Final Report DE-SC0006634:
Quantifying phenotypic and genetic diversity of *Miscanthus sinensis* as a resource for knowledge-based improvement of *M.* ×*giganteus* (*M. sinensis* × *M. sacchariflorus*)

PRINCIPAL INVESTIGATOR: Erik J. Sacks
PI ADDRESS: 279 ERML, MC-051; 1201 W. Gregory Dr.; Urbana, IL 61801
TEL.: 1-217-333-9327. **FAX**: 1-217-244-3637. **EMAIL**: esacks@illinois.edu

CO-INVESTIGATORS:

Joe Brummer	Colorado State University (CSU)	USA
Megan Hall	UC Berkeley (UCB)	USA
Stephen Long	University of Illinois (UIUC)	USA
Junhua Peng	Wuhan Botanical Garden (WBG)	China
Xiaoli Jin	Zhejiang University (ZJU)	China
Toshihiko Yamada	Hokkaido University (HU)	Japan
Chang Yeon Yu	Kangwon National University (KNU)	Korea

COLLABORATORS:

Patrick Brown	University of Illinois (UIUC)	USA
Kweon Heo	Kangwon National University (KNU)	Korea
John Juvik	University of Illinois (UIUC)	USA
Nora Lapitan	Colorado State University (CSU)	USA
Stephen Moose	University of Illinois (UIUC)	USA
Daniel Rokhsar	UC Berkeley (UCB)	USA
Kankshita Swaminathan	University of Illinois (UIUC)	USA
Dean Tiessen	New Energy Farms (NEF)	Canada

DOE/Office of Science Program Office: Office of Biological and Environmental Research
DOE/Office of Science Program Manager Contact: Dr. Catherine M. Ronning

Project objectives:

A. Determine population structure for a core collection of ~500 *M. sinensis* genotypes from throughout the species natural distribution in China, Japan and Korea to identify unique groups. The results of this work will for the first time provide *Miscanthus* breeders knowledge-based hypotheses about which cross-combinations are most likely to be heterotic.
B. Quantify phenotypic variation in *M. sinensis* for key traits (primarily yield, yield-components and stability, flowering time, overwintering ability, low temperature photosynthesis and leaf extension, and drought tolerance) at field trial sites in the US, Canada and Asia, and assess the effects of genotype × environment (GxE) interactions, to facilitate the rational choice of parents for remaking improved versions of *M. ×giganteus* adapted to US production environments.
C. Kick-start *Miscanthus* breeding by identifying genes governing key traits, through the testing of candidate genes from other grasses, and genome-wide association mapping.

Accomplishments, and impacts:

Population structure:

In total, 768 accessions of *M. sinensis* and related out-group species were studied, including 703 *M. sinensis* accessions from their native range in China, Japan, and S. Korea, in addition to US-sourced ornamental cultivars, and US naturalized populations (Table 1). For out-groups, we included the following species: *M. floridulus*, *M. oligostachyus*, *M. sacchariflorus*, and one *Saccharum officinarum* (sugarcane). The accessions were screened with 21,207 RAD-Seq SNPs obtained via the UNEAK pipeline in TASSEL, and 424 GoldenGate SNPs. Additionally, all accessions were screened with ten chloroplast simple sequence repeat markers (SSRs). Population structure was analyzed with STRUCTURE 2.3.4, and via Discriminant Analysis of Principle Components (DAPC) using the R package adegenet.

The results from our study have, for the first time, provided a regional-level understanding of population structure in *M. sinensis*, and insights into its most recent phase of evolution. Initially, we identified six populations of *M. sinensis* from geographically distinct regions in Asia, including two groups in Japan (Fig. 1); a subsequent analysis with greater sampling from Japan revealed three groups in Japan (North, Central, and South; Fig. 2) for a total of seven *M. sinensis* groups. Both nuclear and chloroplast markers indicated that SE China is currently the center of genetic diversity for *M. sinensis*. *An especially impactful conclusion of this study is that nearly all M. sinensis germplasm in the US prior to this study (mostly ornamental cultivars and naturalized populations) originated from a small region of southern Japan. Thus, this study has substantially increased the genetic diversity available for breeding Miscanthus.*

During the last glacial maximum (LGM), when the climate in Asia was colder and drier, the distribution of *M. sinensis* north of tropical Eastern Asia was likely limited to a refugium in present day SE China and coastal areas from S. Taiwan to Hainan that are now mostly submerged under the South China Sea, as these areas had sufficiently warm and moist environments to support temperate steppe and tropical grassland vegetation, respectively. In modern times, *M. sinensis* is rarely observed in climates colder than USDA hardiness zone 5, and thus during the LGM *M. sinensis* would not have been adapted to the boreal climate of the Sichuan basin or land that is now part of Southern Japan. After the LGM, as the climate subsequently warmed at got wetter, *M. sinensis* radiated out into newly hospitable geographies, such as the Sichuan basin, Japan, and the Korean peninsula, evolving into the populations that we observe today. Thus, SE China is likely the origin of the *M. sinensis* populations currently found in temperate eastern Asia. Curiously, the four *M. floridulus* we studied, which were from China, Japan, New Caledonia and Papua New Guinea, grouped with the SE China population of *M. sinensis*, calling into question the validity of two different species names. Moreover, the inclusion of *M. floridulus* in the SE

China population raises the intriguing question of whether *M. sinensis* emerged as a distinct species in SE Asia then migrated to the South Pacific, or if *Miscanthus* originated in New Guinea, which is a center of diversity for its sister genus *Saccharum*, and migrated to Asia; further sampling of *M. floridulus* will be needed to resolve this question.

