Annals of Botany **114**: 97–107, 2014
doi:10.1093/aob/mcu084, available online at www.aob.oxfordjournals.org

ANNALS OF BOTANY
Founded 1887

A footprint of past climate change on the diversity and population structure of *Miscanthus sinensis*

Lindsay V. Clark[1], Joe E. Brummer[2], Katarzyna Głowacka[1,3], Megan C. Hall[4], Kweon Heo[5], Junhua Peng[6], Toshihiko Yamada[7], Ji Hye Yoo[5], Chang Yeon Yu[5], Hua Zhao[6], Stephen P. Long[1] and Erik J. Sacks[1,*]

[1]*University of Illinois, Urbana-Champaign, Urbana, IL 61801, USA,* [2]*Colorado State University, Fort Collins, CO 80523, USA,* [3]*Institute of Plant Genetics, Polish Academy of Sciences, 60-479 Poznań, Poland,* [4]*Bio Architecture Lab, Berkeley, CA 94710, USA,* [5]*Kangwon National University, Chuncheon, Gangwon 200-701, South Korea,* [6]*Huazhong Agricultural University, Wuhan, Hubei 430070, China and* [7]*Hokkaido University, Sapporo, Hokkaido 060-0811, Japan*
** For correspondence. E-mail esacks@illinois.edu*

Received: 6 November 2013 Returned for revision: 7 January 2014 Accepted: 28 March 2014 Published electronically: 10 June 2014

• **Background and Aims** *Miscanthus* is a perennial C_4 grass that is a leading potential feedstock crop for the emerging bioenergy industry in North America, Europe and China. However, only a single, sterile genotype of *M.* × *giganteus* (*M*×*g*), a nothospecies derived from diploid *M. sinensis* (*Msi*) and tetraploid *M. sacchariflorus* (*Msa*), is currently available to farmers for biomass production. To facilitate breeding of *Miscanthus*, this study characterized genetic diversity and population structure of *Msi* in its native range of East Asia.
• **Methods** A total of 767 accessions were studied, including 617 *Msi* from most of its native range in China, Japan and South Korea, and 77 ornamental cultivars and 43 naturalized individuals from the USA. Accessions were evaluated with 21 207 restriction site-associated DNA sequencing single nucleotide polymorphism (SNP) markers, 424 GoldenGate SNPs and ten plastid microsatellite markers.
• **Key Results** Six genetic clusters of *Msi* from geographically distinct regions in Asia were identified. Genetic data indicated that (1) south-eastern China was the origin of *Msi* populations found in temperate eastern Asia, which is consistent with this area probably having been a refugium during the last glacial maximum (LGM); (2) *Msi* migrated directly from south-eastern China to Japan before migrating to the same latitudes in China and Korea, which is consistent with the known sequence of warming post-LGM; (3) ornamental *Msi* cultivars were derived from the southern Japan population, and US naturalized populations were derived from a sub-set of the ornamental cultivars; and (4) many ornamental cultivars previously described as *Msi* have hybrid ancestry from *Msa* and *Msi*, whereas US naturalized populations of *Msi* do not.
• **Conclusions** Population structure of *Msi* was driven by patterns of warming since the LGM, and secondarily by geographical barriers. This study will facilitate germplasm conservation, association analyses and identification of potential heterotic groups for the improvement of *Miscanthus* as a bioenergy crop.

Key words: Andropogoneae, bioenergy, biogeography, plastid microsatellite, climate change, *Miscanthus sinensis*, population genetics, Saccharinae, single nucleotide polymorphism.

INTRODUCTION

Miscanthus × *giganteus* (*M*×*g*) is a high-yielding, perennial, C_4 grass that is an important feedstock crop for the emerging bioenergy industry in the USA and Europe (Somerville *et al.*, 2010). A key distinctive feature of *M*×*g* relative to other Andropogoneae crops is its adaptation to temperate environments as a long-lived perennial, and in particular its ability to maintain exceptionally high rates of photosynthesis at temperatures $<14\,^{\circ}C$, whereas sugarcane, sorghum and maize cannot (Long and Spence, 2013). However, only a single, sterile triploid genotype of *M*×*g*, which was imported from Yokohama, Japan, to Denmark in the 1930s, is currently available to US farmers (Hodkinson, 2002; Głowacka *et al.*, 2014). The development of a new crop based on a single genotype is highly risky, because the emergence of a virulent disease or insect strain could damage a large proportion of plantings in a short time. Moreover, new cultivars that yield more and that are optimized for a range of latitudes and climate zones are also needed.

Miscanthus × *giganteus* is a nothospecies, derived from hybridization between *M. sinensis* (*Msi*) and *M. sacchariflorus* (*Msa*) (Hodkinson *et al.*, 2002*b*). However, little information is available about the genetic diversity and population structure of the parental species of *M*×*g*. An understanding of population structure is essential for enabling selection of parents that maximize heterosis, a strategy that has been especially successful for maize, sorghum and sugarcane (Dillon *et al.*, 2007). Additionally, data on population structure will facilitate association mapping and genomic selection for improved cultivars using molecular marker data.

Miscanthus species *sensu stricto* are native to broad geographies in East Asia and Oceania, with *Msi* the most broadly distributed in Asia (Hodkinson *et al.*, 2002*a*). *Msi* is found from sea level to 2500 m primarily in China, Korea and Japan, but as far north as Sakhalin and as far south as the Indochinese peninsula

(Lee, 1964*a*, *b*; Frontier-Vietnam, 1997; Newman *et al*., 2007; Inthakoun and Delang, 2008; Sun *et al*., 2010). *Msi* is adapted primarily to environments that have an average annual minimum temperature of −28·9 °C or greater (USDA hardiness zone 5 or warmer) and receive ≥750 mm of precipitation annually (Harkevich, 1985; Clifton-Brown *et al*., 2008; Sun *et al*., 2010; Sacks *et al*., 2013). *Msi* is an early colonizer after ecological disturbance in environments that would otherwise support forest (Numata, 1969; Ohtsuka *et al*., 1993; Stewart *et al*., 2009). A broad geographical range suggests opportunities for isolation and differentiation of populations. However, like most *Miscanthus*, *Msi* is self-incompatible and has wind-dispersed pollen and seed, which are traits that would be expected to limit differentiation of populations. The degree of population differentiation has only been partially tested for *Msi* over some of its native range, including studies in South Korea (Qin *et al*., 2013), Japan (Shimono *et al*., 2013), the Izu Islands (Iwata *et al*., 2005, 2006), Taiwan and the Ryukyu Islands (Chou *et al*., 2000), and China (Xu *et al*., 2013; Zhang *et al*., 2013; Zhao *et al*., 2013), and between South Korea and Japan (Slavov *et al*., 2013). *Msi* in Japan were found to be distinct from those in Korea (Slavov *et al*., 2013), and structure was detected among populations in Japan (Shimono *et al*., 2013) and in China (Xu *et al*., 2013; Zhang *et al*., 2013; Zhao *et al*., 2013). Because there was little overlap in geographical sampling or types of molecular marker used among previous studies, it has not been possible to obtain an East Asia-wide perspective of genetic diversity, interpopulation relationships and evolution in *Msi*.

In the USA and Europe, *Msi* is currently a common garden plant, whereas in Asia it is rarely planted, though some wild populations have been managed traditionally for thousands of years by burning or grazing (Stewart *et al*., 2009). Ornamental cultivars of *Msi* were purchased from Japan for import into the USA in the early 1870s (Anonymous, 1876) and became popular in American gardens by the early 1900s (Bailey and Miller, 1901). Since the 1960s, >100 new horticultural cultivars of *Msi* were bred, many by German nurseryman Ernst Pagels (Darke, 1994; Hatch, 2011). Previous molecular studies evaluated genetic diversity among ornamental cultivars of *Msi* (Greef *et al*., 1997; Hodkinson, 2002), but relationships between the ornamental cultivars and wild Asian populations were not explored. In the USA, naturalized populations of *Msi* have become established and, though infrequent, they can be locally abundant, especially in the southern Appalachians (Quinn *et al*., 2010). A recent study compared genetic diversity between native *Msi* populations in Japan and naturalized populations in the USA (Quinn *et al*., 2012), but relationships between naturalized populations, ornamental cultivars and native Asian populations were not determined. Thus, the potential value of naturalized and ornamental *Msi* as a genetic resource for developing improved biomass cultivars of $M \times g$ has been unclear.

