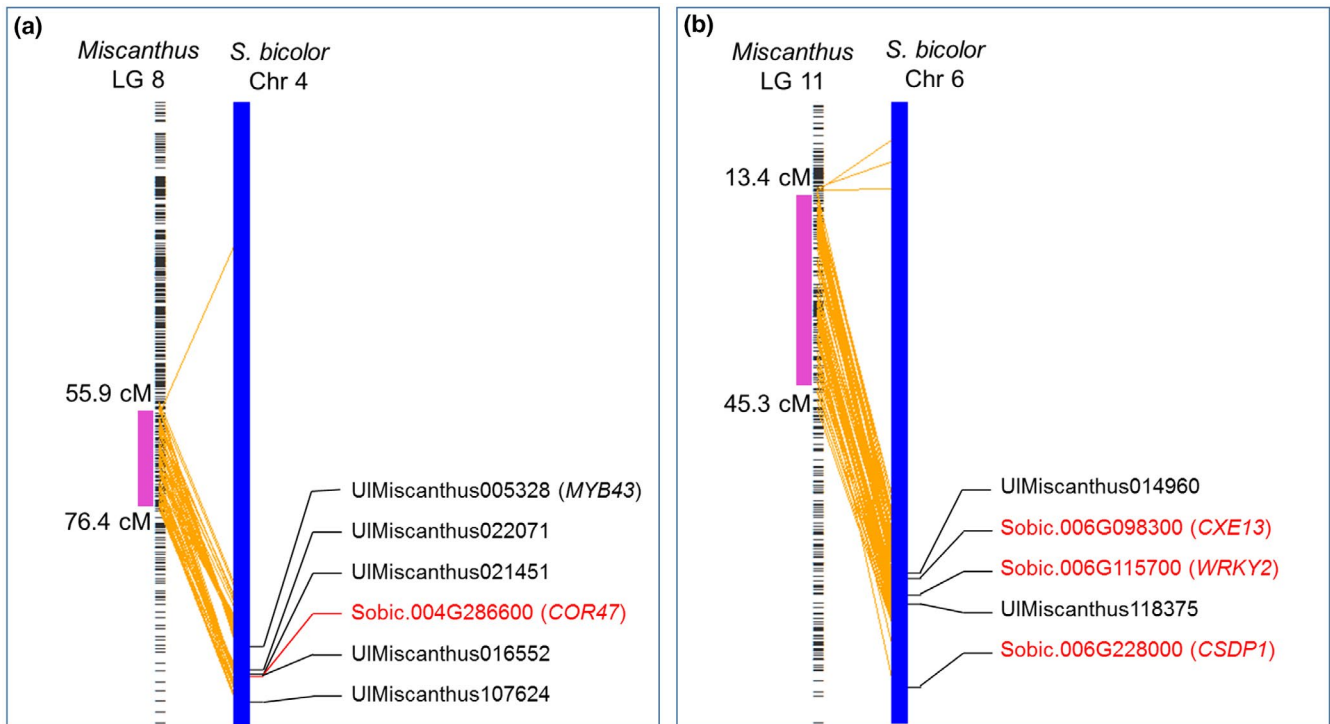
the US ornamental cultivars genetic group, and *M. sacchariflorus* ancestry in these cultivars ranged from 0.00 to 0.36 (Data S1). A significant association was observed between first-winter overwintering ability of the ornamental cultivars and their proportion *M. sacchariflorus* ancestry ( $R^2 = 0.17$ ,  $p = 0.0003$ , Figure S4). Yan et al. (2012) observed that among 48 accessions of *M. sacchariflorus* and 31 of *M. sinensis* from China grown at two northern field trial sites in China, the former was more winter-hardy than the latter. Similarly, in eastern Russia, Clark et al. (2016) did not find wild populations of *M. sinensis* in areas colder than hardiness zone 5, but did find abundant populations of *M. sacchariflorus* through hardiness zone 3. Thus, there appears to be a good case for using cold tolerance genes from *M. sacchariflorus* to improve *M. sinensis*.

However, we should not assume that all *M. sacchariflorus* genotypes are equally good sources of winter hardiness. Previously, we have identified three tetraploid and three diploid *M. sacchariflorus* genetic groups, each associated with different geographical regions in eastern Asia, ranging from the mild Yangtze River region to the cold winters in eastern Russia along the Amur River (Clark et al., 2018). For example, Clifton-Brown et al. (2001) observed that first-winter (1997–1998) survival for a tetraploid *M. sacchariflorus* (EMI no. 5, collected along the Nagara River in Gifu Prefecture, Japan) was only 50% in Sweden and 33% in Denmark, whereas an *M. sinensis* from Hokkaido (Northern Japan group) had 95%–99% survival at the same locations. Similarly, Clifton-Brown et al. (2001) found that nearly all of the triploid M×g tested in Sweden and Denmark died after the first winter, but survival of four *M. sinensis* from Honshu Japan had survival rates of 84%–99%.

If the ultimate goal of breeding *M. sinensis* is to use it as a parent for making improved triploid M×g by crossing it with tetraploid *M. sacchariflorus*, then it would also be desirable to have an understanding of potential heterotic groups within *M. sacchariflorus*, and how introgressed genes from a diploid *M. sacchariflorus* accession into *M. sinensis* might interact with a tetraploid *M. sacchariflorus* from a different genetic group. However, there is currently little information on heterotic groups and gene interactions in *Miscanthus* to guide such breeding strategies.