In the US, we found that nearly half of ornamental cultivars labeled *M. sinensis* were in fact BC_1 or BC_2 hybrids of *M. sinensis* and *M. sacchariflorus*, with the former as the female recurrent parent species. Naturalized *M. sinensis* populations in the US were derived from ornamental *M. sinensis*, but not the backcross interspecific hybrids, suggesting a lack of fitness of the latter. With the exception of *M. sinensis* var. *transmorrisonensis*, which originates from Taiwan and is part of the SE China population, most ornamental and all naturalized *M. sinensis* in the US were derived from a subset of the Southern Japan population, indicating a considerable genetic bottleneck. A nursery industry in Southern Japan that engaged in international trade at least as early as the late 1800's, as it does today, likely accounts for the origin of most ornamental *M. sinensis* germplasm in the US and Europe.

The genetic bottleneck associated with US *M. sinensis* germplasm was a previously unknown limitation to breeding improved bioenergy feedstock cultivars of *Miscanthus.* Triploid *M.* ×*giganteus* (Mxg), which is among the most promising bioenergy crops for the US, is a nothospecies, derived from hybridization between tetraploid *M. sacchariflorus* and diploid *M. sinensis.* The Mxg genotype that currently predominates in the US and Europe, which we call 'Illinois', was introduced to Denmark from Southern Japan in the 1930's. Recently, additional triploid genotypes of Mxg have been bred by us at the Univ. of Illinois and by colleagues in Europe, but all of these have been derived from the Southern Japan *M. sinensis* population. The present study indicates that there is an opportunity to broaden the genetic diversity of Mxg by using *M. sinensis* parents from populations other than Southern Japan. Such use of other *M. sinensis* populations may allow access to new adaptation genes, such as cold-tolerance from the Northern Japan or NE China/Korea populations, and may enable *Miscanthus* breeders to potentially capture greater heterosis than is possible with the Southern Japan population.

In a subsequent study focused on Japan, we found evidence for a gradient of natural introgression of genes from diploid *M. sinensis* into tetraploid *M. sacchariflorus*, especially in Southern Honshu (Fig. 2). This insight indicated that production of unreduced gametes by *M. sinensis* has resulted tetraploid interspecific hybrids that have high genetic diversity. Differing contributions of *M. sinensis* in the tetraploid interspecific hybrids likely influences the length of rhizomes and number of tillers per unit area, which are important biomass traits. Thus, our study suggests that breeders wishing to develop the highest yielding triploid Mxg using Japanese materials should consider the amount of introgression from *M. sinensis* into tetraploid *M. sacchariflorus.*

<u>Material exchange & germplasm conservation</u>:
In September 2011, the collaborators on the project signed a single MTA that is consistent with the Convention on Biological Diversity, allowing the participants to share germplasm, and providing the option for further breeding and commercialization via benefit-sharing with providers. In Q1 of 2012, the collaborators in China, Japan, Korea and the US sent dormant plant materials to our collaborator in Canada for propagation. In Q2 and Q3 of 2012, our collaborator in Canada distributed propagated materials to each of the other collaborators. US quarantine restrictions on clonal materials of grasses required us, as expected, to send *Miscanthus* divisions to USDA's quarantine facility in Beltsville, MD. Given the bottleneck of limited space at USDA's quarantine facility, we used the results of our population structure and genetic diversity analyses to prioritize importation of each group identified and maximize the allelic diversity of the sets prioritized for importation. Using these molecular genetic

guides, we were able to import germplasm that included ~90% of the allelic diversity identified in the entire collection. Capitalizing on our identification of three distinct groups of *M. sinensis* in Japan and leveraging past germplasm collections of seed, which have less stringent quarantine requirements in the US, we conducted a seed increase of Japanese materials using ~200 individual plants in isolated greenhouse bays at Urbana, IL; this seed is now safely conserved in a controlled temperature and humidity seed-storage room at UIUC. Additionally, ~200 clonal accessions from China and Korea have been imported through the USDA quarantine facility.

Quantify phenotypic variation and identifying genes governing key traits:
Flowering time. Replicated field trials were established during spring 2012 in IL, CO, Canada, Japan, and Korea (Table 2). Data on flowering time, yield and yield-component traits was taken in 2013 and 2014. Data from the field trials showed great variation among all the genotypes and groups for each of the traits evaluated.

Flowering time is a key trait for adaptation and yield-potential in grasses. For example, the current commercial cultivar of Mxg is poorly adapted to the southern US because it flowers too early (late July-early August in Alabama and Mississippi). Data from our field trials found a huge range of heading dates for *M. sinensis*, from mid-July to early-November (Fig. 3). Moreover, we observed large differences among groups, with Sichuan Basin and SE China groups flowering latest and considerably later than the Southern Japan group. *Thus, this study points the way to developing Mxg cultivars that will flower later and be better adapted to the southern US than the current Mxg of commerce.*

In addition to identifying accessions and groups that have the desirable late flowering trait, we found 41 significant SNPs for flowering time by GWAS using an FDR-corrected p-value of 0.05. The GWAS was conducted with the software RR-BLUP on 554 *M. sinensis* accessions and 38,597 markers aligned to Sorghum. The marker-trait associations can be used to greatly accelerate breeding progress in *Miscanthus*, because seedlings can now be selected shortly after germination rather than waiting at least two years from planting to obtain phenotypic data.