In this study, we report on current genetic diversity and population structure of *Msi*, a key genetic resource for the newly important bioenergy crop $M \times g$, in the context of climate change since the last glacial maximum (LGM). In total, we studied 767 *Miscanthus* accessions, including 617 *Msi* from their native range in China, Japan and South Korea, in addition to US-sourced ornamental cultivars and US naturalized populations (Supplementary Data Table S1 and Dataset). To provide an evolutionary context, accessions from all three taxonomic sections (Hodkinson *et al*., 2002*a*) of the genus *Miscanthus* were included: *Triarrhena* (12 *Msa*), *Kariyasua* [1 *M. oligostachyus* (*Mol*)] and *Miscanthus* [703 *Msi* and 4 *M. floridulus* (*Mfl*)]. The accessions were screened with 21 207 single nucleotide polymorphism (SNP) markers identified via the UNEAK pipeline in TASSEL (Bradbury *et al*., 2007) from restriction site-associated DNA sequencing (RAD-seq), 424 GoldenGate SNPs developed previously for two ornamental *Msi* cultivars (Swaminathan *et al*., 2012) and ten plastid simple sequence repeat (SSR) markers (de Cesare *et al*., 2010; Jiang *et al*., 2012). Our objectives were (1) to identify centres of genetic diversity for *Msi* and elucidate its history of expansion and migration in Asia; (2) to determine the origins of ornamental and naturalized *Msi* in the USA; and (3) to understand the degree of divergence and hybridization between *Msi* and other *Miscanthus* species. This is the first study to employ tens of thousands of SNP markers to address these questions, and the first to include such a broad geographical sampling and large number of individuals.

MATERIALS AND METHODS

Plant materials and DNA extraction

Seven hundred and sixty-seven *Miscanthus* accessions were studied, most of which were wild-collected *Msi* genotypes from its native range in China, Korea and Japan, but also ornamental cultivars, biomass cultivars, collections from naturalized populations in the USA and germplasm accessions from the US Department of Agriculture (Supplementary Data Table S1). Genotypes of *Msa*, *Mol* and *Mfl* were included for comparison (Table S1). DNA was extracted via a CTAB (cetyltrimethylammonium bromide) protocol (Kabelka *et al*., 2002). Initial phenotypic identification of the species was based on the systematics of Hodkinson *et al*. (1997).

RAD-seq

Sequencing library preparation was based on the protocol of Poland *et al*. (2012), using 96 barcoded adaptor sequences from Thurber *et al*. (2013). Additional details are provided in the Supplementary Data Methods. All sequence data from this study, totalling 1 359 127 782 reads separated by barcode, have been deposited in the NCBI Sequence Read Archive (BioProject SRP026347). Because *Miscanthus* has undergone a genome duplication with respect to *Sorghum*, its closest relative for which a genome sequence is available (Ma *et al*., 2012), we analysed RAD-seq data with the UNEAK pipeline in TASSEL version 3.0.146, due to its ability to distinguish true SNPs from paralogues in organisms for which a reference genome is unavailable (Lu *et al*., 2013). Three *Msi* doubled haploid lines (Głowacka *et al*., 2012; Swaminathan *et al*., 2012) were included in the marker analyses to confirm genotyping quality (i.e. distinguishing between heterozygous and paralogous alleles). A minimum call rate of 50 % was selected based on improved ability of SNPs to detect variation in *Msi* with respect to a 90 % minimum call rate (Supplementary Data Fig. S1). Sequences of retained SNPs, and SNPs removed for appearing heterozygous in doubled haploid lines, are listed in the Supplementary Dataset.

GoldenGate markers

An oligonucleotide assay pool (OPA) containing 672 SNP assays, including 658 previously mapped SNPs (Swaminathan *et al*., 2012), was designed on the Illumina website. A total of 764 *Msi*, *Mfl* and *Msi* var. *transmorrisonensis* individuals were assayed with this OPA at the Functional Genomics Unit of the Roy J. Carver Biotechnology Center at the University of Illinois. *Msa* and *Mol* individuals were omitted from GoldenGate genotyping due to low transferability of the markers across species. Genotypes were curated manually using Illumina Genome Studio software, and markers that were uninterpretable on one or more plates were discarded. The Illumina IDs and sequences of 424 retained markers and 248 discarded markers are listed in the Supplementary Dataset. Although RAD-seq SNPs were used in most analyses due to expectation of ascertainment bias in GoldenGate SNPs, GoldenGate SNPs were used for calculation of F_{IS} due to heterozygote undercalling in the RAD-seq SNPs.

Plastid markers

Ten previously published plastid microsatellite markers (de Cesare *et al*., 2010; Jiang *et al*., 2012; Supplementary Data Methods) were used to analyse diversity and distribution of maternally inherited genes. Size separation of the PCR products by capillary electrophoresis was done on a 3730xl DNA Analyzer (Applied Biosystems, Foster City, CA, USA) with GeneScan 500 LIZ size standard. Marker scoring was accomplished with STRand software v. 2.4.59 (http://www.vgl.ucdavis.edu/STRand; Toonen and Hughes, 2001).

Data analysis

SNP data were analysed using the software Structure (Falush *et al*., 2003), Structure Harvester (Earl and vonHoldt, 2011), adegenet (Jombart *et al*., 2010; Jombart and Ahmed, 2011), TASSEL (Bradbury *et al*., 2007), ape (Paradis *et al*., 2004), mmod (Winter, 2012), pegas (Paradis, 2010), GenAlEx (Peakall and Smouse, 2006), TreeMix (Pickrell and Pritchard, 2012) and an R script implementing an algorithm from PowerMarker (Liu and Muse, 2005). Plastid data were analysed using the software polysat (Clark and Jasieniuk, 2011) and source code from pegas (Paradis, 2010). Additional details regarding data analysis are provided in the Supplementary Data Methods.

RESULTS

Population structure in Asia

Seven genetic clusters (K = 7) based on nuclear SNP data were identified, including six clusters of *Msi* from geographically distinct regions in Asia and one cluster for *Msa* (Fig. 1A–C). Four of the *Msi* clusters were from mainland Asia (south-eastern China, Yangtze-Qinling, Sichuan Basin, Korea/North China), and two were from Japan (southern and northern; Fig. 1A). The software Structure (Falush *et al*., 2003; Evanno *et al*., 2005) and the discriminant analysis of principal components (DAPC) algorithm in the R package adegenet (Jombart *et al*., 2010) both indicated K = 7 across the entire data set, or K = 6 when *Msa* was omitted (Supplementary Data Fig. S2). The *Mol* individual was omitted from these analyses since both algorithms would be expected to have difficulty finding correct clusters when only a single individual belonged to a group. The DAPC analysis produced similar results for RAD-seq and GoldenGate SNPs (Fig. 1B), so the clusters identified using RAD-seq data were chosen for all subsequent analysis due to expectation of ascertainment bias in GoldenGate SNPs. Cluster assignments for all individuals are provided in the Supplementary Dataset. Sub-structure within the south-eastern China and Yangtze-Qinling populations was indicated by their division into sub-regions with different patterns of admixture (Fig. 1A, B), and by significantly higher F_{IS} values (Table 1).

The south-eastern China and the Korea/North China clusters had the greatest nuclear gene diversity among the *Msi* populations, based on estimates of Nei's *D* from RAD-seq SNPs (Table 1). The northern Japan and the Sichuan Basin clusters were the least diverse based on Nei's *D* (Table 1) and were the most differentiated from the other clusters as indicated by relatively high F_{ST} (Table 1) and high pairwise estimates of Jost's *D* (Fig. 2). Low diversity and high differentiation are consistent with northern Japan and the Sichuan Basin being the most geographically isolated clusters.

All accessions were treated as diploid for the purpose of SNP analysis. Based on the RAD-seq SNPs, the average observed heterozygosity per individual was 0·0811 with a standard deviation of 0·0194. The known triploid *M. sinensis* 'Goliath' (Chae *et al*., 2014) had an observed heterozygosity of 0·1680, higher than that of all other accessions, and for this reason was excluded from SNP analysis (Supplementary Dataset).