#### 4.3 | Trait association comparisons between this GWA and a parallel study of three interconnected biparental populations

In our parallel study of three interconnected diploid *Miscanthus* populations (Dong, Liu, et al., 2019), we identified nine QTL for overwintering ability via joint population analysis, whereas in the current GWA analyses, we identified 73 significant SNPs for overwintering ability (Figure S1, Table 4), an eightfold difference. Greater sampling of genetic



**FIGURE 5** Correspondence between QTLs for overwintering ability detected across three interconnected diploid  $F_1$  *Miscanthus* populations (Dong, Liu, et al., 2019) and genome-wide associations (GWA) for overwintering ability in an *M. sinensis* diversity panel. Black dashed bars represent *Miscanthus* linkage groups (LGs), and blue bars represent *S. bicolor* (version 3.0) LGs. Orange lines represent corresponding genomic regions between *Miscanthus* and *Sorghum bicolor*. (a) An overwintering ability QTL on *Miscanthus* LG 8 identified via genetic mapping in three interconnected biparental populations (purple bar) encompassed five marker–trait associations from the GWA analysis (black text). Text inside the parenthesis represents gene symbols. The cold tolerance gene *COR47* (red text; previously identified in *Arabidopsis* by Puhakainen et al. (2004) and Bozovic et al. (2013)) was located within this QTL and was 0.02 Mb away from one GWA hit (UIMiscanthus016552). (b) A QTL for overwintering ability on *Miscanthus* LG 11 identified via genetic mapping in three diploid interconnected biparental populations (purple bar) corresponded to two marker–trait associations from the GWA analysis (black text). Three additional candidate genes (red text) including carboxylesterase 13 (*CEX13*), *WRKY* transcription factor (*WRKY2*), and cold shock domain protein 1 (*CSDP1*) were also located inside this QTL

diversity in the *M. sinensis* germplasm panel (564 genotypes from throughout the species' geographic range) in comparison to the four parents of the interconnected populations, likely contributed to the greater number of significant associations identified. Additionally, the five field trial locations in years 1 and 2 for phenotyping in the GWA study was likely advantageous over the one location in year 3 for phenotyping in the interconnected populations study. Moreover, in linkage analysis of biparental populations, there are only a few opportunities for recombination to occur within each population, and QTLs are only detected based on recombination events that occurred during population development, resulting in relatively low mapping resolution (typically 10–20 cM, Doerge, 2002; Holland, 2007). In contrast, historical recombination and greater genetic diversity can be exploited for high resolution mapping in association analysis (Lipka et al., 2015; Yu & Buckler, 2006), and this appears to have been the case for our studies of overwintering ability in *Miscanthus*.

Two QTLs identified in the interconnected biparental population study included a total of seven SNP–trait associations from the GWA study (Figure 5). One QTL on

*Miscanthus* LG 8 encompassed five marker–trait associations from GWA that aligned to *S. bicolor* chromosome 4 (Figure 5a), and the other QTL on *Miscanthus* LG 11 encompassed two marker–trait associations from GWA that aligned to *S. bicolor* chromosome 6 (Figure 5b). The identification of significant SNPs via GWA within independently identified QTL from three interconnected biparental populations lends strong support to the conclusion that these regions of the genome are important for overwintering ability in *Miscanthus*.

#### 4.4 | Marker-assisted selection with 73 significant markers from GWA and genomic prediction for overwintering ability

Using the 73 markers detected in GWA analyses, both the estimated genotypic values (considering both additive and dominance effects) and the estimated breeding values (considering only additive effect) explained 55%–56% of the variation for observed first-winter survival in this *M. sinensis* germplasm panel (Figure 4c, Figure S2). However, the predictions for overwintering ability based on these 73 markers

could be overestimated because the genotypic and breeding values were calculated from the same dataset that was used to identify the significant SNPs. This phenomenon of inflated prediction is known in as ‘inside trading’ in genomic selection (Arruda et al., 2016). Moreover, given that the dominance effects across these 73 markers ranged from  $-0.59$  to  $0.44$  (Data S2), the similar prediction accuracy between estimated genotypic values and estimated breeding values suggested that the dominance effects of alleles at multiple loci might have cancelled out each other.

The high genomic predictive ability for overwintering ability obtained after accounting for population structure (0.71–0.73; Table 5) indicates that genomic prediction should be highly beneficial for breeding new cultivars of *M. sinensis* with improved winter hardiness. The potential use of genomic prediction in breeding programs has been successfully demonstrated in livestock (Schaeffer, 2006), annual crops (Heffner, Lorenz, Jannink, & Sorrells, 2010; Heffner, Sorrells, & Jannink, 2009), and perennial grass such as switchgrass (*Panicum virgatum* L.) (Lipka et al., 2014) and *M. sinensis* (Slavov et al., 2014, 2018). In switchgrass, high predictive abilities were obtained for most morphological traits such as anthesis date (0.44, Lipka et al., 2014), standability (0.52), and leaf length (0.40). In *M. sinensis*, moderate to high predictive abilities were reported for most traits related to phenology (0.64–0.76), and cell wall composition (0.25–0.62), but were low for dry biomass yield (0.06; Slavov et al., 2014).

Genomic prediction could be more efficient than marker-assisted selection in breeding cold-hardy *Miscanthus*. GWA analyses cannot detect with statistical significance all causative loci underlying a trait of interest. Moreover, in both linkage mapping and GWA analyses, the separation of marker–trait association detection from marker effect estimation results in biased effect estimation (Beavis, 1994; Jannink, Lorenz, & Iwata, 2010; Lande & Thompson, 1990). In contrast, genomic prediction uses all markers for modeling the performance of individuals. Nevertheless, GWA enables identification of candidate genes that can help elucidate gene networks. Thus, GWA and genomic prediction are complementary strategies. Even though the equal variance assumption (i.e., all markers contribute equally to the genetic variances of trait) in ridge regression is unrealistic in a breeding context, superior predictive ability of genomic selection relative to marker-assisted selection has been demonstrated (Arruda et al., 2016; Heffner, Jannink, & Sorrells, 2011; Jannink et al., 2010; Meuwissen, Hayes, & Goddard, 2001).