We also observed substantial sequence variation for a key flowering time, *Heading date 1 (Hd1). Hd1* in *Oryza sativa* acts as repressor to flowering under long day conditions, and we hypothesized that *Hd1* plays a similar role in *M. sinensis.* We compared sequences of *MsiHd1* homologs among 24 wild *M. sinensis* accessions from Japan, 14 from China and 3 from South Korea, covering a broad range of latitudes. Two to five *MsiHd1* alleles in each accession were identified, suggesting that *MsiHd1* consists of at least three loci in the *Miscanthus* genome. Verifying the open reading frame in *MsiHd1*, they were classified as putative functional alleles without mutations or non-functional alleles caused by indels. One of the multiple *MsiHd1* loci is a pseudogene locus without any functional alleles, which we named *MsiHd1b*, and the other loci were considered to be part of the *MsiHd1a* multi-locus family. Interestingly, in most Japanese accessions half or more of the *MsiHd1a* alleles were non-functional, whereas accessions from the East Asian mainland harbored only functional alleles (Fig. 4). As expected from previous studies, there was a latitudinal gradient of heading date for both the accessions from mainland Asia and for those from Japan, with northern accessions heading earlier in the growing season than southern ones when grown in a common garden in Sapporo. The differences in *MsiHd1a* show that dependency on functional *MsiHd1a* alleles is different between accessions from the East Asian mainland and Japan, and suggest that crossing mainland individuals with those from Japan at similar latitudes would likely produce transgressive segregants for flowering time.

Yield and yield components. Great variation in biomass yield was observed among the accessions evaluated in replicated field trials (Fig. 5). With the exception of the Northern Japan group, which had lower yields than most of the other wild *M. sinensis* groups, most of the variation for yield was among accessions within group, indicating the potential to breed high biomass cultivars with adaptation to different environments using genetically diverse sources of breeding materials. The ornamental cultivars and the US naturalized group were also lower yielding on average than all of the Asian native groups, except for the Northern Japan group, indicating the great value in using the wild native populations for breeding high biomass cultivars rather than the ornamental cultivars or the US naturalized populations. Especially notable was the exceptionally high biomass yields of the diploid *M. sacchariflorus* × *M. sinensis* hybrids, indicating that heterosis and/or transgressive segregation typically provides a large and favorable yield boost (about double the yield of the parents). The boost in yield of the interspecific hybrids was substantially conferred by a large increase in vigor and especially the number of culms per plant and per unit area. We identified 12 SNPs significantly associated with biomass yield using an FDR-corrected p-value of 0.05. Given the great time and expense needed to phenotype for yield in breeding populations, these tagged QTL for biomass yield are expected to greatly improve the efficiency of breeding for high biomass *Miscanthus.*

Deliverables:

Publications:

1) Clark, L.V., *K. Głowacka,* J.E. Brummer, *M. Hall*, K. Heo, X. Jin, J. Peng, T. Yamada, J.H. Yoo, C.Y. Yu, H. Zhao, S.P. Long, and E.J. Sacks. (In preparation). Novel genes for flowering time identified in *Miscanthus sinensis*.
2) Clark, L.V., K. Głowacka, J.E. Brummer, M. Hall, K. Heo, X. Jin, J. Peng, T. Yamada, J.H. Yoo, C.Y. Yu, H. Zhao, S.P. Long, and E.J. Sacks. (In preparation). Genome-wide association analyses reveal QTLs for biomass yield in *Miscanthus sinensis.*
3) Clark, L.V., J.E. Brummer, K. Heo, X. Jin, J. Peng, T. Yamada, J.H. Yoo, C.Y. Yu, H. Zhao, S.P. Long, H.D. Upadhyaya, J. Yu and E.J. Sacks. (In preparation). How to select a core collection of *Miscanthus sinensis* to facilitate breeding for biomass: a comparison of selection for marker diversity, phenotypic selection, and genomic selection.
4) Nagano, H., L.V. Clark, H. Zhao, J. Peng, J.H. Yoo, K. Heo, C.Y. Yu, K.G. Anzoua, T. Matsuo, E.J. Sacks, and T. Yamada. 2015. Contrasting allelic distribution of *Co/Hd1* homologs in *Miscanthus sinensis* from the east Asian mainland and the Japanese archipelago. J. Exp. Bot. doi: 10.1093/jxb/erv292.
5) Clark, L.V., J.R. Stewart, A. Nishiwaki, Y. Toma, J.B. Kjeldsen, U. Jørgensen, H. Zhao, J. Peng, J.H. Yoo, K. Heo, C.Y. Yu, T. Yamada, and E.J. Sacks. 2015. Genetic structure of *Miscanthus sinensis* and *M. sacchariflorus* in Japan indicates a gradient of bidirectional but asymmetric introgression. J. Exp. Bot. doi:10.1093/jxb/eru51.
6) Tamura K.I., Y. Sanada, A. Shoji, K. Okumura, N. Uwatoko, K.G. Anzoua, E.J. Sacks, and T. Yamada. 2015. DNA markers for identifying interspecific hybrids between *Miscanthus sacchariflorus* and *Miscanthus sinensis.* Grassland Science doi: 10.1111/grs.12089.
7) Clark, L.V., J.E. Brummer, K. Głowacka, M. Hall, K. Heo, J. Peng, T. Yamada, J.H. Yoo, C.Y. Yu, H. Zhao, S.P. Long, and E.J. Sacks. 2014. A footprint of past climate change on the diversity and population structure of *Miscanthus sinensis.* Annals of Botany 114:97-107.