Strong geographical differentiation was also seen among plastid haplotypes (Fig. 3). Patterns of differentiation (Jost's *D*) were similar for nuclear SNPs and plastid haplotypes, although the degree of differentiation was higher in the plastid genome than in the nuclear genome (Fig. 2; Supplementary Data Table S3). However, the groups with the highest diversity in terms of nuclear SNP data were not necessarily the most diverse in terms of plastid data (Table 1). The most common haplotype in southern China was at the centre of the haplotype network, whereas the most common haplotype north of the Yangtze River was diverged from it by several mutational steps (Fig. 3D, H). Korea and Japan were both split across two divergent regions of the haplotype network (Fig. 3).

Additionally, Neighbor–Joining analysis of the RAD-seq SNP data via TASSEL 3.0.146 (Bradbury *et al*., 2007) identified major groups that were consistent with the DAPC and Structure analyses, and provided additional resolution of population substructure (Fig. 1D). As expected, *Msa* and *Mol* were outgroups to *Msi* in both the Neighbor–Joining analysis (Fig. 1E) and the SSR-derived plastid haplotype network (Fig. 3), providing a useful root to facilitate interpretation. *Mfl* clustered with the earliest branching group of *Msi* (coastal south-eastern China, including Okinawa and Taiwan; Fig. 1D, E) and had plastid haplotypes common in *Msi* (Fig. 3). Given that the plastid haplotype network was centred on southern China (Fig. 3D), and the south-eastern China DAPC cluster was at the centre of the minimum spanning tree between clusters (Fig. 1C), south-eastern China was used as the outgroup in TreeMix (Pickrell and Pritchard, 2012) for further elucidating the divergence and subsequent migration between genetic groups (Fig. 4A).

Eleven accessions from China were found to be interspecific hybrids between *Msa* and *Msi*, as indicated by admixture in

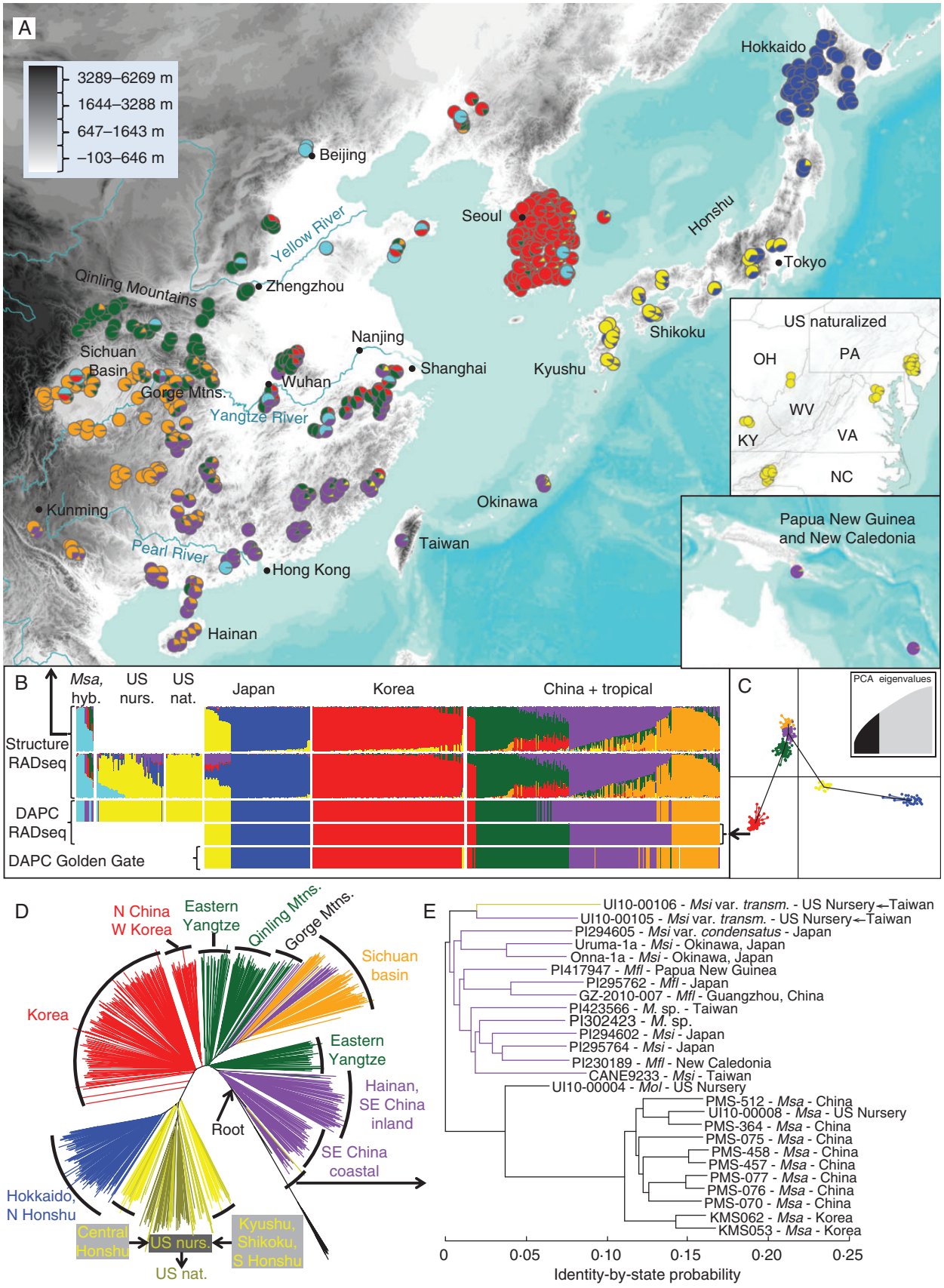
A
3289–6269 m
1644–3288 m
647–1643 m
–103–646 m
Beijing
Yellow River
Seoul
Hokkaido
Honshu
Tokyo
Qinling Mountains
Zhengzhou
Nanjing
Shikoku
Sichuan Basin
Shanghai
Kyushu
Wuhan
Gorge Mtns.
Yangtze River
US naturalized
OH
PA
WV
KY
VA
NC
Okinawa
Kunming
Taiwan
Pearl River
Hong Kong
Hainan
Papua New Guinea and New Caledonia
B
Msa, hyb.
US nurs.
US nat.
Japan
Korea
China + tropical
Structure RADseq
DAPC RADseq
DAPC Golden Gate
C
PCA eigenvalues
D
N China W Korea
Eastern Yangtze
Qinling Mtns.
Gorge Mtns.
Sichuan basin
Korea
Eastern Yangtze
Hainan, SE China inland
Root
SE China coastal
Hokkaido, N Honshu
Central Honshu
US nurs.
Kyushu, Shikoku, S Honshu
US nat.
E
UI10-00106 - Msi var. transm. - US Nursery←Taiwan
UI10-00105 - Msi var. transm. - US Nursery←Taiwan
PI294605 - Msi var. condensatus - Japan
Uruma-1a - Msi - Okinawa, Japan
Onna-1a - Msi - Okinawa, Japan
PI417947 - Mfl - Papua New Guinea
PI295762 - Mfl - Japan
GZ-2010-007 - Mfl - Guangzhou, China
PI423566 - M. sp. - Taiwan
PI302423 - M. sp.
PI294602 - Msi - Japan
PI295764 - Msi - Japan
PI230189 - Mfl - New Caledonia
CANE9233 - Msi - Taiwan
UI10-00004 - Mol - US Nursery
PMS-512 - Msa - China
UI10-00008 - Msa - US Nursery
PMS-364 - Msa - China
PMS-075 - Msa - China
PMS-458 - Msa - China
PMS-457 - Msa - China
PMS-077 - Msa - China
PMS-076 - Msa - China
PMS-070 - Msa - China
KMS062 - Msa - Korea
KMS053 - Msa - Korea
0
0·05
0·10
0·15
0·20
0·25
Identity-by-state probability

TABLE 1. *Diversity and inbreeding of* M. sinensis *clusters*

	N	N_{cp}	GS_{cp}	*D*	F_{IS}	F_{ST}
South-eastern China plus tropical	125	22	0·75 ± 0·04	0·131 ± 0·001	0·1352 ± 0·0005	0·0177 ± 0·0002
Yangtze-Qinling	114	18	0·57 ± 0·06	0·125 ± 0·001	0·1096 ± 0·0006	0·0160 ± 0·0002
Sichuan Basin	58	9	0·74 ± 0·04	0·120 ± 0·001	0·0708 ± 0·0007	0·0295 ± 0·0005
Korea/North China	195	12	0·48 ± 0·04	0·130 ± 0·001	0·0546 ± 0·0004	0·0288 ± 0·0003
Northern Japan	96	6	0·25 ± 0·06	0·113 ± 0·001	0·0636 ± 0·0005	0·0374 ± 0·0005
Southern Japan	32	11	0·87 ± 0·04	0·123 ± 0·001	0·0612 ± 0·0007	0·0079 ± 0·0001
US nurseries, hybrid	34	5	0·28 ± 0·10	0·138 ± 0·001	−0·1356 ± 0·0006	
US nurseries, non-hybrid	42	10	0·80 ± 0·04	0·128 ± 0·001	0·0188 ± 0·0006	
US naturalized	43	3	0·55 ± 0·07	0·119 ± 0·001	0·1139 ± 0·0006	

Groupings in the native range are based on DAPC of RAD-seq data (Fig. 1C). Individuals from US nurseries are split into hybrid and non-hybrid groups based on Structure analysis (Fig. 1B).