Population structure affected genomic prediction of overwintering ability for *M. sinensis*. For whole panel analysis, Method 1, which did not account for population structure, resulted in prediction accuracies that were biased upward relative to Methods 2 (by 8 points) and 3 (by 10 points) which did account for population structure; the observed bias was consistent with prior studies on other crops (Fiedler et al., 2018; Riedelsheimer et al.,

2012; Spindel et al., 2015). To effectively differentiate the overwintering potential of individuals within the Northern Japan and Korea/N China groups, from which nearly all genotypes survived the first winter, higher resolution phenotyping schemes such as a 1–10 ordinal system to capture relative vigor, rather than the current binary system that captured only survival data could be helpful; additionally, challenging these genotypes in colder winter environments would be another promising strategy.

## 4.5 | Candidate genes

GWA analyses provided information for dissecting the genetic mechanism of overwintering ability in *Miscanthus*. Of the 42 marker–trait associations that aligned to the *S. bicolor* genome, many were located near known cold-responsive genes (Table 4). Moreover, some of the candidate genes associated with significant SNPs were found within QTL identified in our parallel study of three interconnected biparental populations (Figure 5). For example, within the overwintering ability QTL on LG 8, one of the significant SNPs from GWA (UIMiscanthus005328, aligned to *S. bicolor* chromosome 4) was within a *MYB* gene (*MYB43*, Table 4, Figure 5a), which is a transcription factor that has been shown to regulate stress response in *Arabidopsis* (Barah et al., 2016). The known cold tolerance gene, *COR47*, was also within the QTL on LG 8 and 0.02 Mb downstream from one GWA hit (UIMiscanthus016552; Figure 5a, Table 4). *COR47* protects thylakoid membranes during freezing by encoding dehydrin proteins, thereby enhancing cold tolerance in *Arabidopsis* (Bozovic, Svensson, Schmitt, & Kohn, 2013; Puhakainen et al., 2004). Upregulation of *COR* genes has also been reported in perennial ryegrass (*Lolium perenne* L. cv. ‘Caddyshack’) during cold acclimation (Zhang, Fei, Warnke, Li, & Hannapel, 2009). Within the overwintering ability QTL on LG 11, three candidate cold-tolerance genes were identified, including carboxylesterase 13 (*CEX13*), WRKY transcription factor (*WRKY2*), and cold shock domain protein 1 (*CSDP1*). Two of these candidate cold-tolerance genes (*CEX13* and *WRKY2*) were closely linked to the significant SNPs from GWA (*CEX13* was 0.53 Mb from UIMiscanthus014960; *WRKY2* was 0.88 Mb from UIMiscanthus118375; Figure 5b, Table 4). Carboxylesterase is an enzyme that regulates biological activities through hydrolysis (Gershter & Edwards, 2007). Although the functional details of carboxylesterase are still unknown, several members of this gene family have been detected following cold treatment in *Arabidopsis* (Wagstaff et al., 2010) and in grape (*Vitis vinifera*) (Xin et al., 2013). WRKY is a large transcription factor family that plays a broad spectrum regulatory role in plant defense regulation, response to abiotic stresses, growth, and development (Agarwal, Reddy, & Chikara, 2011; Seki et al., 2002). *WRKY2* has been shown to have elevated expression in response to osmotic stress in *Arabidopsis* (Jiang & Yu, 2009). In plants, cold shock



domain proteins play essential roles in acquiring freezing tolerance (Sasaki & Imai, 2012); they are among the most evolutionarily conserved nucleic acid-binding domains, predating the divergence of prokaryotes and eukaryotes (Graumann & Marahiel, 1998; Karlson & Imai, 2003).

One theme to emerge from this analysis of candidate genes was the identification of multiple members of gene families known to be involved with plant response to cold stress in or near newly identified SNPs throughout the *Miscanthus* genome that were associated with overwintering ability, which is consistent with the importance of gene duplication in plant evolution (Hu et al., 2003; Ming et al., 2002; Paterson, Lin, et al., 1995; Paterson, Schertz, Lin, Liu, & Chang, 1995). For example, in addition to the cold shock domain locus on *S. bicolor* chromosome 6 that was within the QTL on *Miscanthus* LG 11 and near three significant SNPs for overwintering ability (Figure 5b), another cold shock domain locus on a *S. bicolor* chromosome 3 was 0.01 Mb from marker–trait association UIMiscanthus118370 (Table 4). Similarly, in addition to the *WRKY2* transcription factor on *Miscanthus* LG 11/ *S. bicolor* chromosome 6, another *WRKY* locus (*WRKY54*) was located downstream of marker–trait association UIMiscanthus097991, which aligned to *S. bicolor* chromosome 8 (Table 4). Similar to *COR47* within the QTL on *Miscanthus* LG 8/*S. bicolor* chromosome 6, another locus associated with protecting thylakoid membranes was identified, *COR314*, which was located 0.07 Mb from UIMiscanthus012595 and aligned to *S. bicolor* chromosome 9 (Grundy, Stoker, & Carré, 2015; Li et al., 2018). Additionally, three marker–trait associations (UIMiscanthus118368 aligned to *S. bicolor* Chromosome 2, UIMiscanthus017647 aligned to *S. bicolor* Chromosome 3, and UIMiscanthus005328 aligned to *S. bicolor* Chromosome 4) were all located within or near *MYB* transcription factor genes (Table 4). However, some likely candidate genes were only observed once. For example, the marker–trait association, UIMiscanthus118366 aligned to *S. bicolor* chromosome 2, was only 177 bp downstream from the *JAZ8* gene (jasmonate-zim-domain protein 8), which has increased expression under cold stress in wheat (*Triticum aestivum* L.); moreover, this gene is part of a gene family that regulates responses to biotic and abiotic stresses typically by interacting with *MYB* transcription factors (Wang et al., 2017), which we have also found to play a role in the overwintering ability of *Miscanthus*. Thus, the candidate genes identified here represent testable hypotheses about the genes underlying overwintering ability of *Miscanthus*. Moreover, the functions of the newly identified candidate genes suggest that tolerance to freeze stress is an important component overwintering ability in *Miscanthus*, in addition to likely avoidance strategies.