Presentations:

1) Sacks, E.J. Genomics and breeding of *Miscanthus*. International Plant & Animal Genome XXIII Conference. San Diego, CA. 10-14 January 2015.
2) Clark, L.V., J.E. Brummer, K. Głowacka, M. Hall, K. Heo, J. Peng, T. Yamada, J.H. Yoo, C.Y. Yu, H. Zhao, S.P. Long, and E.J. Sacks. Genome-wide association analysis of flowering time in *Miscanthus sinensis*. International Plant & Animal Genome XXIII Conference. San Diego, CA. 10-14 January 2015.
3) Sacks, E.J. *Miscanthus* Genetic Diversity and Germplasm Resources: Opportunities for Breeding New Bioenergy Cultivars. Tenth Annual Bioenergy Feedstocks Symposium. Champaign, IL. 24-25 September 2014.
4) Clark, L.V., and E.J. Sacks. Diversifying the cultivars of *Miscanthus* and Miscane available for bioenergy feedstocks through breeding and genomics. 4th Pan-American Congress on Plants and BioEnergy. University of Guelph Conference Center. Guelph, ON. 4-7 June 2014.
5) Clark, L.V., J.E. Brummer, K. Glowaca, M. Hall, K. Heo, J. Peng, T. Yamada. J.H. Yoo, C.Y. Yu., H. Zhao, S. Long, and E.J. Sacks. Genetic diversity and population structure of *Miscanthus sinensis*. International Plant & Animal Genome XXII Conference. San Diego, CA. 11-15 January 2014.
6) Nagano, H., N. Uchino, J. Peng, E.J. Sacks, and T. Yamada. Sequence diversity of *Co/Hd1* homologs in *Miscanthus sinensis*. International Plant & Animal Genome XXII Conference. San Diego, CA. 11-15 January 2014.
7) Sugisawa, S., H. Nagano, M. Dwiyanti, E.J. Sacks, and T. Yamada. Relation of Pseudo-Response Regulator (PRR)-like gene to flowering time in *Miscanthus* genome. 124th Meeting of the Japanese Society of Breeding. October 2013.
8) Nagano, H., N. Uchino, E.J. Sacks, and T. Yamada. Paralogue in *Hd1* gene of *Miscanthus sinensis*. 123th Meeting of the Japanese Society of Breeding. March 2013.
9) Clark, L.V., J.E. Brummer, M. Hall, S. Long, J. Peng, T. Yamada. J.H. Yoo, C.Y. Yu., H. Zhao, and E.J. Sacks. Genetic structure of *Miscanthus sinensis* from Asia and the United States. International Plant & Animal Genome XXI. San Diego, CA. 12-16 January 2013.
10) Clark, L.V., T. Yamada, J.R. Stewart, and E.J. Sacks. Genetic diversity and population structure of *Miscanthus* in Japan. 3rd Pan American Congress on Plants and Bioenergy, American Society of Plant Biologists. Champaign, IL. 15-18 July 2012.
11) Sacks, E. Breeding for better bioenergy grasses. 4th International Energy Farming Congress. Papenburg, Germany. 13-15 March 2012.

Training:

Trainee	Under-graduate	Graduate	Post-doctoral	Technical	Institution
Chantel Campbell	X				CSU
Julie Felt	X				CSU
Kristina Gosselin	X				CSU
Kyle Madsen	X				CSU
Lane Todd	X				CSU
Luis Villalobos		X			CSU
Lyndsay Jones		X			CSU
Victoria Marrazzo	X				CSU
Hironori Nagano				X	HU
Kossonou Guillaume Anzoua			X		HU
Maiko Ohta				X	HU
Tomoaki Matsuo	X				HU
Jae Hoo Choi		X			KNU
Ji Hye Yoo		X			KNU
Benjamin Baechle				X	UIUC
Christine Fleener	X				UIUC
Jack Cygan	X				UIUC
Kasia Dubiel	X				UIUC
Katarzyna Głowacka			X		UIUC
Lily Hislop	X				UIUC
Lindsay Clark			X		UIUC
Logan Smith				X	UIUC
Maertens, Colten				X	UIUC
Megan Swanson	X				UIUC
Melina Salgado	X				UIUC
Travis Hurt				X	UIUC
Yingying Zheng	X				UIUC
Hua Zhao			X		WBG
Chen Lin		X			ZJU
Liang Zhu				X	ZJU
Mei Li				X	ZJU
Yifei Jin				X	ZJU
Count	14	5	4	9	32

Collaborations initiated after the project started:

Collaborator	Institution
Jens Bonderup Kjeldsen	Aarhus University
Uffe Jørgensen	Aarhus University
Ryan Stewart	Brigham Young University
Yo Toma	Ehime University
Aya Nishiwaki	University of Miyazaki

Table 1. Entries included in the study of population structure.