N, number of individuals; N_{cp}, number of plastid haplotypes; GS_{cp}, Gini–Simpson index, calculated from plastid haplotype frequencies; *D*, diversity, as calculated from expected heterozygosity of RAD-seq SNPs; F_{IS}, inbreeding coefficient calculated from expected and observed heterozygosity of GoldenGate SNPs; F_{ST}, differentiation from the other five native clusters, estimated from allele frequencies of RAD-seq SNPs using adegenet (Jombart and Ahmed, 2011). Standard errors are given.

Fig. 1A and B. Because plastids are inherited maternally in grasses (Corriveau and Coleman, 1988), we were also able to determine that for nine of the hybrids *Msa* was the maternal parent. With one exception, the hybrids appeared to be F_1 hybrids based on approximate 1:1 ratios of genetic composition determined by Structure (Fig. 1B, Supplementary Dataset). Of the non-hybrid individuals, we did not observe any that had an *Msa* plastid haplotype in combination with an *Msi* nuclear genome, or vice versa (Fig. 3; Supplementary Dataset).

Where did US ornamental cultivars and naturalized Msi *originate?*

Southern Japan was the origin of nearly all ornamental *Msi* cultivars in the USA, as indicated by cluster analysis (Fig. 1B), phylogenetic analysis (Fig. 1D), Jost's *D* statistic (Fig. 2) and similarity of plastid haplotypes (Fig. 3; Supplementary Data Table S3). The two most common *Msi* plastid haplotypes in the USA were unique to southern Japan (Fig. 3A, B) and, among rarer haplotypes found in the USA, none was unique to other regions of Asia (Fig. 3). A single exception to the southern Japan origin of *Msi* ornamental cultivars was *Msi* var. *transmorrisonensis*, which was part of the south-eastern China cluster (Fig. 1), supporting historical documentation that it was introduced from Taiwan (Darke, 1994). Unexpectedly, admixture (Fig. 1B) and principal component analyses (Supplementary Data Fig. S3) revealed that nearly half of the ornamental cultivars labelled as *Msi* were in fact BC_1 or BC_2 hybrids of *Msi* crossed with either *Msa* or *Mol*, and that *Msi* was the female recurrent parent species (Figs 1B and 3). All hybrid ornamentals were also derived from southern Japan with regards to the *Msi* component of their ancestry, with the exception of *Msi* × *Msa* var. *purpurascens* which was derived from the Korea/North China group (Supplementary Dataset).

Naturalized *Msi* populations in the USA were derived from ornamental *Msi* (Figs 1 and 3), but did not have any ancestry from *Msa* or *Mol* (Fig. 1B). Naturalized populations were closely related to non-hybrid cultivars, as indicated by Jost's *D* (Fig. 2). The high F_{IS} value of naturalized accessions (Table 1) indicated population sub-structure within this group.

DISCUSSION

Evolution of Msi *since the LGM*

In this study, nuclear and plastid data indicated that south-eastern China and nearby islands were a centre of radiation for *Msi*, with multiple paths of migration northward into eastern Asia. To understand both the sequence and the approximate timing of divergence and migration for *Msi* populations in Asia, we combined information from our genetic analyses with a series of previously published vegetation maps from the LGM to the present, which were based primarily on analyses of palaeopollen in conjunction with ^{14}C dating (Winkler and Wang, 1993; Adams and Faure, 1997; Ray and Adams, 2001). During the LGM, when the climate in Asia was colder and drier than at present, the distribution of *Msi* in Asia was probably limited to a refugium in coastal areas from southern Taiwan to Hainan that are now mostly submerged under the South China Sea, and on the Indochinese peninsula, because these areas had sufficiently

FIG. 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 USA, which were coloured based on the second bar chart from the top, which contains an additional 120 individuals from the USA 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 coloured based on DAPC groupings as in (C), with the exception of individuals from the USA. (E) Sub-set of Neighbor–Joining tree from (D), showing groups adjacent to the root of the tree.

	SE China plus tropical	Yangtze - Qinling	Sichuan basin	Korea, N China	N Japan	S Japan	US nurseries, hybrid	US nurseries, non-hybrid
Yangtze-Qinling	0·0295 ± 0·0006							
Sichuan Basin	0·0264 ± 0·0006	0·0590 ± 0·0010						
Korea, N China	0·0359 ± 0·0006	0·0187 ± 0·0004	0·0666 ± 0·0011					
N Japan	0·0483 ± 0·0009	0·0582 ± 0·0010	0·0763 ± 0·0013	0·0549 ± 0·0010				
S Japan	0·0298 ± 0·0006	0·0396 ± 0·0007	0·0593 ± 0·0010	0·0348 ± 0·0007	0·0234 ± 0·0006			
US nurseries, hybrid	0·0452 ± 0·0007	0·0525 ± 0·0008	0·0730 ± 0·0011	0·0475 ± 0·0008	0·0437 ± 0·0008	0·0281 ± 0·0006		
US nurseries, non-hybrid	0·0341 ± 0·0006	0·0452 ± 0·0008	0·0630 ± 0·0011	0·0411 ± 0·0007	0·0303 ± 0·0007	0·0103 ± 0·0003	0·0194 ± 0·0005	
US naturalized	0·0431 ± 0·001	0·0518 ± 0·0011	0·0717 ± 0·0014	0·0472 ± 0·0011	0·0389 ± 0·0011	0·0186 ± 0·0008	0·0214 ± 0·0008	0·0050 ± 0·0007

FIG. 2. Jost's *D* statistic showing differentiation between *M. sinensis* genetic groups based on RAD-seq markers. Mean and standard error values are shown across 21 207 loci. Groups are the same as those in Table 1. A colour scale is used to highlight similarities between groups, with green being most similar and red being most dissimilar.

warm and moist environments to support tropical grasslands and forest, whereas the rest of eastern Asia did not (Fig. 4B; Ray and Adams, 2001). For example, during the LGM, southern Japan and the Sichuan Basin harboured open boreal woodlands (Ray and Adams, 2001) comparable with the Siberian taiga today, indicating that the climate there was too cold to be habitable for *Msi*. Near present-day pine (*Pinus*) and birch (*Betula*) forests in the Ussuri and Amur River watersheds of eastern Russia, *Msi* is absent (E. Sacks, pers. obs., 2012 USDA-ARS and N. I. Vavilov Research Institute collection expedition). Similarly, during the LGM, pollen and phytolith data from the Mt. Aso caldera on Kyushu in southern Japan indicated the presence of pine and birch forests and an absence of *Msi* (Hase *et al.*, 2012; Miyabuchi *et al.*, 2012). Consistent with the predicted refugium based on climate during the LGM, *Msi* from south-eastern China was the central branching point of the minimum spanning tree of DAPC clusters (Fig. 1C), and the closest group to the root of the phylogenetic tree (Fig. 1D), indicating that *Msi* radiated from this region into the rest of Asia. Additionally, the south-eastern China population had the greatest number of plastid haplotypes (Table 1), and the most common haplotype in south-eastern China was identified as the most ancestral of the species, given its central location in the haplotype network and its connection to the *Msa* section of the network (Fig. 3).