In summary, screening a large panel of *M. sinensis* germplasm enabled us to determine that the Korea/N China genetic group would likely be a valuable gene pool for cold tolerance, and cold-tolerant genotypes are also frequent in the

Yangtze-Qinling, Southern Japan, and Northern Japan groups. We detected 73 marker–trait associations for overwintering ability using the large *M. sinensis* germplasm panel and observed that *Miscanthus* accessions collected from high latitude locations with cold winters had higher rates of overwintering, and more alleles for overwintering, than accessions collected from low latitude locations with mild winters. Consistency between QTLs and GWA hits suggested that these genomic regions are important for response to cold stress. Many of the candidate genes underlying peak marker–trait associations provided interesting hypotheses for further testing. Similarly, genomic prediction abilities for overwintering ability were high. Given these results, we expect both GWA and genomic prediction to substantially improve breeding efficiency for winter hardiness in *M. sinensis*. Thus, this study represents a significant step toward the development of new *Miscanthus* cultivars that are optimally adapted to temperate regions with cold winters.

## ACKNOWLEDGEMENTS

This research was supported by the Energy Biosciences Institute, USDA HATCH project ILLU-802-311, the National Institute of Food and Agriculture, U.S. Department of Agriculture, under award number 2018-68005-27937, and the DOE Office of Science, Office of Biological and Environmental Research (BER), grant nos. DE-SC0006634 and DE-SC0018420 (Center for Advanced Bioenergy and Bioproducts Innovation). We thank Pat Brown for providing *PstI* adaptors in RAD-seq library preparation. We thank Megan Swanson, Kasia Dubiel, and Melina Salgado for assistance with DNA extraction. Christine Fleener and Lauren Quinn collected naturalized *M. sinensis* in the USA.

## ORCID

Hongxu Dong  <https://orcid.org/0000-0003-4028-3597>

Lindsay V. Clark  <https://orcid.org/0000-0002-3881-9252>

Toshihiko Yamada  <https://orcid.org/0000-0002-7845-6556>

Stephen P. Long  <https://orcid.org/0000-0002-8501-7164>

## REFERENCES

- Agarwal, P., Reddy, M. P., & Chikara, J. (2011). WRKY: Its structure, evolutionary relationship, DNA-binding selectivity, role in stress tolerance and development of plants. *Molecular Biology Reports*, 38(6), 3883–3896. <https://doi.org/10.1007/s11033-010-0504-5>
- Arruda, M. P., Lipka, A. E., Brown, P. J., Krill, A. M., Thurber, C., Brown-Guedira, G., ... Kolb, F. I. (2016). Comparing genomic selection and marker-assisted selection for Fusarium head blight resistance in wheat (*Triticum aestivum* L.). *Molecular Breeding*, 36(84), 1–11. <https://doi.org/10.1007/s11032-016-0508-5>
- Barah, P., B N, M. N., Jayavelu, N. D., Sowdhamini, R., Shameer, K., & Bones, A. M. (2016). Transcriptional regulatory networks in *Arabidopsis thaliana* during single and combined stresses. *Nucleic Acids Research*, 44(7), 3147–3164.