Species	Origin	# of entries
*M. sinensis**	China	296
	Japan**	136
	S. Korea	184
	Taiwan	1
	Ornamental cultivars***	43
	US naturalized	43
M. floridulus	China	1
	Japan****	1
	New Caledonia****	1
	Papua New Guinea****	1
M. sinensis x *M. sacchariflorus* hybrids	China[a]	11
	Ornamental cultivars[b]	34
M. sacchariflorus	China	8
	S. Korea	3
	Ornamental cultivars	1
M. oligostachyus	Ornamental cultivars	1
M. sp.	Taiwan****	1
	unknown****	1
S. officinarum	Malaysia	1
Total		768

*includes varieties *condensatus* (3), *purpurascens* (1) and *transmorrisonensis* (2; presumed to have originated in Taiwan).
**includes 4 cultivars bred for biomass/forage.
***From US nurseries.
****From USDA NPGS.
[a]F_1s and BC_1 found in the wild.
[b]BC_1s and BC_2s with *M. sinensis* as the recurrent parent.

Table 2. Summary of field trial sites.

Site	# genotypes trialed	Lat	Long	Elev (m)	Hardiness Zone
Fort Collins, CO	150	40.7	-105.0	1556	5
Urbana, IL	150	40.1	-88.2	230	6
Leamington, ON	625	42.1	-82.6	196	6
Sapporo, Japan	625	43.1	141.4	11	6
Chunchon, Korea	625	37.5	127.4	75	6-7
Zhuji, China	625	29.8	120.2	58	8

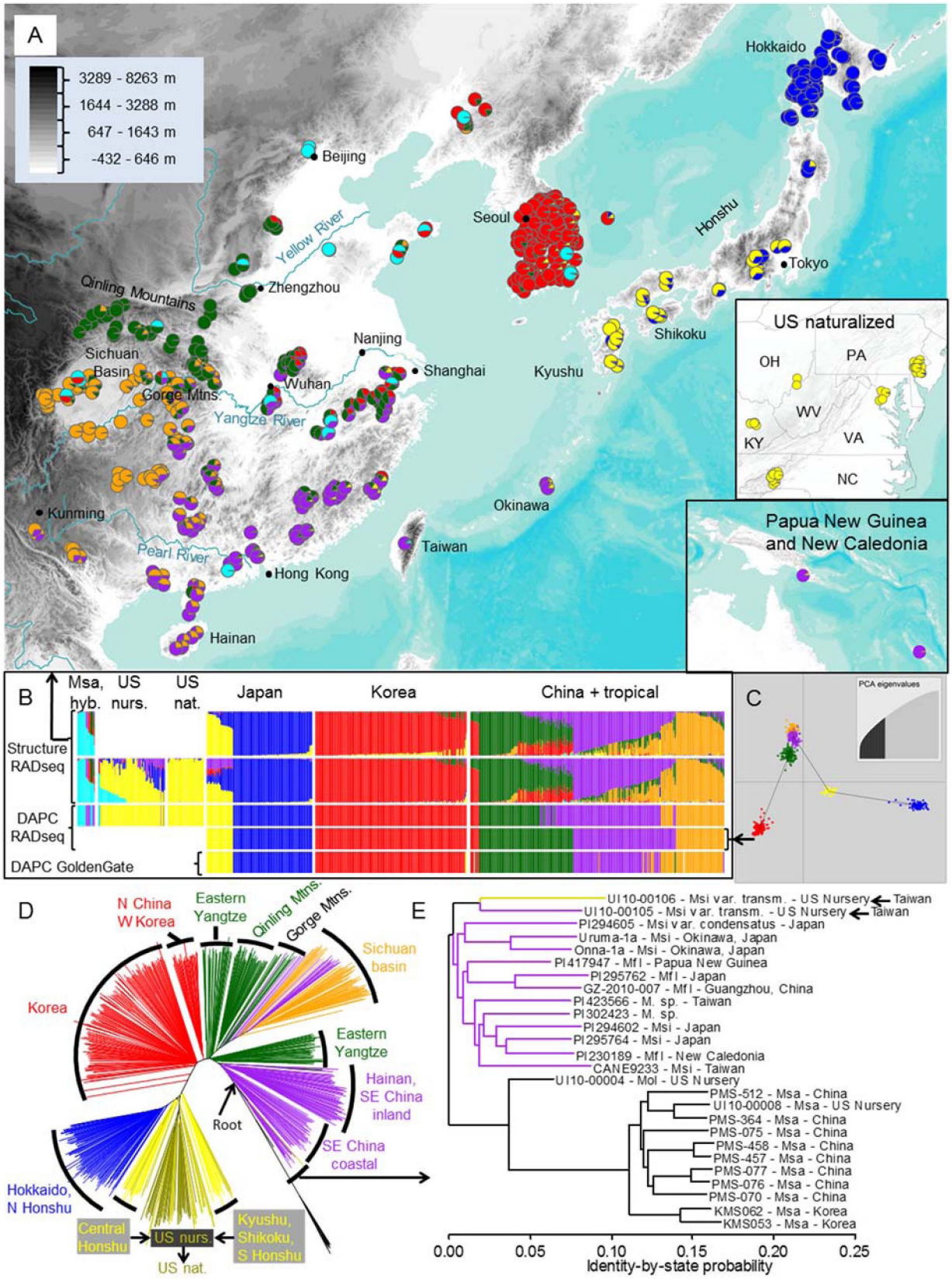

Figure 1. Genetic clustering of *Miscanthus* using 21,207 RAD-seq SNPs and 424 GoldenGate SNPs. Arrows from C to B, B to A, and D to E show connections between analyses as described below. A) Map