After the LGM, warming in Japan preceded warming on mainland Asia, such that by approx. 14 000 years before present (ybp) nearly all of Honshu had a temperate climate favourable for *Msi*, whereas suitable climate on the mainland was still restricted to south of the Yangtze River (Winkler and Wang, 1993; Adams, 1997; Fig. 4B). Consistent with the early warming of Japan, the Japanese *Msi* populations were the most distal on the phylogenetic tree, indicating an early migration, presumably of wind-blown seed, via what was then a narrow sea gap from the ancestral population in south-eastern China (Fig. 1D). Japan may have cooled again during the Younger Dryas 12 800–11 500 ybp (Adams, 1997), which would have caused *Msi* to retreat southward or onto small islands with a warmer maritime microclimate, but nevertheless our data indicate a stronger genetic connection between south-eastern China and Japan than between Korea and Japan (Figs 1C and 2). Moreover, our indirect analyses based on general vegetation maps and molecular genetics are supported by direct evidence from Mt. Aso, where phytolith data indicate that the currently extensive colonization of the caldera by *Msi* is part of a continuous occupation that started approx. 13 500 ybp (Miyabuchi *et al.*, 2012). Genetic isolation of the northern Japan population from the mainland contrasts with evidence for gene exchange between the southern Japan and Korea/North China populations (Figs 1A and 2; Table 1), indicating that the northern Japan population became established in northern Honshu and Hokkaido prior to the arrival of *Msi* in northern China and Korea. Plastid data further support an early split between the south-eastern China population and the Japanese populations, as the haplotype that was most common in northern

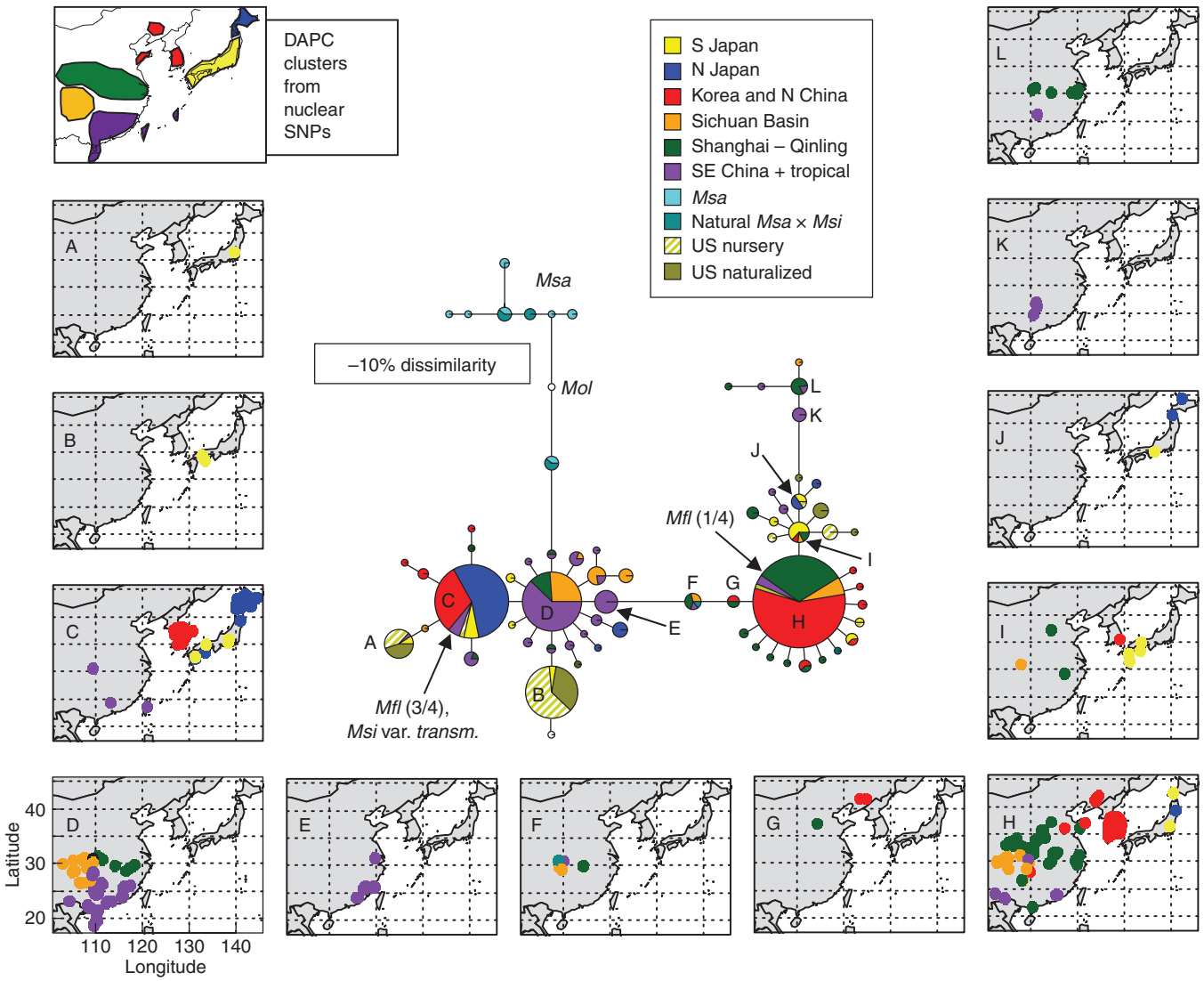

FIG. 3. Plastid haplotype network for 763 *Miscanthus* individuals using ten microsatellite markers. The circle area is proportional to the number of individuals with a given haplotype. Colours from Fig. 1 are used to depict group membership, with the six native *M. sinensis* groups being those from Fig. 1C. The geographical distribution of the six native *M. sinensis* groups, as determined by nuclear SNPs, is shown in the upper left inset. Each circle represents one haplotype. Lengths of lines between circles are proportional to the dissimilarity between haplotypes. Geographical distributions of 12 selected haplotypes are shown in maps A–L.

Japan was related by one mutational step to the most common haplotype in southern China (Fig. 3). The relatively low nuclear genetic diversity of both Japanese populations is consistent with an initial migration of a relatively small founder population from south-eastern China to Japan followed by rapid expansion (Fig. 4B, C; Table 1). Additionally, migration between south-eastern China and southern Japan remains possible in modern times via the Ryukyu Islands.

As a warmer and moister climate rapidly returned to mainland Asia after approx. 10 000 ybp, primary migrations of *Msi* occurred from south-eastern China to the north of the Yangtze River and west to the Sichuan Basin, and north to north-eastern China and Korea (Figs 1 and 4). Analyses of nuclear SNP data with the software TreeMix (Pickrell and Pritchard, 2012) indicate that the Korea/North China and Yangtze-Qinling populations radiated from south-eastern China separately from the populations in Japan and the Sichuan Basin (Fig. 4A). As indicated by the phylogenetic tree, the eastern Yangtze River region was a secondary centre of mainland radiation during this period (Fig. 1D). Admixture between the Sichuan Basin population and portions of the south-eastern China population indicates secondary north–south migrations between these populations (Fig. 1A). Additionally, the plastid haplotype most common in the Korea/North China and Yangtze-Qinling populations is separated by several mutational steps from the haplotype most common in south-eastern China, reflecting the later date of this radiation with respect to the earlier primary migration from south-eastern China to Japan (Fig. 3). Several haplotypes derived from this common haplotype north of the Yangtze are also found in Japan, indicating a secondary migration to Japan from Korea (Figs 3 and 4C).