- Beavis, W. D. (1994). The power and deceit of QTL experiments: lessons from comparative QTL studies. In D. B. Wilkinson (Ed.), *Proceedings of the 49th Annual Corn and Sorghum Research Conference* (pp. 250–265). Washington DC: American Seed Trade Association.
- Boersma, N. N., Dohleman, F. G., Miguez, F. E., & Heaton, E. A. (2015). Autumnal leaf senescence in *Miscanthus* × *giganteus* and leaf [N] differ by stand age. *Journal of Experimental Botany*, *66*(14), 4395–4401.
- Bozovic, V., Svensson, J., Schmitt, J., & Kohn, C. (2013). Dehydrins (LTI29, LTI30, and COR47) from *Arabidopsis thaliana* expressed in *Escherichia coli* protect thylakoid membranes during freezing. *Journal of the Serbian Chemical Society*, *78*(8), 1149–1160.
- Burner, D. M., Tew, T. L., Harvey, J. J., & Belesky, D. P. (2009). Dry matter partitioning and quality of *Miscanthus*, *Panicum*, and *Saccharum* genotypes in Arkansas, USA. *Biomass and Bioenergy*, *33*(4), 610–619. <https://doi.org/10.1016/j.biombioe.2008.10.002>
- Chae, W. B., Hong, S. J., Gifford, J. M., Rayburn, A. L., Sacks, E. J., & Juvik, J. A. (2014). Plant morphology, genome size, and SSR markers differentiate five distinct taxonomic groups among accessions in the genus *Miscanthus*. *GCB Bioenergy*, *6*(6), 646–660.
- Christian, D. G., & Haase, E. (2001). Agronomy of *Miscanthus*. In M. B. Jones, & M. Walsh (Eds.), *Miscanthus for energy and fibre* (pp. 21–45). London, UK: James & James.
- Clark, L. V., Brummer, J. E., Głowacka, K., Hall, M. C., Heo, K., Peng, J., ... Sacks, E. J. (2014). A footprint of past climate change on the diversity and population structure of *Miscanthus sinensis*. *Annals of Botany*, *114*(1), 97–107. <https://doi.org/10.1093/aob/mcu084>
- Clark, L. V., Dwiyantri, M. S., Anzoua, K. G., Brummer, J. E., Ghimire, B. K., Głowacka, K., ... Sacks, E. J. (2019). Biomass yield in a genetically diverse *Miscanthus sinensis* germplasm panel evaluated at five locations revealed individuals with exceptional potential. *GCB Bioenergy*, in press. <https://doi.org/10.1111/gcbb.12606>
- Clark, L. V., Dzyubenko, E., Dzyubenko, N., Bagmet, L., Sabitov, A., Chebukin, P., ... Sacks, E. J. (2016). Ecological characteristics and *in situ* genetic associations for yield-component traits of wild *Miscanthus* from eastern Russia. *Annals of Botany*, *118*(5), 941–955.
- Clark, L. V., Jin, X., Petersen, K. K., Anzoua, K. G., Bagmet, L., Chebukin, P., ... Sacks, E. J. (2018). Population structure of *Miscanthus sacchariflorus* reveals two major polyploidization events, tetraploid-mediated unidirectional introgression from diploid *Miscanthus sinensis*, and diversity centred around the Yellow Sea. *Annals of Botany*, in press. <https://doi.org/10.1093/aob/mcy161>
- Clark, L. V., Stewart, J. R., Nishiwaki, A., Toma, Y., Kjeldsen, J. B., Jørgensen, U., ... Sacks, E. J. (2015). Genetic structure of *Miscanthus sinensis* and *M. sacchariflorus* in Japan indicates a gradient of bidirectional but asymmetric introgression. *Journal of Experimental Botany*, *66*(14), 4213–4225.
- Clifton-Brown, J. C., Chiang, Y. C., & Hodkinson, T. R. (2008). *Miscanthus*: Genetic resources and breeding to enhance bioenergy production. In W. Vermerris (Ed.), *Genetic improvement of bioenergy crops* (pp. 273–294). New York, NY: Springer.
- Clifton-Brown, J. C., & Lewandowski, I. (2000). Overwintering problems of newly established *Miscanthus* plantations can be overcome by identifying genotypes with improved rhizome cold tolerance. *New Phytologist*, *148*(2), 287–294. <https://doi.org/10.1046/j.1469-8137.2000.00764.x>
- Clifton-Brown, J. C., Lewandowski, I., Andersson, B., Basch, G., Christian, D. G., Kjeldsen, J. B., ... Teixeira, F. (2001). Performance of 15 *Miscanthus* genotypes at five sites in Europe. *Agronomy Journal*, *93*(5), 1013–1019.