of collections in the native range. Each accession is represented by a pie chart showing ancestry (Q) among seven genetic clusters as determined by analysis with the software Structure. Elevation in Asia is shown as a grey scale. B) Bar charts showing Q values from Structure analysis or posterior probabilities of assignment to groups based on discriminant analysis of principal components (DAPC). The top bar chart, representing 641 individuals from the native range, was used to create the pie charts seen in A), with the exception of pie charts in the US, which were colored based on the second bar chart from the top, which contains an additional 120 individuals from the United States and four Japanese biomass cultivars. Msa = *M. sacchariflorus*; hyb. = natural *M. sacchariflorus* × *M. sinensis* hybrids; US nurs. = US nurseries and Japanese biomass cultivars; US nat. = US naturalized accessions. C) Scatterplot depicting relationship between the six *M. sinensis* clusters as determined by DAPC of RAD-seq data from 620 *M. sinensis* and *M. floridulus* from the native range. The first two discriminant axes are plotted. Clusters are connected by a minimum spanning tree. Eigenvalues are shown for the first 200 principal components, which were those included in the analysis. D) Neighbor-joining tree of 722 non-hybrid individuals, derived from identity-by-state probabilities in TASSEL 3.0 from RAD-seq data. "Root" indicates the node at which the *M. sacchariflorus* outgroup connects to the rest of the tree. Branches are colored based on DAPC groupings as in C), with the exception of individuals from the United States. E) Subset of neighbor-joining tree from D), showing groups adjacent to the root of the tree.