Analysis of molecular variation (AMOVA; Excoffier *et al.*, 1992) indicated that 69 % of the genetic variation for *Msi* was within clusters, and that 40 % of between-cluster variation was

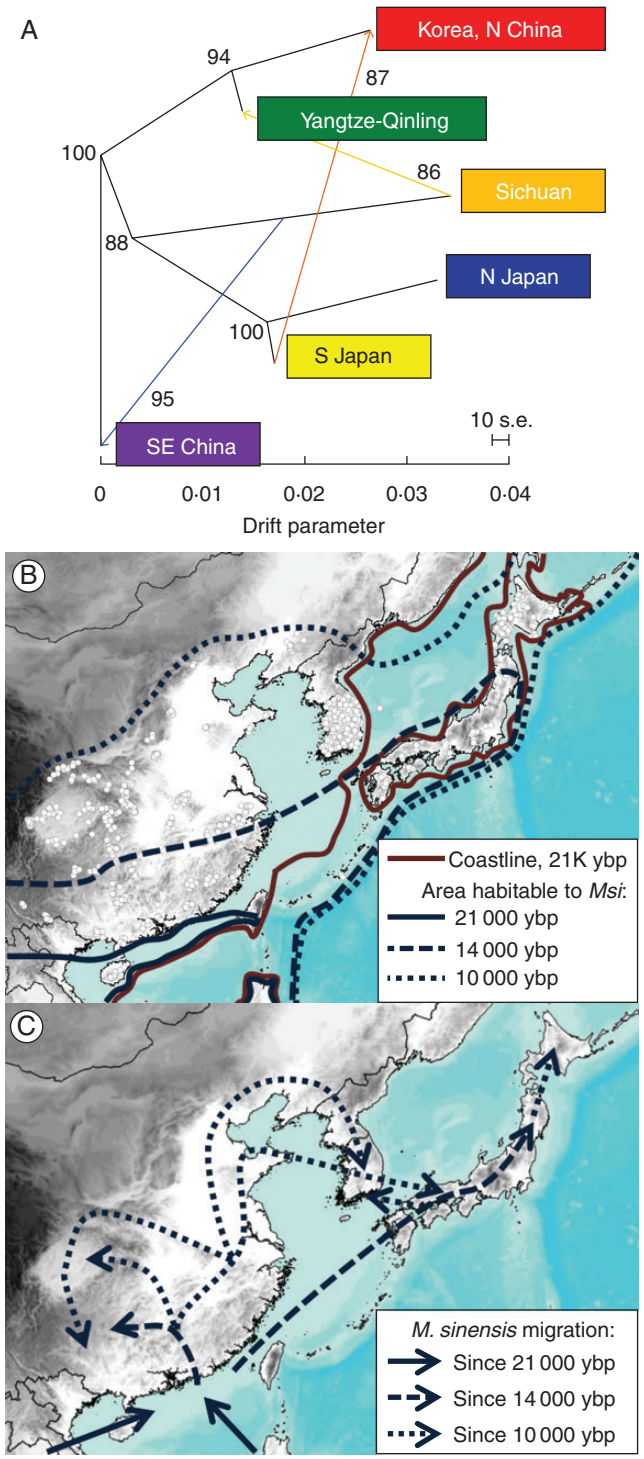

FIG. 4. Model of radiation of *M. sinensis* into Asia. (A) TreeMix results using combined RAD-seq and GoldenGate data sets. The black dendrogram depicts divergence of populations, while the coloured edges (arrows) depict subsequent migrations. Genetic groups are those determined by DAPC of RAD-seq data from the 620 non-hybrid *M. sinensis* collections in the native range. SE China was set as the outgroup, and the number of migration edges was chosen to be three based on the highest bootstrap values. The blue migration edge has the highest weight, and the yellow edge the lowest. One hundred bootstrap replicates were run with 500 SNPs each. Bootstrap values are provided for all nodes and migration edges. (B) Exposed landmass and regions habitable to *Miscanthus* at time points since the last glacial maximum. Outlines of habitable regions are based on published vegetation maps (Winkler and Wang, 1993; Adams, 1997; Adams and Faure, 1997; Ray and Adams, 2001) with dates adjusted according to IntCal04 (Reimer *et al.*, 2004), assuming that *Miscanthus* would grow in temperate, sub-tropical and tropical forests, and tropical grassland, but not in cold coniferous forests, arid steppe, forest-grassland or temperate grassland. White circles indicate sampling locations in this study. (C) *M. sinensis* migrations based on plastid microsatellites (Fig. 3) and nuclear SNP markers (A, Fig. 1D) with timings based on (B).

between mainland Asia and Japanese populations (Supplementary Data Table S2). For outcrossing plant species, genetic variation is typically greater within than among populations. For maize, variation among clusters and sub-clusters accounted for only 10·7 % of genetic variation (Vigouroux *et al.*, 2008), and for switchgrass, variation among genetic pools accounted for 22 % of the total (Zhang *et al.*, 2011). The relatively high proportion of variation (31 %) among clusters for *Msi* may be attributed to having not been subjected to bottlenecks and migrations associated with domestication, as was maize, and to greater isolation by seas (e.g. for Japan and Korea) and mountains (e.g. for the Sichuan Basin) relative to switchgrass (Fig. 1A).

The four *Mfl* individuals in this study, which were from mainland China, Japan, New Caledonia and Papua New Guinea, clustered with the south-eastern China population of *Msi*, in particular the sub-population from Okinawa and Taiwan (Fig. 1D, E), calling into question the validity of having two different species names. Additionally, all plastid haplotypes observed in *Mfl* were also common in *Msi* (Fig. 3). These results were consistent with another molecular phylogenetic study, in which one *Mfl* genotype was embedded in *Msi* genotypes (Hodkinson *et al.*, 2002*a*). In contrast to the more northerly distribution of *Msi*, *Mfl* has a sub-tropical to tropical distribution from southern China throughout the Pacific islands to New Guinea (Clifton-Brown *et al.*, 2008; Sacks *et al.*, 2013). Phenotypic intergradation between *Msi* and *Mfl* where the two species overlap geographically has been a long-standing taxonomic problem (Hooker, 1893; Sun *et al.*, 2010). Though sampling of *Mfl* in the current and previous studies (Hodkinson *et al.*, 2002) was too limited to reach firm conclusions about the relationship of this taxon to *Msi*, the results suggest that modern molecular tools applied to an in-depth study of *Mfl* diversity and population structure in relation to *Msi* would be valuable.

This close relationship between *Mfl* and the earliest branching, sub-tropical *Msi* group (Fig. 1D, E) raises the intriguing question of whether *Msi* and *Mfl* originated in south-east Asia then migrated to the Southern Hemisphere, or if they originated in Australasia and migrated northward. Notably, New Guinea is currently a centre of diversity for the closest relative of *Miscanthus*, *Saccharum* (Grivet *et al.*, 2006), including the domesticated *S. officinarum* and its wild progenitor *S. robustum* (Paterson *et al.*, 2013). However, in support of the Asian origin hypothesis is the fact that species diversity in *Miscanthus* is currently greatest between 30 and 40°N in Asia (Hodkinson *et al.*, 2002*a*). Given the potential for genetic bottlenecks associated with migration over multiple islands, and the large distance between mainland south-eastern Asia and New Guinea, we would expect the *Msi* and *Mfl* populations from the centre of origin to have greater allelic diversity than derived populations. Additional sampling of populations in the Southern Hemisphere would provide a means of testing these competing hypotheses of origin.

Origins of ornamental and naturalized Msi *in the USA*

To the best of our knowledge, this is the first study to report that many of the ornamental *Miscanthus* cultivars sold as *Msi* are actually interspecific backcross hybrids, with *Msi* used as the recurrent female parent. These interspecific hybridizations were probably an effort by plant breeders in Europe and the USA to introgress genes for earlier flowering and greater winter hardiness into an *Msi* population that was otherwise adapted to a mild climate (USDA hardiness zones 8 and 9), and photoperiods associated with approx. 30–35°N. As expected for wide crosses, nuclear genetic diversity was relatively high among the interspecific hybrids, but the low plastid diversity indicated that few *Msi* parents were used (Table 1). The naturalized populations, however, were only derived from *Msi* and not the interspecific backcross cultivars, suggesting either a lack of fitness of interspecific hybrids, or insufficient time for establishment of naturalized populations since the release of backcross interspecific hybrid cultivars to the nursery trade. Low nuclear and plastid genetic diversity with respect to non-hybrid cultivars (Table 1) suggests that the naturalized populations were derived from a small set of founder ornamental genotypes. Different populations of naturalized *Msi* were derived from different sets of ornamental cultivar parents, as indicated by the high F_{IS} value of naturalized accessions (Table 1).

Msi *germplasm resources for crop improvement*

In our set of 617 *Msi* individuals from the native range, 90 % of polymorphic RAD-seq loci (95 % of RAD-seq alleles) could be captured with a relatively small core set of 55 individuals selected to maximize the number of included SNP alleles (Supplementary Data Fig. S4 and Dataset). Self-incompatibility, wind dispersal and the relatively short time for the development of population structure since the LGM all probably contributed to the small number of individuals required to obtain a highly representative core population. In contrast to the wild germplasm, the sets of 43 naturalized individuals and 76 ornamental cultivars from the USA captured only 46 and 68 % of the polymorphism from the native range, respectively (Supplementary Data Fig. S4). The genetic bottleneck associated with US *Msi* germplasm was a previously unknown limitation to breeding improved bioenergy feedstock cultivars of *Miscanthus*. However, our core germplasm set, or one with a similar distribution across genetic clusters and geographical area, would provide an excellent starting point for breeding new $M \times g$ biomass cultivars. Additionally, individual core populations for each of the identified clusters (Supplementary Dataset) could be established to facilitate population improvement in *Msi* while maintaining potential heterotic groups. The individuals in the core populations represent adaptation to a wide range of latitudes and climates, and probably possess desirable alleles for yield traits, pest and pathogen resistance, and abiotic stress tolerance.