- Clifton-Brown, J. C., Stampfl, P. F., & Jones, M. B. (2004). *Miscanthus* biomass production for energy in Europe and its potential contribution to decreasing fossil fuel carbon emissions. *Global Change Biology*, *10*(4), 509–518. <https://doi.org/10.1111/j.1529-8817.2003.00749.x>
- Daetwyler, H. D., Calus, M. P., Pong-Wong, R., de Los Campos, G., & Hickey, J. M. (2013). Genomic prediction in animals and plants: Simulation of data, validation, reporting, and benchmarking. *Genetics*, *193*(2), 347–365. <https://doi.org/10.1534/genetics.112.147983>
- Doerge, R. W. (2002). Mapping and analysis of quantitative trait loci in experimental populations. *Nature Review Genetics*, *3*(1), 43–52. <https://doi.org/10.1038/nrg703>
- Dong, H., Green, S. V., Nishiwaki, A., Yamada, T., Stewart, J. R., Deuter, M., & Sacks, E. J. (2019). Winter hardiness of *Miscanthus* (I): Overwintering ability and yield of new *Miscanthus* × *giganteus* genotypes in Illinois and Arkansas. *GCB Bioenergy*, in press. <https://doi.org/10.1111/gcbb.12588>
- Dong, H., Liu, S., Clark, L. V., Sharma, S., Gifford, J. M., Juvik, J. A., ... Sacks, E. J. (2018). Genetic mapping of biomass yield in three interconnected *Miscanthus* populations. *GCB Bioenergy*, *10*(3), 165–185.
- Dong, H., Liu, S., Clark, L. V., Sharma, S., Gifford, J. M., Juvik, J. A., ... Sacks, E. J. (2019). Winter hardiness of *Miscanthus* (II): Genetic mapping for overwintering ability and adaptation traits in three interconnected *Miscanthus* populations. *GCB Bioenergy*, in press. <https://doi.org/10.1111/gcbb.12587>
- Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. *The Plant Genome*, *4*(3), 250–255. <https://doi.org/10.3835/plantgenome2011.08.0024>
- Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*, *164*(4), 1567–1587.
- Fiedler, J. D., Lanzatella, C., Edmé, S. J., Palmer, N. A., Sarath, G., Mitchell, R., & Tobias, C. M. (2018). Genomic prediction accuracy for switchgrass traits related to bioenergy within differentiated populations. *BMC Plant Biology*, *18*, 142. <https://doi.org/10.1186/s12870-018-1360-z>
- Friesen, P. C., Peixoto, M. M., Lee, D. K., & Sage, R. F. (2015). Sub-zero cold tolerance of *Spartina pectinata* (prairie cordgrass) and *Miscanthus* × *giganteus*: Candidate bioenergy crops for cool temperate climates. *Journal of Experimental Botany*, *66*(14), 4403–4413.
- Gershater, M. C., & Edwards, R. (2007). Regulating biological activity in plants with carboxylesterases. *Plant Science*, *173*(6), 579–588. <https://doi.org/10.1016/j.plantsci.2007.08.008>
- Gifford, J. M., Chae, W. B., Swaminathan, K., Moose, S. P., & Juvik, J. A. (2015). Mapping the genome of *Miscanthus sinensis* for QTL associated with biomass productivity. *GCB Bioenergy*, *7*(4), 797–810.
- Głowacka, K., Adhikari, S., Peng, J. H., Gifford, J. M., Juvik, J. A., Long, S. P., & Sacks, E. J. (2014). Variation in chilling tolerance for photosynthesis and leaf extension growth among genotypes related to the C<sub>4</sub> grass *Miscanthus* × *giganteus*. *Journal of Experimental Botany*, *65*(18), 5267–5278. <https://doi.org/10.1093/jxb/eru287>
- Głowacka, K., Clark, L. V., Adhikari, S., Peng, J., Stewart, J. R., Nishiwaki, A., ... Sacks, E. J. (2015). Genetic variation in *Miscanthus* × *giganteus* and the importance of estimating genetic distance thresholds for differentiating clones. *GCB Bioenergy*, *7*(2), 386–404.
- Głowacka, K., Kaczmarek, Z., & Jeżowski, S. (2012). Androgenesis in the bioenergy plant *Miscanthus sinensis*: From calli induction to plant regeneration. *Crop Science*, *52*(6), 2659–2673.