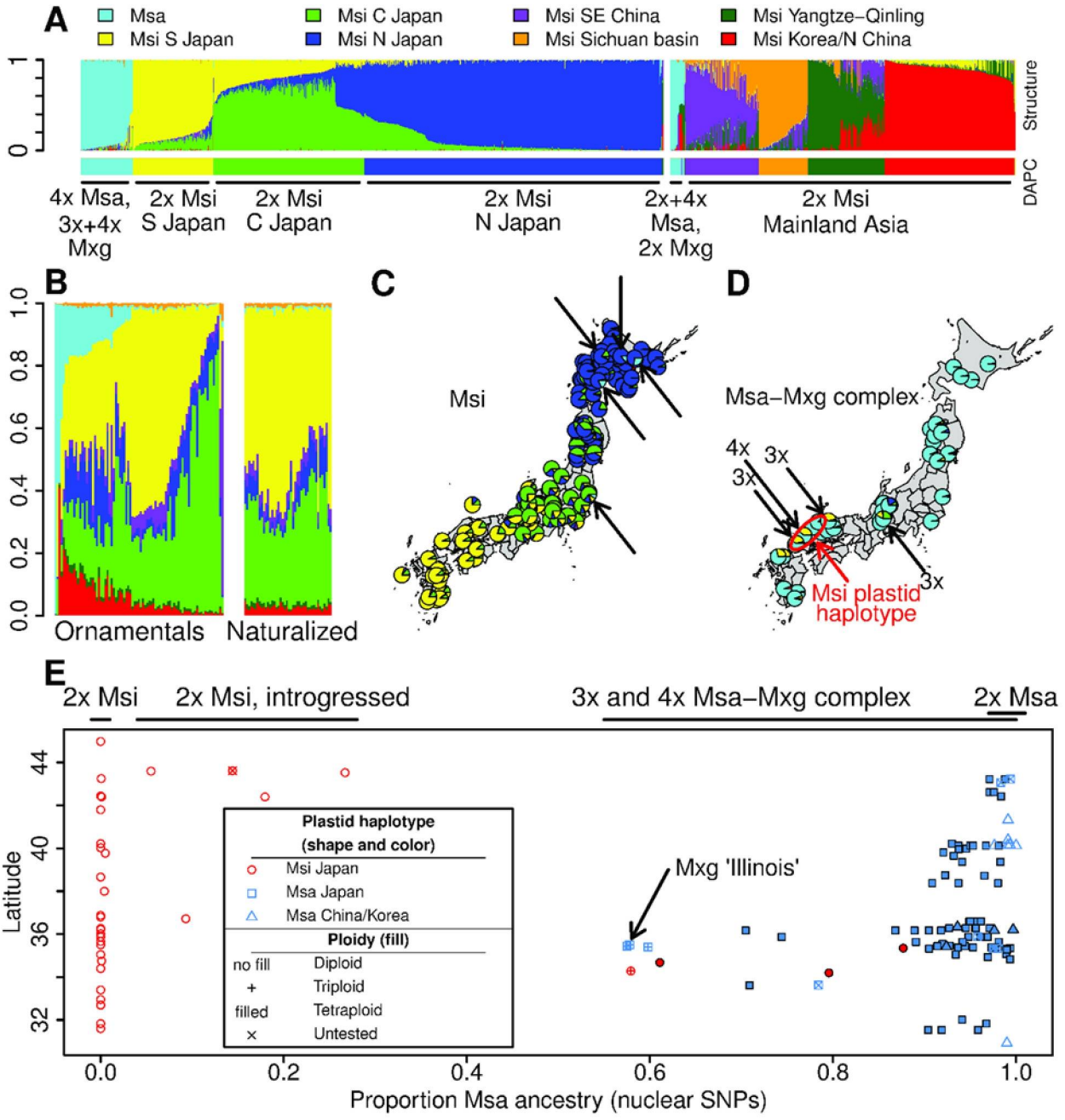

Figure 2. Structure and DAPC results using 20,704 nuclear SNPs. Msi = *Miscanthus sinensis*, Msa = *M. sacchariflorus,* Mxg = *M.* ×*giganteus.* (A) Barplot of Q values (proportion ancestry estimated in Structure) for 745 individuals from the Japan dense-sampling set and 645 individuals from the previously-published region-wide set (Clark *et al.*, 2014). Each of five runs included 253 individuals from the Japan-dense set (one per accession) plus all 645 individuals from the region-wide set; mean Q values are shown for individuals that were present in more than one run. Each of the eight groups is represented by a different color. The narrower bottom bar indicates DAPC group assignments. (B) Mean Q values for 81 ornamental individuals and 42 naturalized individuals from the USA, when the parameters USEPOPINFO and PFROMPOPFLAGONLY were used in Structure to assign ancestry from native populations. (C) Map of Q values for Msi individuals in Japan, including 667 from the Japan-dense

set and 128 from the region-wide set. Five individuals with Msa ancestry 6%-27%, including four diploids and one of undetermined ploidy, are indicated with arrows. (D) Map of Q values for 78 Msa-Mxg complex individuals from Japan, all from the Japan dense-sampling set. Four individuals with Msi ancestry 39-42% are indicated with arrows, and the ploidy determined by flow cytometry is indicated; all other individuals shown were tetraploid except for six of undetermined ploidy. The red ellipse indicates the sampling area for all four Msa-Mxg individuals with an Msi plastid haplotype (other individuals within the ellipse have an Msa plastid haplotype). (E) Latitude vs. Q values for 89 native-collected Msa-Mxg complex individuals, Mxg 'Illinois' (assuming origin in Yokohama, Japan; indicated with an arrow), 28 random Msi individuals that were subjected to flow cytometry, and five Msi individuals with Msa ancestry > 5%. Color and shape of symbols in (E) are used redundantly to indicate plastid haplotype and collection location, and fill is used to indicate ploidy, with filled points outlined in black to make them more easily visible.

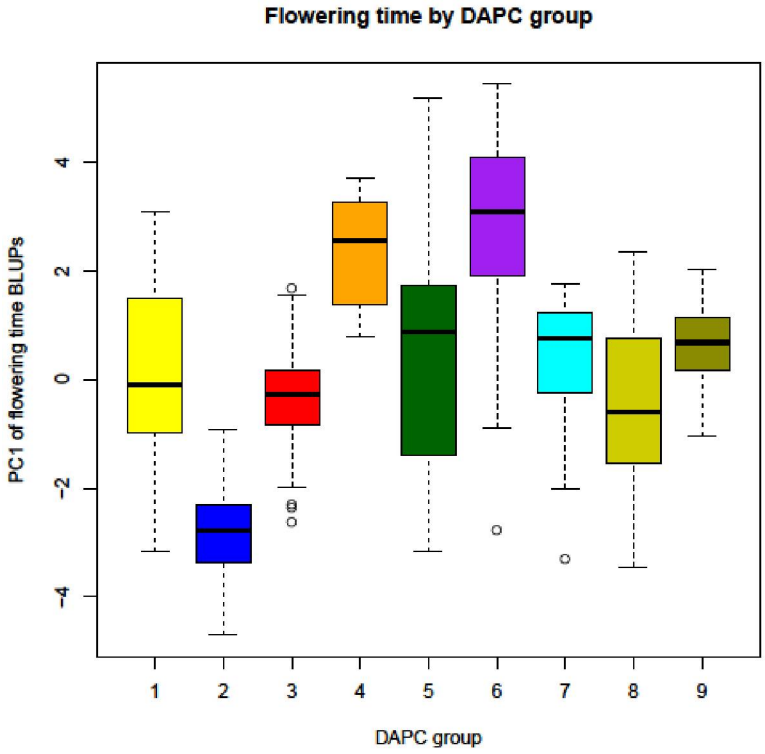

Figure 3. Flowering time boxplots of *M. sinensis* shown by DAPC group (color coding of groups is the same as for Fig. 1). On the PC1 axis, a value of -4 correlates to a median first heading date of about mid-July, 0 to mid-September, +4 to mid-October. Note that the Northern Japan group (blue) is exceptionally early but the Sichuan Basin (orange) and SE China (purple) groups flowered the latest.