Conclusions

Whereas maize, sorghum and sugarcane have population structures dominated by a history of anthropogenic movement in recent millennia (Van Heerwaarden *et al.*, 2011; Morris *et al.*, 2013; Paterson *et al.*, 2013), we found that the population structure of *Msi* appeared driven primarily by patterns of warming since the last glaciation, and secondarily by geographical barriers such as seas and mountains. Knowledge of *Msi* population structure provided by this study will facilitate germplasm conservation, association mapping and genomic selection. Moreover, discovery of distinct *Msi* populations is a first step towards identifying potential heterotic groups. Currently, a few triploid $M \times g$ genotypes have been bred at the University of Illinois but not yet released, and the ancestry of each is entirely from southern Japan. With the identification of six main *Msi* populations in eastern Asia, it may be possible to breed $M \times g$ with greater heterosis and yield by crossing tetraploid *Msa* with *Msi* from populations other than southern Japan. $M \times g$ has proved a remarkably productive and sustainable bioenergy feedstock in Europe and the USA, its high yields minimizing the land area needed for large-scale bioenergy production. However, the single genotype that has been used and trialled at scale, to date, is poorly suited to colder and warmer climates in Europe and the USA. Breeding improved cultivars of perennial bioenergy crops, such as *Miscanthus*, will contribute towards mitigation of anthropogenic climate change, possibly humanity's greatest present challenge.

SUPPLEMENTARY DATA

Supplementary data are available online at www.aob.oxfordjournals.org and consist of the following. Dataset: information on individuals and markers used in the study. Table S1: *Miscanthus* collections used in the present study. Table S2: analysis of molecular variance (AMOVA) of 620 *M. sinensis* and *M. floridulus* individuals sampled in the native range. Table S3: pairwise Jost's *D* statistics showing genetic differentiation between *Msi* groups in terms of plastid haplotype frequencies. Fig. S1: choice of minimum call rate for RAD-seq markers generated from the UNEAK pipeline. Fig. S2: selection of number of clusters for structure and DAPC analysis. Fig. S3: principal component analysis demonstrating hybrid ancestry of *Msi* accessions from US nurseries. Fig. S4: proportion of polymorphic loci captured in core sets vs. number of individuals in the set. Methods: details for DNA extraction, sequencing library preparation, UNEAK pipeline, SNP analysis and plastid markers.

ACKNOWLEDGEMENTS

L.V.C. and E.J.S. are affiliated with the Department of Crop Sciences and the Energy Biosciences Institute. We dedicate this paper to the memory of Dean Tiessen, who as president of New Energy Farms helped support this study and the development of the bioenergy industry. We thank Pat Brown, Jessica Petersen, handling editor Michael Fay, and two anonymous reviewers for their valuable feedback on earlier versions of this manuscript. Kankshita Swaminathan gave helpful advice on SNP genotyping. Pat Brown provided *Pst*I adaptors for RAD-seq. Megan Swanson, Kasia Dubiel and Melina Salgado assisted with DNA extraction and plastid genotyping at the University of Illinois. Christine Fleener and Lauren Quinn collected naturalized *Msi* in the USA. This research was supported by the Office of Science (BER), US Department of Energy [Project ID 0017582].

LITERATURE CITED

Adams J. 1997. Global land environments since the last interglacial. http://www.esd.ornl.gov/ern/qen/nerc.html. Last accessed 5 Nov 2013.

Adams JM, Faure H. 1997. Preliminary vegetation maps of the World since the Last Glacial Maximum: an aid to archaeological understanding. *Journal of Archaeological Science* **24**: 623–647.

Anonymous. 1876. The zebra-striped Eulalia. *The American Agriculturalist* **3**: 460.

Bailey LH, Miller W. 1901. *Cyclopedia of American horticulture*. New York: The MacMillan Co.

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**: 2633–2635.

de Cesare M, Hodkinson TR, Barth S. 2010. Chloroplast DNA markers (cpSSRs, SNPs) for *Miscanthus*, *Saccharum* and related grasses (Panicoideae, Poaceae). *Molecular Breeding* **26**: 539–544.

Chae WB, Hong SJ, Gifford JM, Rayburn AL, Sacks EJ, Juvik JA. 2014. Plant morphology, genome size, and SSR markers differentiate five distinct taxonomic groups among accessions in the genus *Miscanthus*. *GCB Bioenergy* doi:10.1111/gcbb.12101 (in press).

Chou C-H, Chiang Y-C, Chiang T-Y. 2000. Genetic variability and phytogeography of *Miscanthus sinensis* var. *condensatus*, an apomictic grass, based on RAPD fingerprints. *Canadian Journal of Botany* **78**: 1262–1268.

Clark L V, Jasieniuk M. 2011. POLYSAT: an R package for polyploid microsatellite analysis. *Molecular Ecology Resources* **11**: 562–566.

Clifton-Brown J, Chiang Y-C, Hodkinson TR. 2008. *Miscanthus*: genetic resources and breeding potential to enhance bioenergy production. In: Vermerris W, ed. *Genetic improvement of bioenergy crops*. Berlin: Springer, 273–294.

Corriveau JL, Coleman AW. 1988. Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 Angiosperm species. *American Journal of Botany* **75**: 1443–1458.

Darke R. 1994. A century of grasses. *Arnoldia* **54**: 3–11.

Dillon SL, Shapter FM, Henry RJ, Cordeiro G, Izquierdo L, Lee LS. 2007. Domestication to crop improvement: genetic resources for *Sorghum* and *Saccharum* (Andropogoneae). *Annals of Botany* **100**: 975–989.

Earl DA, vonHoldt BM. 2011. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**: 359–361.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2611–2620.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**: 479–491.

Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**: 1567–1587.

Frontier-Vietnam. 1997. *Na Hang Nature Reserve, Tat Ke Sector. Site description and conservation evaluation*. London: Society for Environmental Exploration.

Głowacka K, Kaczmarek Z, Jeżowski S. 2012. Androgenesis in the bioenergy plant *Miscanthus sinensis*: from calli induction to plant regeneration. *Crop Science* **52**: 2659–2673.

Głowacka K, Clark LV, Adhikari S, *et al*. 2014. Genetic variation in *Miscanthus* × *giganteus* and the importance of estimating genetic distance thresholds for differentiating clones. *GCB Bioenergy* (in press).

Greef J, Deuter M, Jung C, Schondelmaier J. 1997. Genetic diversity of European *Miscanthus* species revealed by AFLP fingerprinting. *Genetic Resources and Crop Evolution* **44**: 185–195.

Grivet L, Glaszmann J-C, D'Hont A. 2006. Molecular evidences for sugarcane evolution and domestication. In: Motley T, Zerega N, Cross H, eds. *Darwin's harvest. New approaches to the origins, evolution, and conservation of crops*. New York: Columbia University Press, 49–66.

Harkevich S. 1985. *The vascular plants of Soviet Far East*. Nauka Publisher.

Hase Y, Iwauchi A, Uchikoshiyama U, Noguchi E, Sasaki N. 2012. Vegetation changes after the late period of the Last Glacial Age based on pollen analysis of the northern area of Aso Caldera in central Kyushu, Southwest Japan. *Quaternary International* **254**: 107–117.

Hatch L. 2011. *Hatch's perennials 1.1, cultivars of hardy herbaceous plants*. TCR Press.

Hodkinson TR. 2002. Characterization of a genetic resource collection for *Miscanthus* (Saccharinae, Andropogoneae, Poaceae) using AFLP and ISSR PCR. *Annals of Botany* **89**: 627–636.

Hodkinson TR, Renvoize SA, Chase MW. 1997. Systematics of *Miscanthus*. *Aspects of Applied Biology* **49**: 189–198.

Hodkinson TR, Chase MW, Lledó MD, Salamin N, Renvoise SA. 2002*a*. Phylogenetics of *Miscanthus, Saccharum* and related genera (Saccharinae, Andropogoneae, Poaceae) based on DNA sequences from ITS nuclear ribosomal DNA and plastid *trnL* intron and *trnL-F* intergenic spacers. *Journal of Plant Research* **115**: 381–392.