- Graumann, P. L., & Marahiel, M. A. (1998). A superfamily of proteins that contain the cold-shock domain. *Trends in Biochemical Sciences*, 23(8), 286–290. [https://doi.org/10.1016/S0968-0004\(98\)01255-9](https://doi.org/10.1016/S0968-0004(98)01255-9)
- Greef, J. M., & Deuter, M. (1993). Syntaxonomy of *Miscanthus* × *giganteus* Greef et Deu. *Angewandte Botanik*, 67, 87–90.
- Grundy, J., Stoker, C., & Carré, I. A. (2015). Circadian regulation of abiotic stress tolerance in plants. *Frontiers in Plant Science*, 6, 648. <https://doi.org/10.3389/fpls.2015.00648>
- Heaton, E. A., Dohleman, F. G., & Long, S. P. (2008). Meeting US biofuel goals with less land: The potential of *Miscanthus*. *Global Change Biology*, 14(9), 2000–2014. <https://doi.org/10.1111/j.1365-2486.2008.01662.x>
- Heaton, E. A., Voigt, T., & Long, S. P. (2004). A quantitative review comparing the yields of two candidate C<sub>4</sub> perennial biomass crops in relation to nitrogen temperature and water. *Biomass and Bioenergy*, 27(1), 21–30. <https://doi.org/10.1016/j.biombioe.2003.10.005>
- Heffner, E. L., Jannink, J.-L., & Sorrells, M. E. (2011). Genomic selection accuracy using multifamily prediction models in a wheat breeding program. *The Plant Genome*, 4, 65–75. <https://doi.org/10.3835/plantgenome2010.12.0029>
- Heffner, E. L., Lorenz, A. J., Jannink, J.-L., & Sorrells, M. E. (2010). Plant breeding with genomic selection: Gain per unit time and cost. *Crop Science*, 50, 1681–1690. <https://doi.org/10.2135/cropsci2009.11.0662>
- Heffner, E. L., Sorrells, M. E., & Jannink, J.-L. (2009). Genomic selection for crop improvement. *Crop Science*, 49, 1–12. <https://doi.org/10.2135/cropsci2008.08.0512>
- Hellevik, O. (2009). Linear versus logistic regression when the dependent variable is a dichotomy. *Quality & Quantity*, 43(1), 59–74. <https://doi.org/10.1007/s11135-007-9077-3>
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15), 1965–1978. <https://doi.org/10.1002/joc.1276>
- Hodkinson, T. R., & Renvoize, S. (2001). Nomenclature of *Miscanthus* × *giganteus* (Poaceae). *Kew Bulletin*, 56(3), 759–760.
- Holland, J. B. (2007). Genetic architecture of complex traits in plants. *Current Opinion in Plant Biology*, 10(2), 156–161. <https://doi.org/10.1016/j.pbi.2007.01.003>
- Hu, F. y., Tao, D. y., Sacks, E., Fu, B. y., Xu, P., Li, J., ... Li, Z.-k. (2003). Convergent evolution of perenniality in rice and sorghum. *Proceedings of the National Academy of Sciences of the United States of America*, 100(7), 4050–4054. <https://doi.org/10.1073/pnas.0630531100>
- Jannink, J.-L., Lorenz, A. J., & Iwata, H. (2010). Genomic selection in plant breeding: From theory to practice. *Briefings in Functional Genomics*, 9(2), 166–177. <https://doi.org/10.1093/bfpg/elq001>
- Jiang, W., & Yu, D. (2009). *Arabidopsis* WRKY2 transcription factor may be involved in osmotic stress response. *Acta Botanica Yunnanica*, 31(5), 427–432.
- Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, 11, 94. <https://doi.org/10.1186/1471-2156-11-94>
- Jørgensen, U., Mortensen, J., Kjeldsen, J. B., & Schwarz, K. (2003). Establishment, development and yield quality of fifteen *Miscanthus* genotypes over three years in Denmark. *Acta Agriculturae Scandinavica, Section B-Soil & Plant Science*, 53(4), 190–199.
- Karlson, D., & Imai, R. (2003). Conservation of the cold shock domain protein family in plants. *Plant Physiology*, 131(1), 12–15. <https://doi.org/10.1104/pp.014472>
- Lande, R., & Thompson, R. (1990). Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics*, 124(3), 743–756.
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), 357–359. <https://doi.org/10.1038/nmeth.1923>
- Lewandowski, I., Clifton-Brown, J. C., Scurlock, J. M. O., & Huisman, W. (2000). *Miscanthus*: European experience with a novel energy crop. *Biomass and Bioenergy*, 19(4), 209–227. [https://doi.org/10.1016/S0961-9534\(00\)00032-5](https://doi.org/10.1016/S0961-9534(00)00032-5)
- Li, Q., Byrns, B., Badawi, M. A., Diallo, A. B., Danyluk, J., Sarhan, F., ... Fowler, D. B. (2018). Transcriptomic insights into phenological development and cold tolerance of wheat grown in the field. *Plant Physiology*, 176(3), 2376–2394. <https://doi.org/10.1104/pp.17.01311>
- Linde-Laursen, I. B. (1993). Cytogenetic analysis of *Miscanthus* ‘Giganteus’, an interspecific hybrid. *Hereditas*, 119(3), 297–300. <https://doi.org/10.1111/j.1601-5223.1993.00297.x>
- Lipka, A. E., Kandianis, C. B., Hudson, M. E., Yu, J., Drnevich, J., Bradbury, P. J., & Gore, M. A. (2015). From association to prediction: Statistical methods for the association and selection of complex traits in plants. *Current Opinion in Plant Biology*, 24, 110–118.
- Lipka, A. E., Lu, F., Cherney, J. H., Buckler, E. S., Casler, M. D., & Costich, D. E. (2014). Accelerating the switchgrass (*Panicum virgatum* L.) breeding cycles using genomic selection approaches. *PLoS ONE*, 9, e112227.
- Lu, F., Lipka, A. E., Glaubitz, J., Elshire, R., Cherney, J. H., Casler, M. D., ... Costich, D. E. (2013). Switchgrass genomic diversity, ploidy, and evolution: Novel insights from a network-based SNP discovery protocol. *PLoS Genetics*, 9, e1003215. <https://doi.org/10.1371/journal.pgen.1003215>
- Ly, D., Hamblin, M., Rabbi, I., Melaku, G., Bakare, M., Gauch, H. G., ... Jannink, J.-L. (2013). Relatedness and genotype × environment interaction affect prediction accuracies in genomic selection: A study in cassava. *Crop Science*, 53(4), 1312–1325. <https://doi.org/10.2135/cropsci2012.11.0653>
- Maughan, M., Bollero, G., Lee, D. K., Darmody, R., Bonos, S., Cortese, L., ... Voigt, T. (2011). *Miscanthus* × *giganteus* productivity: The effects of management in different environments. *GCB Bioenergy*, 4(3), 253–265.
- Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157(4), 1819–1829.
- Ming, R., DelMonte, T. A., Hernandez, E., Moore, P. H., Irvine, J. E., & Paterson, A. H. (2002). Comparative analysis of QTLs affecting plant height and flowering among closely-related diploid and polyploid genomes. *Genome*, 45(5), 794–803. <https://doi.org/10.1139/g02-042>
- Nie, G., Huang, L., Zhang, X., Taylor, M., Jiang, Y., Yu, X., ... Zhang, Y. (2016). Marker-trait association for biomass yield of potential biofuel feedstock *Miscanthus sinensis* from Southwest China. *Frontiers in Plant Science*, 7, 802. <https://doi.org/10.3389/fpls.2016.00802>
- Paterson, A. H., Bowers, J. E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., ... Rokhsar, D. S. (2009). The *Sorghum bicolor* genome and the diversification of grasses. *Nature*, 457(7229), 551–556.
- Paterson, A. H., Lin, Y. R., Li, Z., Schertz, K. F., Doebley, J. F., Pinson, S. R., ... Irvine, J. E. (1995). Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. *Science*, 269(5231), 1714–1718.