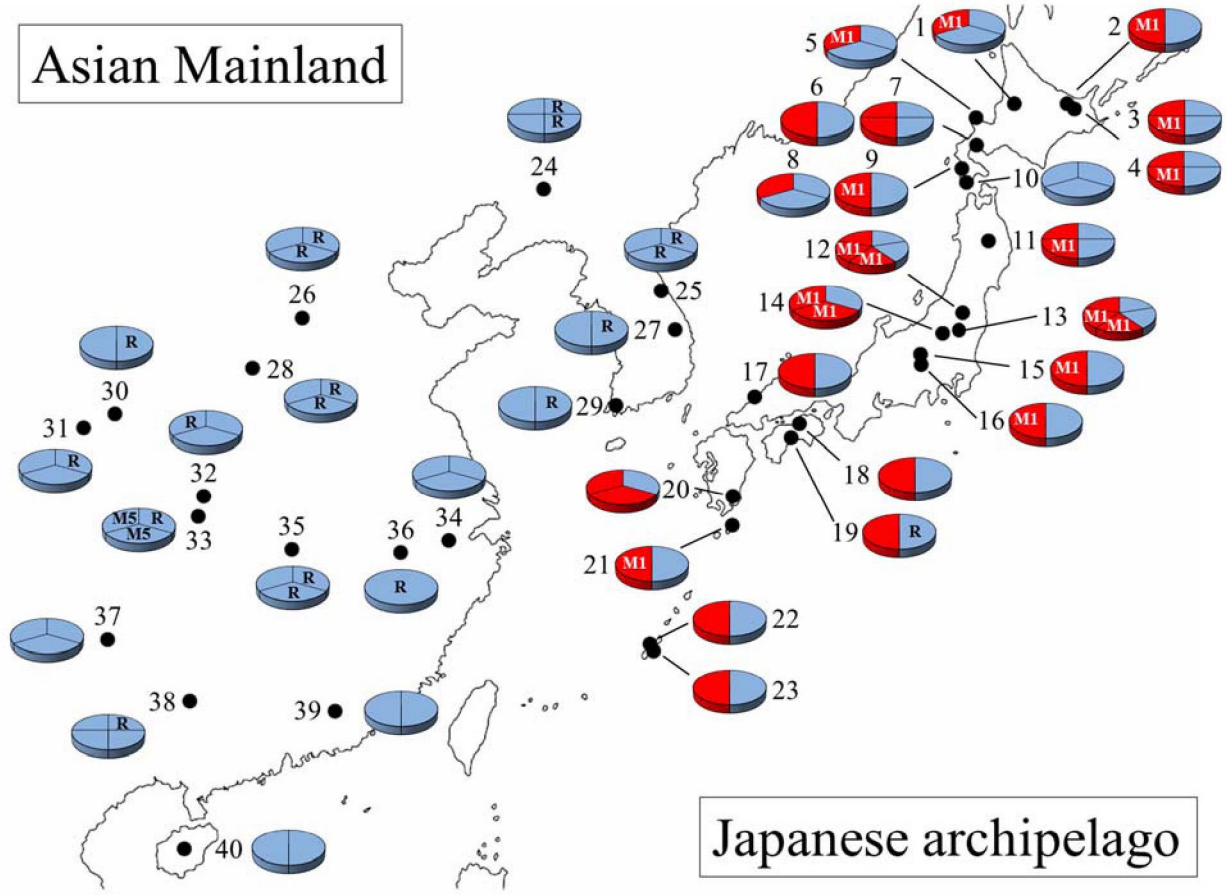

Figure 4. Geographic distribution of *Hd1a* alleles in *M. sinensis* accessions from East Asia. Information regarding *MsiHd1b* was omitted for simplification. Pie charts with one to five pieces represent the number of detected alleles in *MsiHd1a*. Blue indicates functional alleles and red indicates non-functional alleles. R = Revertant allele from *MsiMITE1*, M1 = *MsiMITE1*, and M5 = *MsiMITE5*.

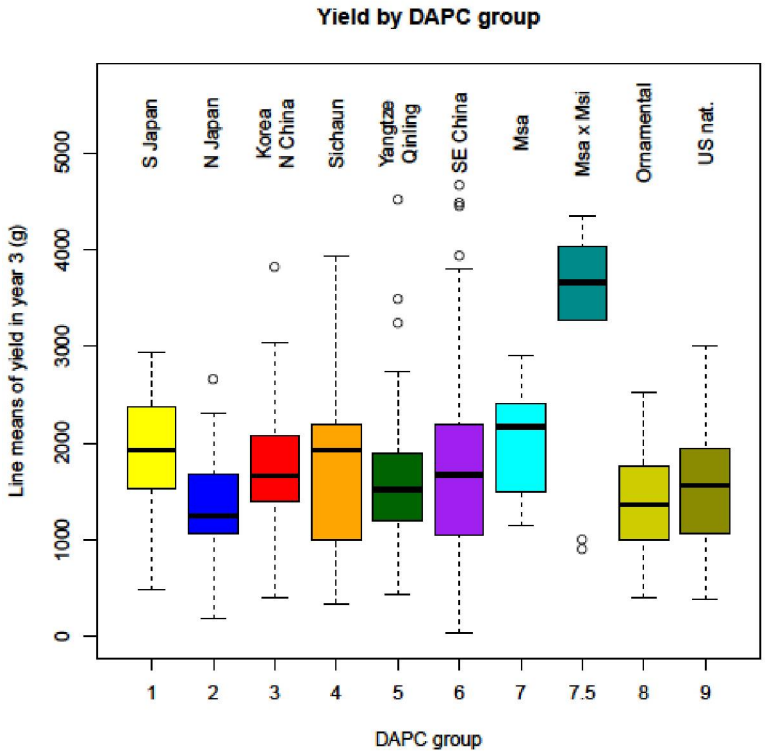

Figure 5. Boxplots of average biomass yield per plant (dry weight in g) for accessions of *M. sinensis* shown by DAPC group. Data summarizes yield over three northern location trial sites: Leamington, ON; Sapporo, Japan; and Chunchon, Korea.