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

Hooker JD. 1893. Tab. 7304. *Miscanthus sinensis. Curtis's Botanical Magazine* No. 583, vol. **119**, July. Available online at http://www.biodiversitylibrary.org/item/14247.

Inthakoun L, Delang CO. 2008. *Lao Flora: a checklist of plants found in Lao PDR with scientific and vernacular names*. Morrisville, NC: Lulu Press.

Iwata H, Kamijo T, Tsumura Y. 2005. Genetic structure of *Miscanthus sinensis* ssp. *condensatus* (Poaceae) on Miyake Island: implications for revegetation of volcanically devastated sites. *Ecological Research* **20**: 233–238.

Iwata H, Kamijo T, Tsumura Y. 2006. Assessment of genetic diversity of native species in Izu Islands for a discriminate choice of source populations: implications for revegetation of volcanically devastated sites. *Conservation Genetics* **7**: 399–413.

Jiang J-X, Wang Z-H, Tang B-R, Xiao L, Ai X, Yi Z-L. 2012. Development of novel chloroplast microsatellite markers for *Miscanthus* species (Poaceae). *American Journal of Botany* **99**: e230–e233.

Jombart T, Ahmed I. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**: 3070–3071.

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.

Kabelka E, Franchino B, Francis DM. 2002. Two loci from *Lycopersicon hirsutum* LA407 confer resistance to strains of *Clavibacter michiganensis* subsp. *michiganensis*. *Phytopathology* **92**: 504–510.

Lee Y. 1964*a*. Taxonomic studies of the genus *Miscanthus*: relationships among the section, subsection, and species, part 1. *Journal of Japanese Botany* **39**: 196–205.

Lee Y. 1964*b*. Taxonomic studies on the genus *Miscanthus*: relationships among the section, subsection, and species, part 2: enumeration of species and varieties. *Journal of Japanese Botany* **39**: 257–265.

Liu K, Muse SV. 2005. PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* **21**: 2128–2129.

Long SP, Spence AK. 2013. Toward cool C_4 crops. *Annual Review of Plant Biology* **64**: 701–722.

Lu F, Lipka AE, Glaubitz J, Elshire R, *et al*. 2013. Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. *PLoS Genetics* **9**: e1003215.

Ma X-F, Jensen E, Alexandrov N, *et al*. 2012. High resolution genetic mapping by genome sequencing reveals genome duplication and tetraploid genetic structure of the diploid *Miscanthus sinensis*. *PLoS One* **7**: e33821.

Miyabuchi Y, Sugiyama S, Nagaoka Y. 2012. Vegetation and fire history during the last 30,000 years based on phytolith and macroscopic charcoal records in the eastern and western areas of Aso Volcano, Japan. *Quaternary International* **254**: 28–35.

Morris GP, Ramu P, Deshpande SP, *et al*. 2013. Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proceedings of the National Academy of Sciences USA* **110**: 453–458.

Newman M, Ketphanh S, Svengsuksa B, *et al*. 2007. *A checklist of the vascular plants of Lao PDR*. Edinburgh: Royal Botanic Garden of Edinburgh.

Numata M. 1969. Progressive and retrogressive gradient of grassland vegetation measured by degree of succession – Ecological judgement of grassland condition and trend IV. *Vegetatio* **19**: 96–127.

Ohtsuka T, Sakura T, Ohsawa M. 1993. Early herbaceous succession along a topographical gradient on forest clear-felling sites in mountainous terrain, central Japan. *Ecological Research* **8**: 329–340.

Paradis E. 2010. pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* **26**: 419–20.

Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**: 289–290.

Paterson AH, Moore PH, Tew TL. 2013. The gene pool of *Saccharum* species and their improvement. In: Paterson AH, ed. *Genomics of the Saccharinae*. New York: Springer, 43–71.

Peakall R, Smouse PE. 2006. genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**: 288–295.

Pickrell JK, Pritchard JK. 2012. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics* **8**: e1002967.

Poland JA, Brown PJ, Sorrells ME, 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.

Qin Y, Kabir MA, Wang HW, *et al*. 2013. Assessment of genetic diversity and relationships based on RAPD and AFLP analyses in *Miscanthus* genera landraces. *Canadian Journal of Plant Science* **93**: 171–182.

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

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

Ray N, Adams JM. 2001. A GIS-based vegetation map of the world at the last glacial maximum (25,000–15,000 BP). *Internet Archaeology* **11**. http://intarch.ac.uk/journal/issue11/rayadams_toc.html. Last accessed: 5 Nov 2013.

Reimer PJ, Baillie MGL, Bard E, *et al*. 2004. IntCal04 terrestrial radiocarbon age calibration, 0–26 cal Kyr bp. *Radiocarbon* **46**: 1029–1058.

Sacks EJ, Juvik JA, Lin Q, Stewart JR, Yamada T. 2013. The gene pool of *Miscanthus species* and its improvement. In: Paterson AH, ed. *Genomics of the Saccharinae*. New York: Springer, 73–101.

Shimono Y, Kurokawa S, Nishida T, Ikeda H, Futagami N. 2013. Phylogeography based on intraspecific sequence variation in chloroplast DNA of *Miscanthus sinensis* (Poaceae), a native pioneer grass in Japan. *Botany* **91**: 449–456.

Slavov G, Robson P, Jensen E, *et al*. 2013. Contrasting geographic patterns of genetic variation for molecular markers vs. phenotypic traits in the energy grass *Miscanthus sinensis. GCB Bioenergy* **5**: 562–571.

Somerville C, Youngs H, Taylor C, Davis SC, Long SP. 2010. Feedstocks for lignocellulosic biofuels. *Science* **329**: 790–792.

Stewart JR, Toma Y, Fernández FG, Nishiwaki A, Yamada T, Bollero G. 2009. The ecology and agronomy of *Miscanthus sinensis*, a species important to bioenergy crop development, in its native range in Japan: a review. *GCB Bioenergy* **1**: 126–153.

Sun Q, Lin Q, Yi Z-L, Yang Z-R, Zhou F-S. 2010. A taxonomic revision of *Miscanthus sl* (Poaceae) from China. *Botanical Journal of the Linnean Society* **164**: 178–220.

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

Thurber CS, Ma JM, Higgins RH, Brown PJ. 2013. Retrospective genomic analysis of sorghum adaptation to temperate-zone grain production. *Genome Biology* **14**: R68.

Toonen RJ, Hughes S. 2001. Increased throughput for fragment analysis on an ABI PRISM (R) automated sequencer using a membrane comb and STRand software. *Biotechniques* **31**: 1320–1324.

Van Heerwaarden J, Doebley J, Briggs WH, *et al*. 2011. Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proceedings of the National Academy of Sciences USA* **108**: 1088–1092.

Vigouroux Y, Glaubitz JC, Matsuoka Y, Goodman MM, Sánchez GJ, Doebley J. 2008. Population structure and genetic diversity of New World maize races assessed by DNA microsatellites. *American Journal of Botany* **95**: 1240–1253.

Winkler MG, Wang PK. 1993. The late-Quaternary vegetation and climate of China. In: Wright HE, ed. *Global climates since the Last Glacial Maximum*. Minneapolis: University of Minnesota Press, 221–264.

Winter DJ. 2012. MMOD: an R library for the calculation of population differentiation statistics. *Molecular Ecology Resources* **12**: 1158–1160.

Xu W-Z, Zhang X-Q, Huang L-K, Nie G, Wang J-P. 2013. Higher genetic diversity and gene flow in wild populations of *Miscanthus sinensis* in southwest China. *Biochemical Systematics and Ecology* **48**: 174–181.

Zhang QX, Shen YK, Shao RX, *et al*. 2013. Genetic diversity of natural *Miscanthus sinensis* populations in China revealed by ISSR markers. *Biochemical Systematics and Ecology* **48**: 248–256.

Zhang Y, Zalapa JE, Jakubowski AR, *et al*. 2011. Post-glacial evolution of *Panicum virgatum*: centers of diversity and gene pools revealed by SSR markers and cpDNA sequences. *Genetica* **139**: 933–948.

Zhao H, Wang B, He J, Yang J, Pan L, Sun D, Peng J. 2013. Genetic diversity and population structure of *Miscanthus sinensis* germplasm in China. *PLoS One* **8**: e75672.