- Paterson, A. H., Schertz, K. F., Lin, Y. R., Liu, S. C., & Chang, Y. L. (1995). The weediness of plants: molecular analysis of genes influencing dispersal and persistence of johnsongrass, *Sorghum halepense* (L.) Pers. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(13), 6127–6131.
- Poland, J. A., Brown, P. J., Sorrells, M. E., & Jannink, J.-L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE*, *7*, e32253. <https://doi.org/10.1371/journal.pone.0032253>
- Puhakainen, T., Hess, M. W., Mäkelä, P., Svensson, J., Heino, P., & Palva, E. T. (2004). Overexpression of multiple dehydrin genes enhances tolerance to freezing stress in *Arabidopsis*. *Plant Molecular Biology*, *54*(5), 743–753. <https://doi.org/10.1023/B:PLAN.0000040903.66496.a4>
- Rayburn, A. L., Crawford, J., Rayburn, C. M., & Juvik, J. A. (2009). Genome size of three *Miscanthus* species. *Plant Molecular Biology Reporter*, *27*, 184–188. <https://doi.org/10.1007/s11105-008-0070-3>
- Riedelsheimer, C., Czedik-Eysenberg, A., Grieder, C., Lisec, J., Technow, F., Sulpice, R., ... Melchinger, A. E. (2012). Genomic and metabolic prediction of complex heterotic traits in hybrid maize. *Nature Genetics*, *44*(2), 217–220. <https://doi.org/10.1038/ng.1033>
- Riedelsheimer, C., Endelman, J. B., Stange, M., Sorrells, M. E., Jannink, J.-L., & Melchinger, A. E. (2013). Genomic predictability of interconnected biparental maize populations. *Genetics*, *194*(2), 493–503. <https://doi.org/10.1534/genetics.113.150227>
- Sacks, E. J., Juvik, J. A., Lin, Q., Stewart, J. R., & Yamada, T. (2013). The gene pool of *Miscanthus* species and its improvement. In A. H. Paterson (Ed.), *Genomics of the Saccharinae* (pp. 73–101). New York, NY: Springer.
- Sasaki, K., & Imai, R. (2012). Pleiotropic roles of cold shock domain proteins in plants. *Frontiers in Plant Science*, *2*, 116. <https://doi.org/10.3389/fpls.2011.00116>
- Schaeffer, L. (2006). Strategy for applying genome-wide selection in dairy cattle. *Journal of Animal Breeding and Genetics*, *123*, 218–223. <https://doi.org/10.1111/j.1439-0388.2006.00595.x>
- Segura, V., Vilhjálmsson, B. J., Platt, A., Korte, A., Seren, Ü., Long, Q., & Nordborg, M. (2012). An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nature Genetics*, *44*(7), 825–830. <https://doi.org/10.1038/ng.2314>
- Seki, M., Narusaka, M., Ishida, J., Nanjo, T., Fujita, M., Oono, Y., ... Shinozaki, K. (2002). Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high salinity stresses using a full-length cDNA microarray. *The Plant Journal*, *31*(3), 279–292. <https://doi.org/10.1046/j.1365-313X.2002.01359.x>
- Slavov, G. T., Davey, C. L., Bosch, M., Robson, P. R. H., Donnison, I. S., & Mackay, I. J. (2018). Genomic index selection provides a pragmatic framework for setting and refining multi-objective breeding targets in *Miscanthus*. *Annals of Botany*, in press. <https://doi.org/10.1093/aob/mcy187>
- Slavov, G. T., Nipper, R., Robson, P., Farrar, K., Allison, G. G., Bosch, M., ... Jensen, E. (2014). Genome-wide association studies and prediction of 17 traits related to phenology, biomass and cell wall composition in the energy grass *Miscanthus sinensis*. *New Phytologist*, *201*(4), 1227–1239.
- Somerville, C., Youngs, H., Taylor, C., Davis, S. C., & Long, S. P. (2010). Feedstocks for lignocellulosic biofuels. *Science*, *329*(5993), 790–792.
- Spindel, J., Begum, H., Akdemir, D., Virk, P., Collard, B., Redoña, E., ... McCouch, S. R. (2015). Genomic selection and association mapping in rice (*Oryza sativa*): Effect of trait genetic architecture, training population composition, marker number and statistical model on accuracy of rice genomic selection in elite, tropical rice breeding lines. *PLoS Genetics*, *11*(2), e1004982. <https://doi.org/10.1371/journal.pgen.1004982>
- Swaminathan, K., Chae, W., Mitros, T., Varala, K., Xie, L., Barling, A., ... Moose, S. P. (2012). A framework genetic map for *Miscanthus sinensis* from RNAseq-based markers shows recent tetraploidy. *BMC Genomics*, *13*, 142. <https://doi.org/10.1186/1471-2164-13-142>
- U.S. Department of Agriculture. (2012). US plant hardiness zone map. Retrieved from <http://planthardiness.ars.usda.gov>
- Wagstaff, C., Bramke, I., Breeze, E., Thornber, S., Harrison, E., Thomas, B., ... Rogers, H. (2010). A specific group of genes respond to cold dehydration stress in cut *Alstroemeria* flowers whereas ambient dehydration stress accelerates developmental senescence expression patterns. *Journal of Experimental Botany*, *61*(11), 2905–2921. <https://doi.org/10.1093/jxb/erq113>
- Wang, Y., Qiao, L., Bai, J., Wang, P., Duan, W., Yuan, S., ... Zhao, C. (2017). Genome-wide characterization of JASMONATE-ZIM DOMAIN transcription repressors in wheat (*Triticum aestivum* L.). *BMC Genomics*, *18*(1), 152. <https://doi.org/10.1186/s12864-017-3582-0>
- Xin, H., Zhu, W., Wang, L., Xiang, Y., Fang, L., Li, J., ... Li, S. (2013). Genome wide transcriptional profile analysis of *Vitis amurensis* and *Vitis vinifera* in response to cold stress. *PLoS ONE*, *8*, e58740. <https://doi.org/10.1371/journal.pone.0058740>
- Yan, J., Chen, W., Luo, F., Ma, H., Meng, A., Li, X., ... Sang, T. (2012). Variability and adaptability of *Miscanthus* species evaluated for energy crop domestication. *GCB Bioenergy*, *4*(1), 49–60. <https://doi.org/10.1111/j.1757-1707.2011.01108.x>
- Yu, J., & Buckler, E. S. (2006). Genetic association mapping and genome organization of maize. *Current Opinion in Biotechnology*, *17*(2), 155–160. <https://doi.org/10.1016/j.copbio.2006.02.003>
- Zhang, C., Fei, S. Z., Warnke, S., Li, L., & Hannapel, D. (2009). Identification of genes associated with cold acclimation in perennial ryegrass. *Journal of Plant Physiology*, *166*(13), 1436–1445. <https://doi.org/10.1016/j.jplph.2009.03.001>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Dong H, Clark LV, Lipka AE, et al. Winter hardiness of *Miscanthus* (III): Genome-wide association and genomic prediction for overwintering ability in *Miscanthus sinensis*. *GCB Bioenergy*. 2019;11:930–955. <https://doi.org/10.1111/gcbb.12615>